Jets in hot nuclear matter

Resumming multiple emissions in QCD

Adam Takacs

Thesis for the degree of Philosophiae Doctor (PhD)
University of Bergen, Norway
2023

UNIVERSITY OF BERGEN

Jets in hot nuclear matter

Resumming multiple emissions in QCD

Adam Takacs

Thesis for the degree of Philosophiae Doctor (PhD)
at the University of Bergen

Date of defense: 20.06.2023

Year: 2023

Title: Jets in hot nuclear matter

Name: Adam Takacs

Print: Skipnes Kommunikasjon / University of Bergen

Acknowledgements

When I have read the acknowledgments of previous Ph.D. students, I always wondered why they are so oversaturated with emotions, but look where I'm going now... Leaving home and moving to a new country to become a researcher is not an easy ride. Doing a Ph.D. makes life rather miserable. Starting with a stressful application procedure, and for the lucky ones, getting the opportunity to leave their surroundings behind and move to a new country. Looking back to the beginning, after all these years, the good memories are shining like the morning star in the dark sky. Acquiring the Ph.D. was fun, I'm happy with my decision, and I will recommend it to others.

First of all, I'm grateful to the senior colleagues who surrounded and supported me during my progression. I'm grateful to Konrad Tywoniuk. As a Ph.D. advisor, he invited me to the field and introduced me to a formalism, that is state-of-the-art and relevant in heavy-ion physics. His up-to-date knowledge provided a very broad, and at the same time, detailed overview of our field. He is always supportive and open to developing conversations. I'm also thankful to Gregory Soyez, from whom I learned so much about jet physics. He made it possible to experience pioneering research. I'm also grateful to my previous supervisors Gergely Gábor Barnaföldi, Denes Molnar, and Bence Kocsis, who introduced me to collaborative working and quality research.

I'm thankful to my collaborators and dear friends, Alba, Dani, and Paul, who are excellent researchers and people. They are the reason why I enjoy conferences more than festivals. I'm so grateful that they always have time for me and my—usually picky—questions. Without them, this thesis wouldn't be the same (it would be much thinner). I would also like to thank my friends and coworkers here in Bergen, who brightened my every day both during and after working hours, Alex, Helge, Johanne, Johannes, and Simon.

And last but not least, I would like to thank my mom, Ati, and my grandmother, who always supported me during my education, and also Anne Elin, who made me fall in love with Norway, literally.

This thesis was supported by a Starting Grant from Trond Mohn Foundation (BFS2018REK01), the MCnetITN3 H2020 Marie Curie Initial Training Network, contract 722104, and I'm thankful for the hospitality of Gregory Soyez at the Institut de Physique Théorique (IPhT), and for Leif Lönnblad at Lund University.

Abstract

In high-energy heavy-ion collisions, protons and neutrons melt and form the quark-gluon plasma. The modification of jets, propagating through this deconfined medium, has been extensively studied at the CERN-LHC, and the BNL-RHIC colliders experiments. Over the last decade, extensive knowledge has piled up in the theory of jet modification. This thesis presents a consistent and state-of-the-art perspective of jet modification based on perturbative QCD. Considering recent progress toward a more accurate description of jets in high-energy physics, jet modification in the medium is reviewed by focusing on their all-order perturbative structure and defining the accuracy of quenched jet observables.

Abstrakt på norsk

Under tung-ion-kollisjonar med høg energi smeltar protonar og nøytronar saman og dannar eit kvark-gluon-plasma. Modifiseringa av jeter, som forplantar seg gjennom det avgrensa mediet, har vorte grundig forska på ved CERN-LHC, og BNL-RHIC i kollisjonseksperiment. I løpet av det siste tiåret har man samla opp omfattande kunnskap om teorien om jet-modifikasjonar. Denne avhandlinga presenterer eit konsistent og toppmoderne perspektiv på jet-modifikasjonar basert på perturbasjons-QCD. Tatt i betraktning nylege framsteg mot ei meir nøyaktig forklaring av jeter i høgenergifysikk, vert jet-modifikasjonar i mediet gjennomgått ved å fokusera på deira perturbasjonsstruktur i alle orden og definere grannsemda til observerbare jet-strålar.

List of publications

1. Paul Caucal, Alba Soto-Ontoso, and Adam Takacs, *Dynamical grooming meets LHC data*, JHEP**07**, 020, 2021.

2. Frederic Dreyer, Gregory Soyez, and Adam Takacs, *Quarks and gluons in the Lund plane*, JHEP**08**, 177, 2022.

3. Adam Takacs and Konrad Tywoniuk, *Quenching effects in the cumulative jet spectrum*, JHEP**10**, 038, 2021.

4. Paul Caucal, Alba Soto-Ontoso, and Adam Takacs, *Dynamically groomed jet radius in heavy-ion collisions*, Phys.Rev.D**105**, 114046, 2021.

5. Johannes Hamre Isaksen, Adam Takacs, and Konrad Tywoniuk, *A unified picture of medium-induced radiation*, JHEP**02**, 156, 2023.

Papers that have been published during the Ph.D. but were not included in the dissertation:

6. Adam Takacs, Daniel Pablos, and Konrad Tywoniuk, *Resolving the spacetime structure of jets with medium*, PoS**387**, 161, 2020.

7. Adam Takacs and Denes Molnar, *Suppression of elliptic flow without viscosity*, arXiv:1906.12311, 2019.

Contents

Acknowledgements **i**

Abstract **iii**

Abstrakt på norsk **v**

List of publications **vii**

1 Introduction **1**

2 Resumming multiple emissions in high-energy collisions **5**

2.1 Some notation to start with 5

2.2 Generalities on perturbative calculations 8

2.3 Infrared divergences 11

2.4 Factorization in the soft and collinear limits 14

2.5 Jets and their substructure 17

2.6 Applications 18

2.6.1 Dynamical grooming 18

2.6.2 Quark and gluon-jet classification at NLL accuracy 26

2.7 Conclusions of vacuum physics 29

3 Resumming multiple emissions in heavy-ion collisions 31

3.1 Jet particles in the quark-gluon plasma 32

3.1.1 The Lagrangian of heavy-ion collisions 32

3.1.2 The background field 34

3.1.3 Effective Feynman rules in the medium 36

3.1.4 Momentum broadening in medium: 38

3.1.5 Medium-induced emissions 40

3.2 Infrared divergences in the medium 44

3.3 Factorization in the soft limit in the medium 46

3.4 Medium-induced cascade 47

3.5 Jets and their substructure in the medium 49

3.6 Applications 51

3.6.1 Jet energy-loss and the quenched jet spectrum 51

3.6.2 Dynamical grooming in heavy-ion collisions 56

3.7 Conclusion of medium physics 60

4 Outlook 63

5 Scientific results 75

5.1 Dynamical Grooming meets LHC data 77

5.2 Quarks and gluons in the Lund plane 127

5.3 Quenching effects in the cumulative jet spectrum 169

5.4 Dynamically groomed jet radius in heavy-ion collisions 209

5.5 A unified picture of medium-induced radiation 233

Chapter 1

Introduction

The Large Hadron Collider (LHC) at CERN is currently the most advanced experiment to study the energy frontier of physics. By colliding particles with higher and higher energies, one can investigate smaller and smaller distances inside them, and eventually resolve new particles and interactions.[1] Experiments in the energy frontier taught us about the existence of most particles, starting with the atomic nuclei and their constituents: protons (1917) and neutrons (1932). More recently discovering quarks (1968) and gluons (1979), which build up hadrons such as the proton and neutron themselves. Since the 1970s, the Standard Model (SM) of particle physics has been established, combining electromagnetic, weak, and strong forces in a unique framework. Nowadays, the SM is mastered by the high-energy physics community.

Proton-proton (pp) collisions at the LHC happen at such high energy, it resolves their inner structure: quarks and gluons. Furthermore, due to the high energy (see also Einstein's famous formula $E = mc^2$), hundreds of new particles are created in these collisions. The Standard Model is challenged daily in the LHC experiments and by the whole high-energy community. Researchers try to find flaws between SM predictions and experimental observations. These predictions rely on small couplings (or rare interactions) of the underlying theories. For the weak and electromagnetic interaction, this is the clear case. It is not that obvious for the strong interaction that describes quarks and gluons (partons). Luckily, the strong interaction is asymptotically free, meaning the coupling decreases as one goes to more energetic scatterings. This decrease was observed,

[1] Let us take two clocks. If we smash them together, they will break apart. By studying the outgoing pieces, in principle, one could reconstruct how clocks work. Relating to particle physics, it is easy to see the importance of high energy. It is necessary to break the clocks/particles apart to study their inner structure. One can break the clock into smaller pieces by going higher energies to learn more details. Later we will see that confinement in this picture translates to observing outgoing intact clocks instead of broken pieces.

for example, in high-energy electron-positron (e^-e^+) collisions. One of the consequences of asymptotic freedom is that the protons consist of three independent, point-like particles (quarks) that have been measured in electron-proton (ep) collisions. On the other hand, for lower energy scatterings, the coupling increases, and it makes calculations difficult. In this thesis, we will see how to avoid this problem. The coupling cannot grow forever as we do not observe infinitely many hadrons.

The theory of quarks and gluons is called quantum chromodynamics (QCD), as partons interact through their color charge that is similar to electric charge. Finally, low-energy quarks and gluons are separately not observed. This phenomenon is referred to as color confinement, as low-energy partons are captured in color-charge-neutral groups.

Both SM predictions and experimental analyses rely heavily on jets. Jets are collimated sprays of hadrons appearing in the detectors after high-energy collisions. They are usually a sign of the creation of a colored parton, which emits further ones and undergoes a parton shower. Eventually, these partons transform, due to color confinement, into detectable hadrons (hadronize) much before they reach the detectors. Most collision events at the LHC (or at least those which are studied) contain jets, and therefore their reconstruction and description are essential.

The LHC was built for a better understanding of the electroweak (EW) sector of the SM. Electroweak particles, e.g. H, W, and Z bosons, can decay into quarks and initiate jets. Jets, therefore, can be used to study EW processes. Moreover, the substructure of jets can distinguish between the different jet-initiating processes. For example, weak bosons and b-quarks typically decay into two light quarks and produce two jets often ending up in a single fat jet. The Higgs boson and top quarks produce 3 jets that can end up an even fatter single jet. Jet substructure can distinguish EW decays.

Moreover, the general picture of a scattering that involves beyond the standard model (BSM) physics, is the creation of some new particle that decays into EW bosons and quarks and induces jets.[2] Therefore, discovering a new particle reduces to a classification task, separating signal from background events by analyzing the topology of events and the substructure of jets as it is sketched in Fig. 1.1. The SM background, therefore, has to be described precisely for future discoveries.

Other than protons, the nuclei of heavy ions (Pb, and Au) also collide at LHC and Relativistic Heavy-Ion Collider (RHIC) at BNL. As large nuclei consist of hundreds of protons and neutrons, these collisions are approximately the superposition of hadron collisions resulting in tens of thousands of detected particles. Jets are found in heavy-ion

[2]There are BSM analyses which rather focus on leptonic decays, or missing energy detection and therefore the do not need jets.

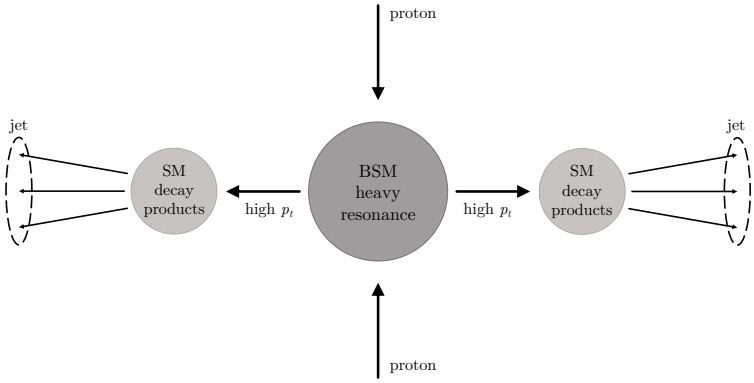


Figure 1.1: The sketch of the creation and detection of a BSM particle using jets. The figure is borrowed from Ref. [1].

collisions too, showing the fundamental presence of strong interaction. They sit on top of a large background as most nucleon scattering produces low-energy (mini) jets. The probability of creating a high-energy jet is rapidly decreasing with the jet energy. The background is comparable with the jet energy, and therefore, it is a highly nontrivial task to analyze jets in this environment. Figure. 1.2 shows a typical event containing two back-to-back jets in *pp* and *PbPb* collisions. The latter has a huge background in which the jets are much less visible. The histograms denote the deposited energy in that particular direction.

A phenomenological separation of particles is usually made to distinguish between jet and background particles. Particles with high energy $\gtrsim 5$ GeV are hard probes, originate from a hard scattering or a jet, and therefore have a perturbative origin. On the other side, bulk particles are soft $\lesssim 5$ GeV and non-perturbative.[3] Surprisingly, the most successful phenomenological description of bulk particles comes from relativistic hydrodynamics. It was concluded therefore, heavy-ion collisions form the quark-gluon plasma (QGP).

How do we know that the created matter is fluid? The angular and momentum correlations among outgoing bulk particles show long-range correlations ($\Delta\phi \sim \frac{\pi}{2}$) referred to as flow [3]. Measurements show a strong correlation between the flow and the centrality (impact parameter) of collisions. The spatial anisotropy in the initial state transforms into an angular (and momentum) correlation of outgoing particles. This correlation is a characteristic feature of fluid motion.[4]

[3] In the collision experiment, the distinction between these two regions is not black or white, and one has to be more rigorous when defining a systematic separation of jets and background (similar to grooming in jet physics).

[4] Take the example of colliding water droplets: in the collision, they form a drop, that oscillates after

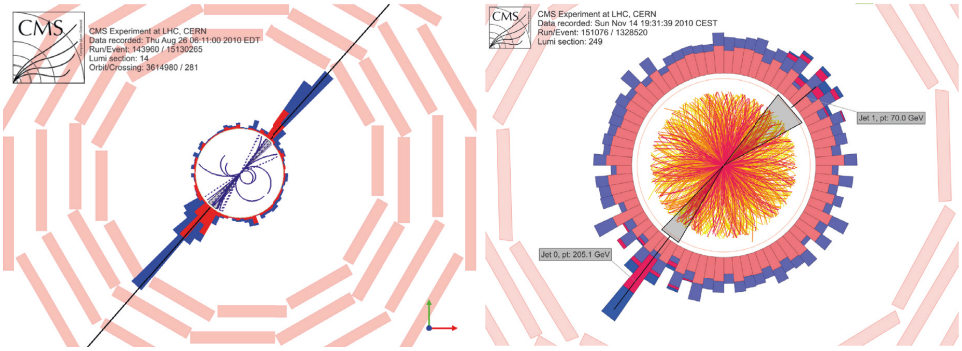


Figure 1.2: A typical collision event containing jets in pp and in $PbPb$ collisions in the CMS experiment, from Ref. [2].

How do we know that the created liquid is the quark-gluon plasma? As fluids are close to local thermal equilibrium, one can study the phase diagram of nuclear matter. Lattice techniques are ideal to evaluate the thermal properties of strongly interacting matter. It has been shown that the phase transition from interacting hadron gas to the deconfined QGP is around 150 MeV. From the hydrodynamic description (or by fitting the bulk's energy spectrum), one can also phenomenologically extract the local temperature of the liquid, which is higher than the transition temperature. Simulations showed that the local temperature of the created plasma is at most 500 MeV. This temperature is roughly the typical non-perturbative scale of QCD, $\Lambda_{\mathrm{QCD}} = 250$ MeV. The QGP created in current experiments is, therefore, strongly coupled.

Jets in heavy-ion collisions are modified as they form inside the quark-gluon plasma. Experiments measure less energetic hadrons and jets in heavy-ion than in pp collisions. This phenomenon is referred to as jet suppression or jet quenching. The modern interpretation of quenching is that jets transfer energy to the QGP, and shift towards lower energy.

In this thesis, the theory of jets will be introduced by focusing on their all-order perturbative structure. We will see that jets and their substructure provide a direct comparison between experimental observables and first principle calculations. We also present the state-of-the-art picture of jet modification in the quark-gluon plasma.

its creation. This oscillation (angular correlation) is the flow, and its magnitude depends on the size of the initial deformation.

Chapter 2

Resumming multiple emissions in high-energy collisions

This chapter introduces the underlying theory of jets, focusing on perturbative QCD at fixed and to all orders. We study the infrared structure of QCD diagrams, and that amplitudes with multiple legs factorize in the infrared limit. This factorization makes possible the all-order treatment of multiparticle production and, thus, the description of jets with high accuracy.

As a practical introduction to jet substructure calculation, by the end of the chapter, dynamically groomed jet observables will be evaluated at fixed and to all orders in perturbation theory for the first time, reviewing essential ingredients of state-of-the-art resummation and their contribution to accuracy. Finally, a novel analysis will be introduced on quark/gluon jet discrimination, lifting the veil on neural network classifiers whose superior performance originates in uncontrolled physics (e.g. uncontrolled logarithms and non-perturbative physics).

2.1 Some notation to start with

Quick summary of QCD

This section quickly summarizes the most fundamental properties of QCD. We introduce quark and gluon fields, the QCD Lagrangian, perturbation theory, running coupling, and renormalization. As these formulas are in many textbooks[1], the reader can skip this section, however, it is necessary to fix notations in the beginning.

[1] A particularly good book in quantum field theory is in Ref. [4].

Let us quickly summarize the most essential ingredient of jet physics, the QCD Lagrangian, which is

$$\mathcal{L}_{\rm QCD} = \sum_f \bar{q}_j^f(x)[i\gamma^\mu D_\mu - m_f]_{jk} q_k^f(x) - \frac{1}{4} F^{\mu\nu a} F^a_{\mu\nu}\,, \tag{2.1}$$

where $q_i^f(x)$ is the quark field, with $f = (u, d, s, \dots)$ flavor, $i = (1, \dots, N_c)$ color, and m_f mass (zero in this thesis). As quarks are Dirac spinors, the γ^μ gamma matrices act on their field, satisfying the algebra $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu} I_4$, where the Minkowski metric is $g^{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$ and I_4 is the 4×4 identity of the spinor fields. The Lagrangian is invariant under the $\mathrm{SU}(N_c)$ gauge transformation giving rise to gluons and interactions. The $A^a_\mu(x)$ is the gluon gauge field $a = (1, \dots, N_c^2 - 1)$, and its field strength is $F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^{abc} A^a_\mu A^b_\nu$, where $f^{abc} t^c = -i[t^a, t^b]$ are $\mathrm{SU}(N_c)$ algebra generators. The covariant derivative results in the coupling between quark and gluon fields that is, $D_\mu = \partial_\mu - igt^a A^a_\mu(x)$, where g is the QCD coupling, and the color matrices $t^a = \lambda^a/2$, with λ^a Gell-Mann matrices $a = (1, \dots, N_c^2 - 1)$ acting on the color fields.

By assuming small coupling ($g \ll 1$), one can define perturbation theory and corresponding Feynman rules. These rules will not be reviewed here, but one can find them in Ref. [4].

The QCD Lagrangian has several interesting features. Ultraviolet (UV) divergences can be removed by renormalizing the parameters m_f and g. The renormalization procedure makes the coupling run (momentum dependent) following the equation

$$Q^2 \frac{\mathrm{d}\alpha_s(Q^2)}{\mathrm{d}Q^2} = -\beta(Q^2)\alpha_s(Q^2)\,, \tag{2.2}$$

where $\alpha_s = \frac{g^2}{4\pi}$. The running coupling at 1-loop is

$$\alpha_s(Q^2) = \frac{\alpha_s}{1 + \beta_0 \alpha_s \ln(Q^2/Q_0^2)}\,, \tag{2.3}$$

where in the $\overline{\rm MS}$ scheme $\beta_0 = \frac{11C_A - 2nf}{12\pi}$, and $\alpha_s = \alpha_s(Q_0^2)$ is an experimental input to fix the value. It does not matter if $Q^2 < Q_0^2$, and usually $\alpha_s(m_Z^2) = 0.1184$ is taken. For $\Lambda^2_{\rm QCD} = Q_0^2 \exp[-1/(\beta_0\alpha_s)]$, one gets the textbook form of the running coupling, $\alpha_s(Q^2) = 1/[\beta_0 \ln(Q^2/\Lambda^2_{\rm QCD})]$. For high energies, $Q \gg \Lambda_{\rm QCD} \approx 0.25$ GeV, the coupling is small $\alpha_s \ll 1$ that is referred to as asymptotic freedom, where $\Lambda_{\rm QCD}$ is the non-perturbative scale of QCD.

The QCD beta function can also be evaluated at higher orders $\beta = \beta_0 + \alpha_s(Q^2)\beta_1 + \dots$. At 2-loop order $\beta_1 = \frac{17C_A^2 - 5C_A n_f - 3C_F n_f}{24\pi^2}$ in the $\overline{\rm MS}$ scheme. One can make an expansion

in α_s, resulting in an implicit equation for $\alpha_s(Q^2)$,

$$\alpha_s(Q^2) = \left[\frac{1}{\alpha_s} + \beta_0 \ln\frac{Q^2}{Q_0^2} - \frac{\beta_1}{\beta_0}\ln\left(\frac{\alpha_s(Q^2)}{\alpha_s}\frac{\beta_0+\beta_1\alpha_s}{\beta_0+\beta_1\alpha_s(Q^2)}\right)\right]^{-1} . \tag{2.4}$$

By evaluating at first-order successive approximation, one gets the approximate formula for 2-loop running coupling

$$\alpha_s(Q^2) = \frac{\alpha_s}{1+\alpha_s\beta_0\ln(Q^2/Q_0^2)}\left[1 - \alpha_s\frac{\beta_1}{\beta_0}\frac{\ln(1+\alpha_s\beta_0\ln(Q^2/Q_0^2))}{1+\alpha_s\beta_0\ln(Q^2/Q_0^2)}\right] , \tag{2.5}$$

where all terms are neglected with corrections $\mathcal{O}(\alpha_s^3)$. The Eq. (2.5) is useful to perform all-order resummation and to understand higher-order running effects. For example, Eq. (2.5) trivially shows that the 2-loop correction is suppressed by additional $\alpha_s \ll 1$.

Thrust as a typical observable

This chapter will use the notation of a typical/general event shape observable. As this might be abstract for those who are not familiar with the notation, we introduce the thrust and will always come back to this as an example. The thrust is defined for a whole event as

$$T = \max_{\vec{n}} \frac{\sum_i |\vec{p}_i\cdot\vec{n}|}{\sum_i |\vec{p}_i|} \approx \max_{\vec{n}} \frac{\sum_i |\vec{p}_i\cdot\vec{n}|}{Q} \tag{2.6}$$

where $p_i^\mu = (E_i, \vec{p}_i)$ are the final state momenta, and in the second step masses are neglected. The maximizing vector $\vec{n} \mapsto \vec{n}_T$ is the thrust axis, that separates the event into left and right hemispheres perpendicular to it. When $T \approx \frac{1}{2}$, the event is spherical, while it is back-to-back-jet-like for $T \approx 1$ and $\vec{n}_T \sim \max_i \vec{p}_i$. As the two-jet limit is the leading order contribution (we will see later in Fig. 2.2), it will be useful to define 1-thrust as

$$v_T \equiv 1 - T \approx 1 - \frac{|\vec{q}_L\cdot\vec{n}_T| + |\vec{q}_R\cdot\vec{n}_T|}{Q} = \frac{q_L^2}{Q} + \frac{q_R^2}{Q} + \mathcal{O}\left(\frac{q_L^2 q_R^2}{Q^4}\right) , \tag{2.7}$$

where we evaluated Eq. (2.6) using the thrust axis $\vec{n}_T$, and the hemisphere momenta are $q_{L/R} = \sum_{j\in S_{L/R}} p_j$. The thrust is, therefore, the sum of hemisphere masses that is the event shape generalization of the jet mass. The last term in the last step is the recoil between the two hemispheres will be neglected. For additional soft or collinear emissions, $v_T \approx 0$, while for hard 3-jet-like corrections, it is $v_T \sim 1$.

We will see the importance of an observable to be soft and collinear safe. One can show that the thrust satisfies the criteria of (i) introducing a soft particle with arbitrary orientation (λk^μ), the thrust will change as $\Delta\tau \sim \lambda$ that vanishes in the $\lambda \to 0$ limit. (ii) introducing a collinear splitting with arbitrary energy ($p_i \to p_{i1} + p_{i2}$), the definition

of $\vec{n}_T$ would only change if p_{i1} and p_{i2} ends in different hemispheres, that is not possible in the $\vec{p}_{i1} \parallel \vec{p}_{i2}$ limit.

2.2 Generalities on perturbative calculations

High-energy proton-proton collisions (pp) are well described in the collinear factorization theorem [1, 5]. This theory relies on asymptotic freedom and separates low-energy (non-perturbative) from high-energy (perturbative) dynamics[2]. Figuratively, the exclusive cross-section of $h_1 + h_2 \to 1, \ldots, n$ process in the collinear factorization looks as

$$\sigma_n = \sum_{a,b} f_a^{h_1} \otimes f_b^{h_2} \otimes \sigma_{ab\to n}(Q^2) + \mathcal{O}\left(\frac{\Lambda^2_{\text{QCD}}}{Q^2}\right) , \tag{2.8}$$

where $f_a^{h_1}$ denotes the a parton's distribution in h_1 hadron, $\sigma_{ab\to n}$ is the partonic cross-section with n final state partons, and $\sum, \otimes$ denote general sums and integrals over internal states. The Q^2 is the scale that appears in the hard scattering. The power of Eq. (2.8) relies on the fact, the matrix element is independent of the incoming hadronic states. Vice versa, the parton distribution functions are non-perturbative but independent of the scattering process (universal). By parametrizing the non-perturbative parton distribution functions in experiments (e.g. in electron-proton collisions (ep)), one can universally use them to predict pp collisions. The last term includes corrections to the factorized picture including non-perturbative effects, and we will discuss them later.[3]

The task left for theorists is to calculate the scattering amplitudes for different processes. This is usually done in the fixed-order expansion scheme of the strong coupling ($\alpha_s \ll 1$), that formally is

$$\sigma_n = \sigma_{n,0} + \alpha_s \sigma_{n,1} + \alpha_s^2 \sigma_{n,2} + \ldots . \tag{2.9}$$

The $\sigma_{n,0}$ denotes the leading order (LO) Born-level cross-section with n final states, and $\sigma_{n,i}$ contains the next-to^i-leading order (N^{i}LO) correction. The N^{i}LO term includes $n+i$ emissions where i of them are integrated out.

[2] A typical LHC collision involves the whole Standard Model. Our goal is to understand jets, thus, we focus on the most essential ingredient.

[3] A similar, complementing argument exists for hadronic final states. There, the partonic cross section is combined with non-perturbative fragmentation functions, which describe the probability of an outgoing parton c ending up in a hadron. Fragmentation functions are similarly universal and can be parametrized experimentally (e.g. in e^-e^+ collision). Fragmentation functions are not used frequently in jet physics, as jets contain multiparton final states. Hadronization corrections to jets are more differential than the relatively simple fragmentation function.

Even though Feynman rules can be completely automated, evaluating the expansion in Eq. (2.9) is highly non-trivial. Different kinds of divergences appear. Loop diagrams exhibit ultra-violet (UV) divergences, that are removable by redefining (renormalizing) the coupling and quark masses. Infrared divergences also appear in both real and virtual diagrams.[4] These divergences only cancel for special observables which are referred to as infrared safe (IRC safe).

Take a general observable v, that is $v \to 0$ in the back-to-back jet limit (e.g. 1-thrust). This observable is defined through some combination of final state momenta $v \equiv V(p_1, \dots, p_n)$. The corresponding distribution is

$$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}v} = \sum_{n=2}^{\infty} \frac{1}{\sigma} \int \mathrm{d}\sigma_n(p_1, \dots, p_n) \cdot \delta(v - V(p_1, \dots, p_n)) \,, \tag{2.10}$$

where σ_n is the exclusive probability with n final states, and the $\int \delta(\dots)$ sets the measurement. The IRC safety condition is given by the premise that v cannot change by adding extra final state particles with infinitesimally small energy or angle. For example, multiplicities are usually not IRC-safe, as adding extra particles change their value. Several codes are available to calculate IRC safe scattering amplitudes up to NNLO accuracy [6, 7].

When there is no hierarchy of scales, the fixed-order expansion scheme works extremely well. In the observable, this is presented as $v \sim 1$. In the case of the thrust, 3-jet-like corrections give a hemisphere mass that is roughly the scale in the hard scattering $v_T \sim \frac{q^2_{L/R}}{Q^2} \approx 1$. On the other hand, when there is a big gap between the exchanged momentum in the hard-scattering and the measured observable large logarithmic subseries appear in the perturbative expansion. In the case of the thrust, soft and collinear emissions contribute at $v_T \sim \frac{q^2_{L/R}}{Q^2} \ll 1$. This is connected with the infrared structure of QCD. Real and virtual diagrams both exhibit IRC divergences, which cancel and leave large infrared logarithms behind in the form of $\ln \frac{1}{v}$. Then $\alpha_s L^2 > 1$, even for $\alpha_s \ll 1$ ($L = \ln \frac{1}{v}$) and the perturbative expansion, presented in Eq. (2.9), breaks down. The cumulative cross-section (physical probability) then can be written in a more general form

$$\begin{aligned} \Sigma(v) &= \int_0^v \mathrm{d}v' \frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}v'} \\ &= 1 + \alpha_s(\sigma_{12}L^2 + \sigma_{11}L + \sigma_{10}) + \alpha_s^2(\sigma_{24}L^4 + \sigma_{23}L^3 + \dots) + \dots \,. \end{aligned} \tag{2.11}$$

The most dominant terms are $\alpha_s^n L^{2n}$, double-logarithmic (DL), while $\alpha_s^n L^{2n-i}$ are the

[4] Infrared divergences include both soft and collinear (IR&C) divergences. Sometimes, soft divergences also referred to as infrared divergence and can be confusing.

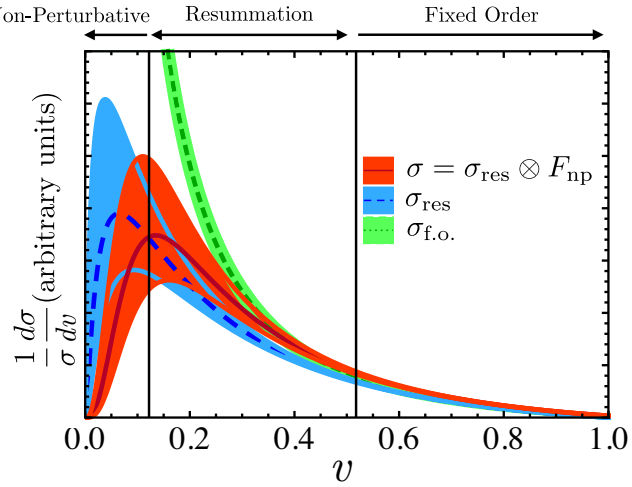


Figure 2.1: The distribution of a typical event shape observable. Fixed order and all logarithmic order resummed formulas constrain different values of the observable. Non-perturbative effects are also illustrated. The figure is from Ref. [11].

N^{i}DL corrections.[5] Keeping all DL terms (i.e. resuming the most dominant logarithmic subseries), one gets a more accurate description of the observable for $v < \exp[-\alpha_s^{1/2}]$, while for bigger v, the fixed order expansion is satisfactory. This all-order treatment of large logarithms is referred to as resummation. In more general, resummation in jet physics is referred to as evaluating subseries of the perturbative expansion.

Similar to fixed-order expansion, theorists have to calculate observables at higher logarithmic accuracy. As mentioned, all-order resummation is needed when emissions are soft or collinear. This typically happens at small values of the observable (but not that small to reach non-perturbative dynamics). Several algorithms and parton showers are available to resum multiple emissions [8–10]. The theoretical argument behind the experimental observation of jets is a bunch of collinear partons detected as hadrons.

The distribution of a typical observable is shown in Fig. 2.1. The fixed order expansion gives a good description for $v \sim 1$, and diverges in the infrared limit ($v \to 0$). The all-order, logarithmic resummed result gives a better, finite description for $v \ll 1$, but above the non-perturbative limit. The final result is the combination of all-order logarithmic resummation, fixed-order matching, and non-perturbative corrections.

The fixed order and resummed expansions denote the same physical phenomenon, how-

[5]Some observables (e.g. thrust) can be written in the exponential form $\sigma(v) = \sigma_0 g_0 \exp[L g_1(\alpha_s L) + g_2(\alpha_s L) + \ldots]$. Therefore historically the naming is defined on the exponent resulting in the leading-logarithmic scheme that is N^{i}LL for g_{i+1}.

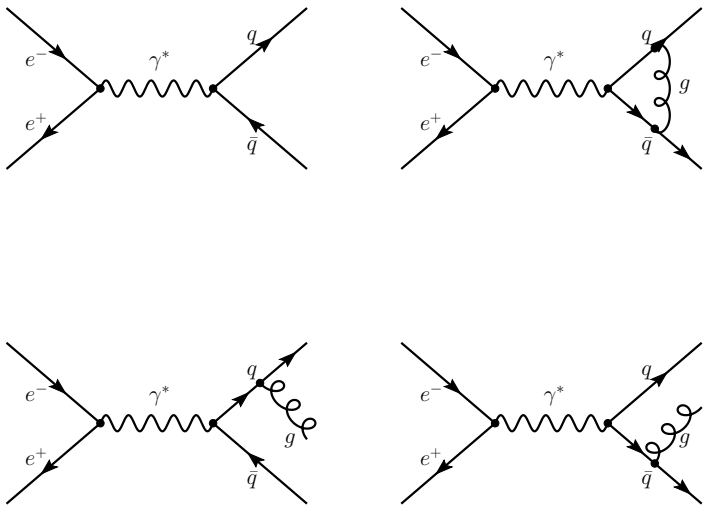


Figure 2.2: Selection of matrix-elements of $e^-e^+ \to q\bar{q}g$ at NLO: LO term (upper left corner), the virtual correction (upper right corner), and real emission diagrams (lower line). Additional virtual diagrams are present when the loop is on the (anti)quark leg.

ever, they are valid in different regions of the phase space, and therefore they are complementary. At the end of this chapter, we will test the general statement we made in this section and apply them in LHC phenomenology for different examples.

2.3 Infrared divergences

Jets appear as a bunch of collimated hadrons measured in the detectors. To understand jets, one has to study scattering amplitudes with multi-particle final states preferably at higher orders of pQCD. As we will see, when final-state particles become soft or collinear with each other, matrix elements may exhibit divergences. This seems to be worrisome as the QCD coupling grows in this infrared region. The QCD perturbation theory, therefore, loses its predictive power and becomes meaningless at small energy (long distances). Can we rely on perturbative calculations of cross-sections? The answer is NO in general. A subset of observables, however, are well-defined in the perturbation theory to all orders (free of IRC divergences) and thus calculable at any order in pQCD. For reviews on perturbative expansions to all-order and IRC safety, see Refs. [1, 8, 12].

Let us recall NLO corrections to the 2-jet production in electron-positron annihilation (e^-e^+) from Ref. [12]. We calculate real and virtual contributions of $e^-e^+ \to q\bar{q}g$ and integrate out the final-state gluon. The total cross-section at LO ($e^-e^+ \to q\bar{q}$) can be written as $\sigma_{q\bar{q},0}(s) = \frac{4\pi\alpha}{3s} N_c \sum_q e_q^2$, where the fine structure constant α, quark electric charge e_q, number of colors N_c, and center-of-mass energy $s = (p_q + p_{\bar{q}})^2$ (first diagram in Fig. 2.2). After summing and averaging over the final and initial state spins, the real contribution is (corresponding to the second line in Fig. 2.2),

$$\overline{\sum_{\text{spin}}} |\mathcal{M}_{\text{real}}|^2 \sim \left(\frac{s_{qg}}{s_{\bar{q}g}} + \frac{s_{\bar{q}g}}{s_{qg}} + \frac{2 s_{q\bar{q}} s_{q\bar{q}g}}{s_{qg} s_{\bar{q}g}} \right) , \tag{2.12}$$

where $s_{ij} = (p_i + p_j)^2$, $s_{ijk} = (p_i + p_j + p_k)^2$. The first two terms are direct emissions (the same emitting parton captures the gluon in the complex conjugate), while the third term is the cross terms. The above amplitude becomes singular when $E_g \to 0$ (soft pole) or when $\theta_{qg}, \theta_{\bar{q}g} \to 0$ (collinear pole). There are two collinear poles as the gluon can be collinear with either p_q or $p_{\bar{q}}$.

Let us rewrite Eq. (2.12) if the emitted gluon is soft ($E_g \to 0$),

$$\overline{\sum_{\text{spin}}} |\mathcal{M}_{\text{real}}|^2 \approx |\mathcal{M}_{q\bar{q}}|^2 4\pi\alpha_s \mu^{2\varepsilon} 2C_F \frac{p_q \cdot p_{\bar{q}}}{(p_q \cdot p_g)(p_{\bar{q}} \cdot p_g)} , \tag{2.13}$$

showing that the soft gluon emission factories from the LO matrix element $\mathcal{M}_{q\bar{q}}$. The momentum ratio in Eq. (2.13) is referred to as the soft antenna emission factor. Factorization is present in the virtual term too. Furthermore, a similar formula can be written in the collinear limit. There the antenna factor is replaced by the DGLAP (or collinear) splitting functions.

By studying the angular structure of Eq. (2.13), one can use the covariant phase space element $[\mathrm{d}p_g] = \frac{\mathrm{d}^3 p_g}{(2\pi)^3 2E_g} = \frac{E_g^2 \mathrm{d}E_g \mathrm{d}\phi \mathrm{d}\cos\theta}{(2\pi)^3 2E_g}$ and integrate out the azimuth of the emitted gluon. The cross-section of the real contribution in the soft and collinear limit is then

$$\sigma^{\text{real}}_{q\bar{q}g} = \sigma_{q\bar{q}} \cdot \frac{\alpha_s}{2\pi} \frac{2C_F}{E_g} \frac{2}{\theta^2_{qg}} \Theta(\theta_{q\bar{q}} - \theta_{qg}) \cdot \mathrm{d}E_g \mathrm{d}\theta^2_{qg} . \tag{2.14}$$

The kinematic factor $\frac{2C_F}{E_g}$ is the soft limit of the DGLAP splitting function, and $\frac{2}{\theta^2_{qg}}$ is the collinear limit of the antenna formula. The last Θ function tells that the soft gluon is emitted at smaller angles than the $q\bar{q}$ dipole, and it originates from the interference of real diagrams and it is a result of color conservation [5, 13]. The factorization and the angular ordering hold for multiple gluon emissions, resulting in the so-called coherent branching algorithm. This algorithm is the basis of resumming multiple emissions [14], and therefore, analytic calculations and numeric parton showers.

When the gluon momentum in Eq. (2.14) is integrated out, one has to evaluate $\int_0^{E_{\max}} \frac{\mathrm{d}E_g}{E_g}$ and $\int_0^{\theta_{\max}} \frac{\mathrm{d}\theta_{qg}^2}{\theta_{qg}^2}$. This is usually the case for inclusive observables. Both of these integrals are logarithmically divergent for real (and virtual) diagrams. When can we get finite results?

Let us calculate the inclusive $e^-e^+ \to q\bar{q} + X$ cross-section as an example, where the gluon momentum is integrated out. The real and virtual contributions can be written as

$$\begin{aligned}
\alpha_s \sigma^{\text{real}}_{q\bar{q}+X,1} &= \sigma_{q\bar{q},0} \frac{\alpha_s}{2\pi} \frac{C_F}{\Gamma(1-\varepsilon)} \left(\frac{4\pi\mu^2}{s}\right)^{\varepsilon} \left(\frac{2}{\varepsilon^2} + \frac{3}{\varepsilon} - \pi^2 + \frac{19}{2} + \mathcal{O}(\varepsilon)\right) , \\
\alpha_s \sigma^{\text{virt}}_{q\bar{q}+X,1} &= \sigma_{q\bar{q},0} \frac{\alpha_s}{2\pi} \frac{C_F}{\Gamma(1-\varepsilon)} \left(\frac{4\pi\mu^2}{s}\right)^{\varepsilon} \left(-\frac{2}{\varepsilon^2} - \frac{3}{\varepsilon} + \pi^2 - 8 + \mathcal{O}(\varepsilon)\right) .
\end{aligned} \tag{2.15}$$

As we saw in Eq. (2.12), integrating out the gluon will give divergences when the gluon energy is small, or it is collinear with any of the $q/\bar{q}$. To obtain a finite result, we use dimensional regularization to regulate these infrared divergences,[6] $d = 4 - 2\varepsilon$, $\varepsilon < 0$. Renormalization also proceeded, with its scale μ in the 1-loop running coupling, so we have already removed the UV divergence. The IRC logarithmic divergences appear as poles in $\varepsilon \to 0$ both in real and virtual terms. These singularities are the same between the real and virtual terms, and they cancel in the inclusive cross-section,

$$\sigma_{q\bar{q}+X} = \sigma_{q\bar{q},0} \left(1 + \alpha_s \frac{3C_F}{4\pi} + \mathcal{O}(\alpha_s^2)\right) . \tag{2.16}$$

Here, Eq. (2.16) corresponds to the inclusive form of the exclusive $n \mapsto q\bar{q}$ cross-section in Eq. (2.9), and inclusive cross-sections are generally absent of infrared divergences. Logarithmic divergences are canceled in a peculiar way, no large logarithms are left behind as is the case for many other observables as we will see later.

Finally, let us see how logarithmic divergences paper in general inclusive observables. Collecting real and virtual terms together from Eq. (2.15) (before integrating the gluon momentum), the NLO distribution of a general v observable can be written as

$$\begin{aligned}
\frac{1}{\sigma_{q\bar{q},1}} \frac{\mathrm{d}\sigma}{\mathrm{d}v} = \frac{1}{2s} &\int [\mathrm{d}p_q][\mathrm{d}p_{\bar{q}}] |\mathcal{M}_{q\bar{q}}|^2 \delta(v - V(p_q, p_{\bar{q}})) \\
&+ \frac{1}{2s} \int [\mathrm{d}p_q][\mathrm{d}p_{\bar{q}}] |\mathcal{M}_{q\bar{q}}|^2 \int [\mathrm{d}p_g] 8\pi\alpha_s C_F \frac{p_q \cdot q_{\bar{q}}}{(p_q \cdot p_g)(p_{\bar{q}} \cdot p_g)} \\
&\times \left[\delta(v - V(p_q, p_{\bar{q}}, p_g)) - \delta(v - V(p_q, p_{\bar{q}}))\right] .
\end{aligned} \tag{2.17}$$

Here, we introduced the measurement function $V(p_1, \ldots, p_i)$ that tells how to calculate

[6] Dimensional regularization with $\varepsilon > 0$ can be used to regulate UV divergences.

v from i final-state particles.[7] The first line describes the $q\bar{q}$ (LO Born) cross-section and the measurement of these 2 final states. The second line uses the factorized limit of the soft gluon emission, and V measures on 3 particles, while the last negative term is from the virtual corrections, and therefore, it acts on the 2-particle phase space. The $\int[\mathrm{d}p]$ phase space integration is used accordingly to how many final states are present. We saw both real and virtual NLO correction terms diverge logarithmically in Eq. (2.15), and therefore v is well defined only if the last line in Eq. (2.17) cancels in the IRC limit. Generally, for a well-defined observable, V_i has to satisfy (where i now labels the number of final state particles)

$$V_{m+1}(\dots, p_i, p_j, \dots) \to V_m(\dots, p_i + p_j, \dots)\,, \quad p_i \,||\, p_j\,, \tag{2.18}$$

$$V_{m+1}(\dots, p_i, \dots) \to V_m(\dots, p_{i-1}, p_{i+1}, \dots)\,, \quad p_i \to 0\,. \tag{2.19}$$

Then the observable v is IRC safe [8]. For example, inclusive observables $V_i \equiv 1$ are IRC safe as we saw σ in Eq. (2.16). Even though, IRC divergences cancel for these observables, large (but finite) logarithms can still remain. As both soft and collinear limits diverge logarithmically, each emission contributes at most 2 large logs. This is the reasoning behind the $\alpha_s^n L^{2n}$ terms in Eq. (2.11).

2.4 Factorization in the soft and collinear limits

In the previous section, we showed the factorization of a $e^-e^+ \to q\bar{q}g$ scattering separately in the soft and collinear limit. In this section, we recall that all scattering amplitude factorizes in the infrared corner of the emission phase space. Following Ref. [15], we show that the dipole factorization formula reproduces the correct factorization in both soft and collinear limits. The dipole factorization is the basis of modern fixed-order pQCD calculations and parton showers.

Take a scattering process with m final states. Corrections to this include an extra final-state particle. The $|\mathcal{M}_{m+1}|^2$ has a singular structure on the momentum p_j, when (i) p_j is soft, and (ii) when p_j is collinear with any of the other final state particles (a generalization of Eq. (2.12)). Surprisingly the singular behavior is independent of the exact scattering process, and it has an opposite sign between real and virtual terms. A factorization formula can be formally written in these two corners of the phase space [15],

$$|\mathcal{M}_{m+1}|^2 \to |\mathcal{M}_m|^2 \otimes \boldsymbol{D}_{ij,k}\,. \tag{2.20}$$

[7] For thrust, $V_T(p_q, p_{\bar{q}}) = \max_{\vec{n}} \frac{|\vec{p}_q\cdot\vec{n}|+|\vec{p}_{\bar{q}}\cdot\vec{n}|}{|\vec{p}_q|+|\vec{p}_{\bar{q}}|}$, and $V_T(p_q, p_{\bar{q}}, p_g) = \max_{\vec{n}} \frac{|\vec{p}_q\cdot\vec{n}|+|\vec{p}_{\bar{q}}\cdot\vec{n}|+|\vec{p}_g\cdot\vec{n}|}{|\vec{p}_q|+|\vec{p}_{\bar{q}}|+|\vec{p}_g|}$.

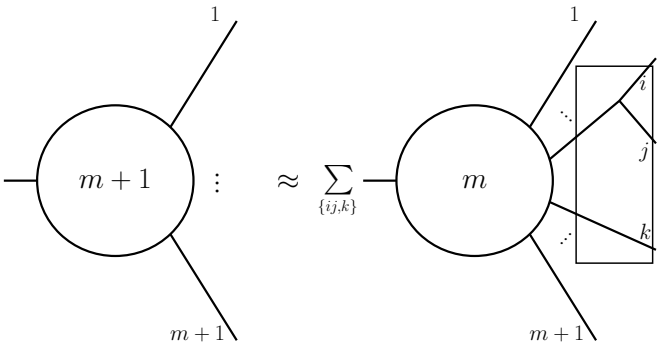


Figure 2.3: The pictorial representation of the dipole factorization formula Eq. (2.20).

Here, $\boldsymbol{D}_{ij,k}$ contains the singular factor, and it only depends on the momenta and quantum numbers of three partons i, j, k. The i and j are the "emitters" while k is the "spectator", which is why Eq. (2.20) is called the dipole factorization formula. A sketch of this factorization is shown in Fig. 2.3, where the box denotes $\boldsymbol{D}_{ij,k}$.

To illustrate the derivation of Eq. (2.20), take the soft limit first. In this case $p_j^\mu = \lambda q$ with q arbitrary 4-momentum, and $\lambda \to 0$. The matrix element square in this limit is

$$|\mathcal{M}_{m+1}|^2 \approx -\frac{1}{\lambda^2} 4\pi\mu^{2\epsilon}\alpha_s \, {}_m\langle 1, \ldots, \not{j}, \ldots, m+1 | [\boldsymbol{J}^\mu(q)]^\dagger \boldsymbol{J}_\mu(q) | 1, \ldots, \not{j}, \ldots, m+1 \rangle_m \,. \tag{2.21}$$

The bra-ket denotes the projection onto the final states where p_j was removed. The renormalization scale μ comes from the renormalization of the coupling in the $\overline{\text{MS}}$ scheme. The formula with dimensional regulator ($d = 4 - 2\epsilon$). The eikonal current is $\boldsymbol{J}_\mu(q) = \sum_{i=1}^m \boldsymbol{T}_i \frac{p_i^\mu}{p_i q}$, where $\boldsymbol{T}_i$ is the color charge operator. The soft singularity represents in the λ^{-2} pole that is the generalization of the one we found in Eq. (2.15). The matrix element $|\mathcal{M}_{m+1}|^2$ decouples from $|\mathcal{M}_m|^2 = {}_m\langle \ldots || \ldots \rangle_m$. The square of the eikonal currents can be written as $[\boldsymbol{J}^\mu(q)]^\dagger \boldsymbol{J}_\mu(q) = \sum_{k,i} \boldsymbol{T}_k \cdot \boldsymbol{T}_i \frac{p_k p_i}{(p_k q)(p_i q)}$. The evaluation of color states involves non-diagonal components, therefore, the factorization is not complete. Additional singularity appears in the collinear limit if $p_k || q$ or $p_i || q$. To represent all poles, we rewrite it in the following way

$$|\mathcal{M}_{m+1}|^2 \approx -\frac{1}{\lambda^2} 8\pi\mu^{2\epsilon}\alpha_s \sum_i \frac{1}{p_i q} \sum_{k \neq i} {}_m\langle 1, \ldots, m+1 | \boldsymbol{T}_k \cdot \boldsymbol{T}_i \frac{p_k \cdot p_i}{(p_i + p_k) q} | 1, \ldots, m+1 \rangle_m \,. \tag{2.22}$$

Now both soft λ^{-2}, and collinear $1/(p_i q)$ singularities are shown at the same time. The dipole structure also emerges: the radiated soft gluon is q, while the direction of the collinear singularity denotes the emitter p_i. The spectator p_k accounts for the color correlation. The dipole structure is pictorially represented in Fig. 2.3.

The factorization in the collinear limit works in a similar way, however, one has to define the collinear direction with care. We use the so-called Sudakov momentum decomposition. In this representation p_i (emitter) and p_j (emission) final-state particles become collinear when $k_\perp \to 0$ and

$$p_i^\mu = z p^\mu + k_\perp^\mu - \frac{k_\perp^2}{z}\frac{n^\mu}{2p\cdot n}\,, \tag{2.23}$$

$$p_j^\mu = (1-z) p^\mu - k_\perp^\mu - \frac{k_\perp^2}{1-z}\frac{n^\mu}{2p\cdot n}\,, \tag{2.24}$$

$2p_i \cdot p_j = \frac{k_\perp^2}{z(1-z)}$, and $p_{i,j}^2 = 0$. The p_j assigns the collinear direction, while n is an auxiliary light-like vector to complete the momentum decomposition. In the collinear limit ($k_\perp \to 0$), therefore,

$$|\mathcal{M}_{m+1}|^2 \to \frac{1}{p_i p_j} 4\pi\mu^{2\epsilon}\alpha_s\, {}_m\langle 1,\ldots,m+1|\hat{P}_{(ij),i}(z,k_\perp;\epsilon)|1,\ldots,m+1\rangle_m\,. \tag{2.25}$$

In the final state, i and j have been combined into one (ij). The $\hat{P}(z,k_\perp;\epsilon)$ is the collinear DGLAP splitting function in dimensional regularization. For example, if it is a $q \to qg$ splitting,

$$\langle s|\hat{P}_{qg}(z,k_\perp;\varepsilon)|s'\rangle = \delta_{ss'} C_F \left[\frac{1+(1-z)^2}{z} - \varepsilon z\right]\,. \tag{2.26}$$

The collinear singularity is presented as $\frac{1}{p_i q} \sim \frac{1}{k_\perp^2}$, while the soft ones sits in the splitting function at $z \to 0, 1$ as $\frac{1}{z}$ and $\frac{1}{1-z}$.

In summary, we presented how multi-parton final-states factorize in the soft or collinear limits separately. Factorization lets us perform Markov chain algorithms to calculate soft and collinear final states with many particles by iterating soft and collinear splittings. The dipole factorization formula is especially useful, as it reproduces both soft and collinear limits, and therefore modern parton showers are all based on some form of the dipole formulae.

2.5 Jets and their substructure

Previous sections introduced fixed-order expansion and all-order logarithmic resummation in the description of event shapes. In this section, we apply this knowledge to jets and their substructure. As jets are represented in the majority of LHC analyses and they serve as a bridge between experimental measurements and fundamental theoretical calculations. For theory reviews see Refs. [1, 11, 16], while for experimental overview see Refs. [17].

Jets are defined through the jet-finder (or -reconstruction) algorithm. This algorithm has to translate into an IRC-safe observable, $V(p_1, \ldots, p_n)$, to result in theoretically meaningful objects [16]. The most used, generalized k_t family of jet finders, is based on the iterative $2 \to 1$ recombination of particles minimizing a distance measure [18]

$$d_{ij} = \min(p_{t,i}^{2p}, p_{t,j}^{2p}) \frac{\Delta R_{ij}^2}{R^2} \,, \tag{2.27}$$

where $p_{t,i}, p_{t,j}$, are the transverse momenta of two jet particles, ΔR_{ij} the distance between them in the azimuth-rapidity plane. Additionally, $d_{iB} = p_{t,i}^{2p}$ is the distance from the "beam". Typical parameters of the jet finders are the cone size of jets R, exponent p, recombination scheme of particles, the cutoff parameter of how many jets one wants in a collision event, or what is the smallest jet momentum p_t one accept in an analysis. One can easily see the IRC safety of Eq. (2.27). In the soft limit $p_{t,j} \to 0$, it p_j combines with the closest particle, without changing its energy. In the collinear limit $\Delta R_{ij} \to 0$, combining the particles immediately.

Many jet observables are borrowed from event shapes, for example, jet shapes, angularities, or energy-energy correlation functions. The jet mass $m_{\text{jet}}^2 = (\sum_{i \in \text{jet}} p_i)^2$ is very similar to the thrust, that sums hemisphere masses. A practical aspect of jets is their typically small cone size $R < 1$; thus, collinear splittings and their factorization are the main interest.

An important difference between event shape observables and jets is that the latter is restricted to a region of phase space (namely a finite cone in the (η, ϕ) plane). This phase space cut makes all jets observable to be non-global [19]: they are sensitive to some emissions while ignoring others. A jet therefore seemingly does not know about emissions happening outside of its cone. This is an incorrect picture of jets and it is generalizable to all orders in perturbation theory. It is more correct to interpret them as a (global) event shape observable that is projected onto several jets. An example of their limitation is the following, a real gluon emission is treated differently if is in/out of the jet cone, while the virtual emission does not know if it is in/out of the jet (remember

Eq. (2.17)). This results in the incomplete cancellation of IRC divergences and the appearance of large non-global logarithms. These corrections cannot be described by iterating soft and collinear splitting inside jets. One can only resum them by going back to the event shape picture, and treating jets accordingly (emissions can appear in the jet cone from outside of it).

Furthermore, most of our formulas involve final state partons. Jets, and event shapes, however, are measured on hadrons. Can we relate parton-level calculations to measuring hadrons? The answer is NO in general. The factorization formula we started with in Eq. (2.8) has corrections. A typical correction, for example, goes as $(\Lambda_{\mathrm{QCD}}/Q)^p$, where $p > 0$ is some power depending on the observable. When Q is big, corrections are subleading, and the factorized, perturbative (parton-level) description is well justified. When we take an event shape observable, a new scale gets introduced that is v itself (e.g. hemisphere mass for the thrust). The perturbative expansion and therefore is well defined if all scales are well separated $\Lambda^2_{\mathrm{QCD}} \ll v \cdot Q^2 \ll Q^2$. Non-perturbative correction arises then $v \cdot Q^2 \sim \Lambda^2_{\mathrm{QCD}}$, meaning either the jet energy (roughly Q) is small, or the observable is deeply in the infrared $v \to 0$. Non-perturbative corrections such as hadronization are suppressed for high-energy jets AND not "too-small" observables. This was illustrated in Fig. 2.1.

Including non-global logarithms and non-perturbative corrections are possible, but it is usually complicated. To reduce corrections, additional grooming techniques will be introduced in the next section.

2.6 Applications

Previous sections quickly summarized fixed and all-order logarithmic resummed perturbative QCD calculations. In this section, two applications of these calculations will be applied to LHC phenomenology. Firstly, in Sec. 2.6.1, we will see a high-order calculation of a novel jet substructure observable. Finally, in Sec. 2.6.2, we will see how all-order perturbative expansion presents and helps with the interpretability of neural network-based jet classifiers.

2.6.1 Dynamical grooming

Although most jet observables are IRC-safe by construction, observables can still be sensitive to non-perturbative effects such as hadronization. Grooming techniques have

been introduced to suppress non-perturbative corrections. Grooming is similar to IRC safety, but it uses a finite, not-too-low scale below which it prevents emissions. It prevents emissions that would be non-perturbative. For example, SoftDrop grooming [20] removes emissions from jets which $z_i < z_{\text{cut}}\theta_i^{\beta}$, where z_i is the splitting energy fraction, and θ_i is the splitting angle compared to the jet axis. For finite grooming parameter z_{cut}, the resulting jet becomes less sensitive to non-perturbative effects. Furthermore, grooming reduces wide-angle emissions in jets which are typically non-global corrections and challenging to handle.

Similar to grooming, tagging emissions is also a powerful tool in jet physics. For example, heavy EW objects (W, Z or Higgs boson) typically decay to quarks and initiate jets. Due to their big boost, quarks get collinear producing a single fat jet. The jet substructure can disentangle these fat jets from single-parton-initiated ones (QCD background). In a tagging setup, SoftDrop, for example, discards emissions below the cut $z_i < z_{\text{cut}}\theta_i^{\beta}$ and returns with the kinematics of the widest angle splitting. This emission is a good proxy for the original hard scale that characterizes the jet-initiating scattering process.

In our work [21], we studied a new observable, and grooming technique called dynamical grooming [22, 23]. It is defined through the procedure

1. Find a jet with anti-k_t algorithm [24];
2. Recluster the jet constituents with the C/A algorithm [25, 26];
3. Go through the reconstruction history and return with the kinematics of the emission that is the "hardest"

$$\kappa^{(a)} = \max_{i,j\in\text{jet}} \frac{p_{ti}p_{tj}}{(p_{ti}+p_{tj})^2}\frac{p_{ti}}{p_{t,\text{jet}}}\left(\frac{\Delta R_{ij}}{R}\right)^a . \tag{2.28}$$

Here, p_{ti} and p_{tj} are the transverse momenta and angle $\theta_{ij} = \Delta R_{ij}/R$ of the splitting $((ij) \to i + j)$. The reclustering in the 2nd step reorganizes the branching history in an angular-ordered way reducing some of the non-global logarithms. For IRC safety, $a > 0$ is necessary. Putting it in words, DyG returns the "hardest" emission inside the jet, where the hardness is $\max z\theta^a$. For example, splitting with the biggest transverse momentum is $\kappa^{(a=1)} = k_t$, or with the biggest mass $\kappa^{(a=2)} = m^2$. In the case of $\kappa^{(-\infty)}$, DyG returns the widest angle emission similar to SoftDrop without any cuts on the phase space. Dynamical grooming is, in fact, a tagging procedure, but it was named to be a groomer as it discards many soft emissions inside the jet. Possible applications of DyG are in EW boson tagging, or in heavy-ion collisions. DyG can tag emissions that formed early vs. late in the medium by setting $a = 2$, which corresponds to inverse formation

time. The DyG can also be used to tag splittings that are perturbative ($\kappa^{(a)} \sim Q$), medium-involved ($\kappa^{(a)} \sim Q_{\rm med}$), or non-perturbative ($\kappa^{(a)} \sim \Lambda_{\rm QCD}$).

To illustrate the all-order structure of DyG observables, we recall its form in the soft and collinear (double-logarithmic) limit. In this limit, Eq. (2.28) simplifies to $\max_i z_i\theta_i^a$, where we neglected the degradation of the leading particle's energy, and the relative angle is measured to the jet axis. At leading order, there is, at most, one emission inside the jet that is trivially the hardest

$$\begin{aligned}\frac{1}{\sigma}\frac{\mathrm{d}\sigma^{\rm LO}}{\mathrm{d}v_g}\bigg|_a &= \int_0^1 \mathrm{d}\kappa^{(a)}\ \frac{1}{\sigma}\frac{\mathrm{d}\sigma^{\rm LO}}{\mathrm{d}v_g}\bigg|_{\kappa^{(a)}} \\ &= \int_0^1 \mathrm{d}\kappa \int \mathrm{d}\theta\mathrm{d}z\widetilde{P}_i(z,\theta)\left[\delta(\kappa^{(a)}-z\theta^a)\delta(v_g-V(z,\theta))-\delta(v_g)\right]. \end{aligned} \tag{2.29}$$

In the first line, we made it explicit that the tagged emission is with $\kappa^{(a)} = z\theta^a$, and then integrated over the hardness. The second line uses the LO probability of the splitting $i \to i+g$ and the measurement function. The observable v_g can be for example the angle ($\delta(\theta_g-\theta)$) or the transverse momentum ($\delta(k_{t,g}-z\theta)$) of the hardest splitting. The last term with $-\delta(v_g)$ is the virtual contribution to the zero-bin as it has no emission inside the jet. At higher orders in the IRC limit, the n particle final state can be written as independent angular-ordered emissions. This is the so-called coherent branching algorithm (see Sec. 2.3). It results in the distribution of $\kappa^{(a)}$ hardest splittings

$$\begin{aligned}\frac{1}{\sigma}\frac{\mathrm{d}\sigma^{\rm DLA}}{\mathrm{d}v_g}\bigg|_{\kappa^{(a)}} &= \sum_{n=0}^{\infty}\frac{1}{n!}\prod_{m=1}^{n}\int \mathrm{d}\theta_m\mathrm{d}z_m\widetilde{P}_i(z_m,\theta_m)\Theta(\theta_m>\theta_{m+1}) \\ &\quad\cdot\left[\delta(\kappa^{(a)}-\max_m(z_m\theta_m^a))\delta(v_g-V(\{z_1,\theta_1\},\ldots,\{z_n,\theta_n\}))-\delta(v_g)\right]^m. \end{aligned} \tag{2.30}$$

The first line represents n independent, angular-ordered gluon emissions, where $n!$ is necessary as gluons are indistinguishable. The second line contains the DyG criterium and the measurement function. The last negative terms are again the virtual ones.[8] The virtual corrections can be resummed trivially, resulting $\exp[-\int \mathrm{d}\theta\mathrm{d}z\widetilde{P}(z,\theta)]$.[9] Let us denote the hardest emission with (z_g,θ_g). The deltas simplify $\delta(\kappa^{(a)}-z_g\theta_g^a)\delta(v_g-V(z_g,\theta_g))$. Therefore, the deltas became independent of all integrals $\int \mathrm{d}\theta_m\mathrm{d}z_m$, and therefore all

[8] Equation (2.30) is somewhat symbolic in the interpretation of the δ^n function, as any of the n emissions can be real or virtual. Equation (2.30) accounts for all of their combinations. Our derivation is complete, however, describing an infinite sum of diagrams in a two-line expression is challenging.

[9] Recall the definition of exponential functions

$$e^{A+B} = \sum_n \frac{1}{n!}(A+B)^n \equiv \sum_n \frac{1}{n!}A^n \cdot \sum_m \frac{1}{m!}B^m = e^A\cdot e^B. \tag{2.31}$$

$m \neq g$ can be resummed trivially. This results in $\exp[\int \mathrm{d}\theta \mathrm{d}z \widetilde{P}(z,\theta)\Theta(\kappa^{(a)} - z\theta^a)]$, where we kept in mind that these emissions are softer than $\kappa^{(a)}$. The two exponential factors can be combined, resulting

$$\begin{aligned}\frac{1}{\sigma}\frac{\mathrm{d}\sigma^{\mathrm{DLA}}}{\mathrm{d}v_g}\bigg|_{\kappa^{(a)}} &= \int \mathrm{d}\theta_g \mathrm{d}z_g \widetilde{P}(z_g,\theta_g)\delta(\kappa^{(a)} - z_g\theta_g^a)\delta(v_g - V(z_g,\theta_g))e^{-\int \mathrm{d}\theta \mathrm{d}z \widetilde{P}(z,\theta)\Theta(z\theta^a - \kappa^{(a)})} \\ &\quad + \delta(v_g)e^{-\int \mathrm{d}\theta \mathrm{d}z \widetilde{P}(z,\theta)}\,. \end{aligned} \tag{2.32}$$

Here, the second line comes again from virtual terms exclusively, representing the no emission inside the jet probability. This term only contributes to normalization. As we will self-normalize, we omit this term from now on. Finally, let us take the hardness integral $\int \mathrm{d}\kappa^{(a)}$, and simplify the notation by introducing the Sudakov factor. The resulting cross-section is the starting point in Ref. [21]

$$\begin{aligned}\frac{1}{\sigma}\frac{\mathrm{d}\sigma^{\mathrm{DLA}}}{\mathrm{d}v_g}\bigg|_a &= \int_0^1 \mathrm{d}\kappa^{(a)}\, \frac{1}{\sigma}\frac{\mathrm{d}\sigma^{\mathrm{DLA}}}{\mathrm{d}v_g}\bigg|_{\kappa^{(a)}} \\ &= \int_0^1 \mathrm{d}\kappa^{(a)} \int_0^1 \mathrm{d}\theta_g \int_0^1 \mathrm{d}z_g \widetilde{P}_i(z_g,\theta_g)\Delta_i(z_g\theta_g^a)\delta(\kappa^{(a)} - z_g\theta_g^a)\delta(v_g - V(z_g,\theta_g))\,, \end{aligned} \tag{2.33}$$

where the LO splitting probability and the no-emission Sudakov factor in the soft and collinear limit are

$$\widetilde{P}_i(z,\theta) = \frac{\alpha_s}{\pi}\frac{2C_i}{\theta z}\,, \tag{2.34}$$

$$\ln \Delta_i(\kappa^{(a)}) = -\int_0^1 \mathrm{d}\theta' \int_0^1 \mathrm{d}z' \widetilde{P}_i(z',\theta')\Theta(z'\theta'^a - \kappa^{(a)})\,. \tag{2.35}$$

The Sudakov factor prevents emissions harder than $\kappa^{(a)}$ in the jet. Finally, the observable is, for example, the transverse momentum $V(z,\theta) = z\theta \equiv k_{t,g}$, or angle $V(z,\theta) = \theta \equiv \theta_g$ of the hardest splitting. Illustratively, Eq. (2.33) tags an emission (z,θ) inside the jet, while the Sudakov factor makes sure there is no harder emission than (z,θ). Finally, V measures v_g observable, and all possible hardness is integrated over. This is sketched in Fig. 2.4.

To study the all-order perturbative structure of DyG observables, we use the cumulative cross-section (physical probability), that we introduced in Eq. (2.11). The transverse

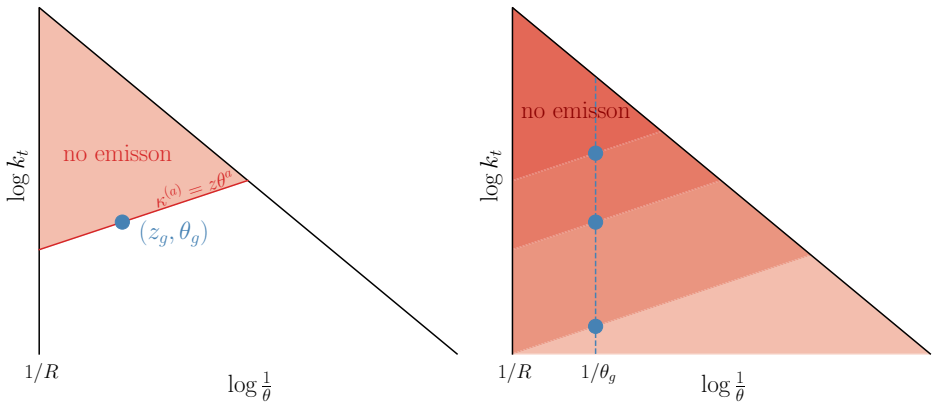


Figure 2.4: The phase space of a dynamically groomed observable. The boundary on the left side is the jet cone ($\theta < R$), while the diagonal border is the maximal splitting energy ($z < 1$). *Left:* the blue bullet represents the tagged hardest splitting ($\widetilde{P}_i(z,\theta)$ in Eq. (2.33)). The red-shaded area has no emissions, and it represents the Sudakov factor ($\Delta_i(\kappa^{(a)})$ in Eq. (2.33)). *Right:* The illustration of integrating to all possible hardness, while θ_g observable kept fixed ($\int \mathrm{d}z\mathrm{d}\theta\delta(v_g - \theta_g)$ in Eq. (2.33)).

momentum of the a-hardest splitting is

$$\begin{aligned}\Sigma(k_{t,g}|a) = &\int_0^{k_{t,g}} \mathrm{d}k'_{t,g} \frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}k'_{t,g}} \\ &\frac{1}{a-1}\left[ae^{\frac{\bar{\alpha}L^2}{a}} - e^{-a\bar{\alpha}L^2} - \sqrt{\pi a\bar{\alpha}L^2}\left[\mathrm{erf}\left(-\sqrt{\frac{\bar{\alpha}L^2}{a}}\right) - \mathrm{erf}\left(-\sqrt{a\bar{\alpha}L^2}\right)\right]\right] \\ &\approx 1 - \bar{\alpha}L^2 + \frac{1+a+a^2}{6a}(\bar{\alpha}L^2)^2 + \mathcal{O}(\alpha_s^3 L^6)\,, \end{aligned} \tag{2.36}$$

where $\bar{\alpha} = \frac{\alpha_s C_i}{\pi}$, and $L = \ln\frac{1}{k_{t,g}}$. The result is finite for $a > 0$ (collinear safety). The second line shows that Eq. (2.33) resums $(\alpha_s L^2)^n$ double-logarithms to all order[10]. We expected the DL structure, as the coherent branching algorithm involves soft and collinear emissions. Both limits diverge logarithmically. As DyG is IRC-safe, these divergences cancel between real and virtual diagrams. The observable is left with large logarithms. The IRC limits contribute to one logarithm at every α_s order, and therefore we resumed both logarithms. Similarly, the angle of the a-hardest splitting is

$$\begin{aligned}\Sigma(\theta_g|a) &= e^{-a\bar{\alpha}L^2} - \sqrt{\pi a\bar{\alpha}L^2}\left[\mathrm{erf}\left(-\sqrt{a\bar{\alpha}L^2}\right) + 1\right] \\ &\approx 1 - \sqrt{\pi a\bar{\alpha}L^2} + a\bar{\alpha}L^2 + \mathcal{O}(\alpha_s^2 L^4)\,, \end{aligned} \tag{2.37}$$

where $L = \ln\frac{1}{\theta_g}$, and again the double-logarithmic resummation is presented. Surpris-

[10] As Σ does not exponentiate, the traditional $N^i LL$ counting is not applicable here.

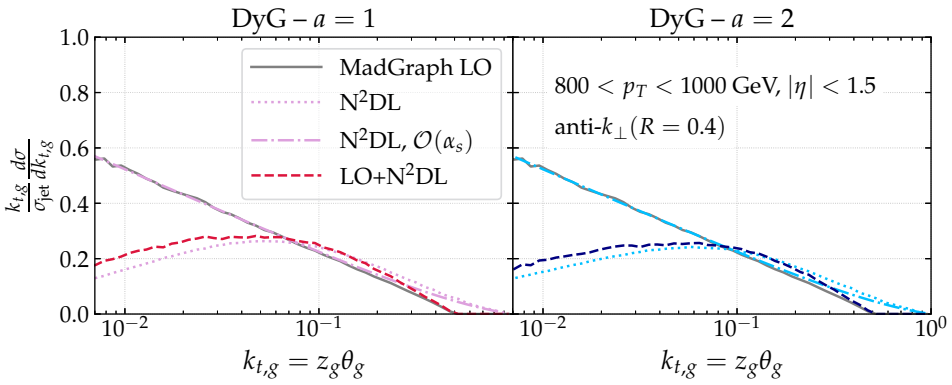


Figure 2.5: The fixed order, N^{2}DL resummed and mathced $k_{t,g}$ distribution for different hardness criteria. Reprinted from Ref. [21].

ingly, we see a non-analytic behavior in the $\sqrt{\alpha_s}$ term. This non-analyticity is known for Sudakov safe observables, which lack IRC safety, and thus they are meaningless at fixed-order. Their all-order form is, however, meaningful. The understanding behind the appearance of the non-analytic terms is not complete yet [27, 28].

As we showed, the coherent branching algorithm is powerful to resum double logarithms. Based on the more general factorization in Sec. 2.4, achieving higher logarithmic accuracy is also possible by relaxing the soft or collinear logarithms. For event shapes, this is referred to as the CAESAR and ARES algorithms [8, 9] that can be applied for jets too. Let us recall the N^{p}DL resummation scheme from Eq. (2.11). If our targeted accuracy is N^{2}DL ($\alpha_s^n L^{2n-2}$), one needs the following additional ingredients in the coherent branching algorithm:

- **Kinematics:** use the full definition of Eq. (2.28).
- **Double-logarithmic approximation:** using 1-loop splitting functions in the soft-collinear limit with fixed coupling resums DL terms $\alpha_s^n L^{2n}$, for $n \geq 1$.
- **Hard-collinear splitting:** releasing the soft limit in the 1-loop splitting function by $\delta\widetilde{P}_i^{\rm hc}(z,\theta) = \widetilde{P}_i(z > e^{-B_i}, \theta)$, with $B_q = \frac{2}{C_F}$, $B_g = \frac{11}{12} - n_f \frac{T_r}{3C_A}$. It gives corrections as $\alpha_s^n L^{2n-1}$ for $n \geq 1$.
- **1-loop running coupling:** the 1-loop running coupling was introduced in Eq. (2.3). The correct scale is the transverse momentum of the splitting $Q = z\theta p_{t,\rm jet}$. It gives corrections as $\alpha_s^n L^{2n-1}$ for $n \geq 2$.
- **2-loop running coupling:** the 2-loop running coupling was introduced in Eq. (2.5). It gives corrections as $\alpha_s^n L^{2n-2}$ for $n \geq 1$.

- **2-loop splitting function:** the soft limit of the 2-loop splitting function can be included as an effective coupling (CMW scheme [14]), $\delta\alpha_s^{\rm CMW} = \alpha_s^{2\ell} - \frac{K}{2\pi}(\alpha_s^{1\ell})^2$, $K = (\frac{67}{18} - \frac{\pi^2}{6})C_A - \frac{5n_f}{9}$. It gives corrections as $\alpha_s^n L^{2n-2}$ for $n \geq 3$.

- **Non-global soft splittings:** jets are non-global observables (see Sec 2.5). Partons outside of the studied jet cone can influence the jet substructure, for example, a wide-angle soft emission of another jet can pollute ours. These emissions interfere with the jet clustering algorithm resulting in clustering (or boundary) logarithms. Event shape techniques to resum them. It gives corrections as α_s, and $\alpha_s^n L^{2n-2}$ for $n \geq 2$.

- **Matching:** similar to non-global logarithms, one needs the global picture to correctly describe the jet cross-section. We use the dijet cross-section at LO that fixes terms at $\mathcal{O}(\alpha_s)$. For this, we used MadGraph5 [7] and performed $\log R$ matching [29]. An important detail is that Sudakov safe observables (like θ_g) do not have the fixed-order expansion, and their matching is, therefore, opened question. We performed the matching on the double differential $\mathrm{d}^2\sigma/(\mathrm{d}z_g \mathrm{d}\theta_g)$ IRC-safe counterpart [27] that is the joint probability of measuring (z_g, θ_g) simultaneously.

In Fig. 2.5, there are examples for the N^{2}DL resummation of the $k_{t,g}$ distribution for two different hardness criteria: hardest transverse momentum ($a = 1$) and mass ($a = 2$) in the jet. The same observable is also shown at LO using MadGraph, and the effect of matching changing the tail. There is a shift in the endpoint of the $k_{t,g}$ distribution in Fig. 2.5. This comes from the fact that we performed the resummation neglecting the energy degradation $(1 - z)$ term in Eq. (2.28) and corrected it after the matching. This is a standard trick in the matching of observables that are not bounded by 1 ($0 \leq v_g \leq v_{g,\max} < 1$).

Finally, to compare with the ALICE measurement [30], non-perturbative effects have to be considered. For example, at a relatively low $p_{t,\rm jet}^{\rm ch} = 70$ GeV, non-perturbative corrections become significant even for groomed observables. We used Pythia8 [31] and HERWIG7 [32], which both include some modeling of these non-perturbative effects tuned to describe LHC data. We extracted the modification between parton and hadron, and parton and charged hadron distributions and corrected our analytics. The resulting $k_{t,g}$ and θ_g distributions are in Fig. 2.6 for the same two hardness presented in the previous plot. The plots show sizeable non-perturbative (NP) corrections, coming from the relatively low jet momentum. As one goes to higher jet p_t, these corrections vanish. All in all, there is a good agreement between our all-order perturbative QCD calculation and the measured distributions from the ALICE collaboration [30].

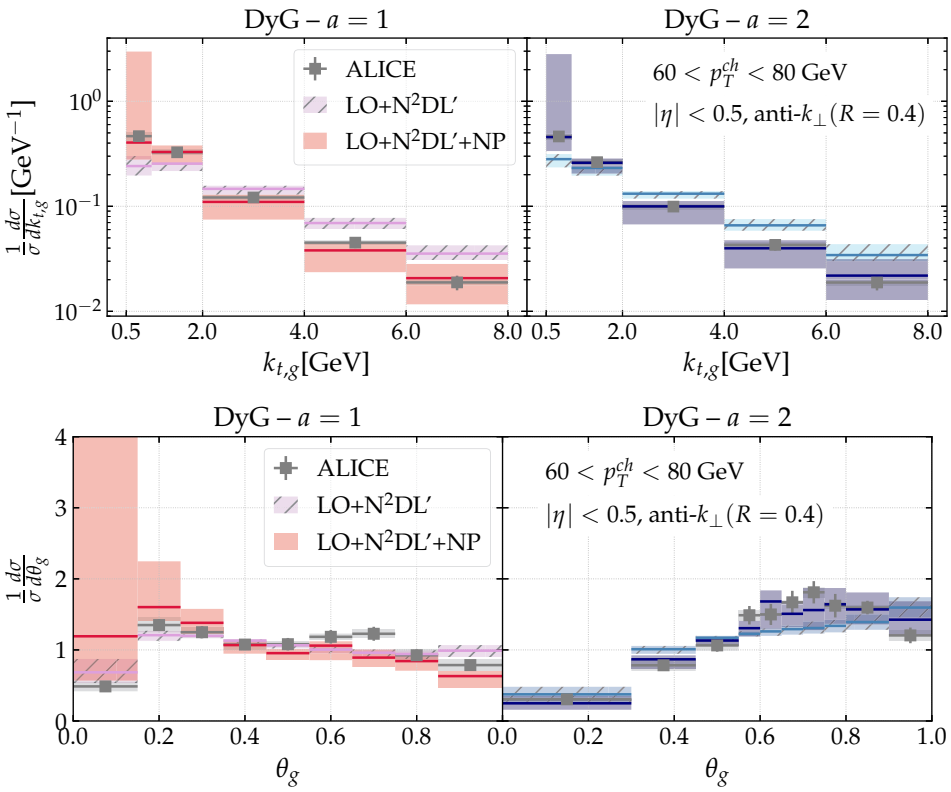


Figure 2.6: The $k_{t,g}$ and θ_g distributions for different hardness criteria. Non-perturbative effects are important in such low jet momentum. Reprinted from Ref. [21].

2.6.2 Quark and gluon-jet classification at NLL accuracy

Distinguishing quark-jets from gluon-jets is an essential task in LHC phenomenology. New particles from beyond the Standard Model (BSM) typically decay to quarks and fragment into jets, while the background is typically gluon jet dominated [1, 11, 17]. A highly accurate description of the QCD background is, therefore, necessary. The naive definition of quark and gluon jets at LO coincides with the flavor of the jet initiating parton. This picture becomes ambiguous at higher orders, as quarks and gluons can transform into each other through successive splittings. Moreover, other jets in the event can influence the assigned flavor. A more rigorous definition of the jet flavor is needed, which, again, goes back to defining jet flavor as an event shape observable [33–35].

Clearly, experiments would like to distinguish quark and gluon jets and understand the background with high precision for BSM searches. Usually, quark/gluon-jet classifiers (or taggers) are based on one or many jet substructure observables (e.g., jet mass, jet shape), which are sensitive to quark and gluon emissions in different regions of the phase space. For example, gluon jets are typically wider than quarks and have more emissions inside them due to their different color factors C_i. Cuts are defined on the observable to make quark or gluon jet enriched samples. There are tremendous theoretical efforts in improving taggers on the perturbative ground. In vain, the simplest neural network (NN) based classifiers outperform any perturbative attempt, and therefore they are widely preferred in LHC analyses [11, 17].

Classifiers are based on the quark/gluon jet tagging likelihood ratio. It is the ratio of probabilities of the given jet being quark or gluon-like. As the likelihood is an exclusive probability (see Eq. (2.9)), we calculated it at NLL accuracy [36]. This likelihood ratio is the ideal discriminant at NLL accuracy. Practically it means that any other discriminant (including well-trained NN ones) will, at best, recover the ideal likelihood's performance in the strict NLL limit. This hypothesis can prove that NN classifiers learn NLL physical features during their training. Moreover, outperforming the ideal discriminant (in the not strict NLL limit) would prove that NN learns beyond NLL features from the training samples, which are usually out of control.

Accessing the NLL limit numerically in parton showers has been explored recently. In Ref. [37], the authors developed the PanScales parton shower that can take numerically the limit $\alpha_s \to 0$, while keeping the product fixed $\alpha_s L = \text{fix}$. For a given value of the observable $L = \ln \frac{1}{v}$. This limit removes any higher-order term than NLL ($\alpha_s^n L^n$) in the all-order expansion (see Eq. (2.11)).

To calculate the quark/gluon-jet likelihood ratio at NLL (in the strongly ordered limit),

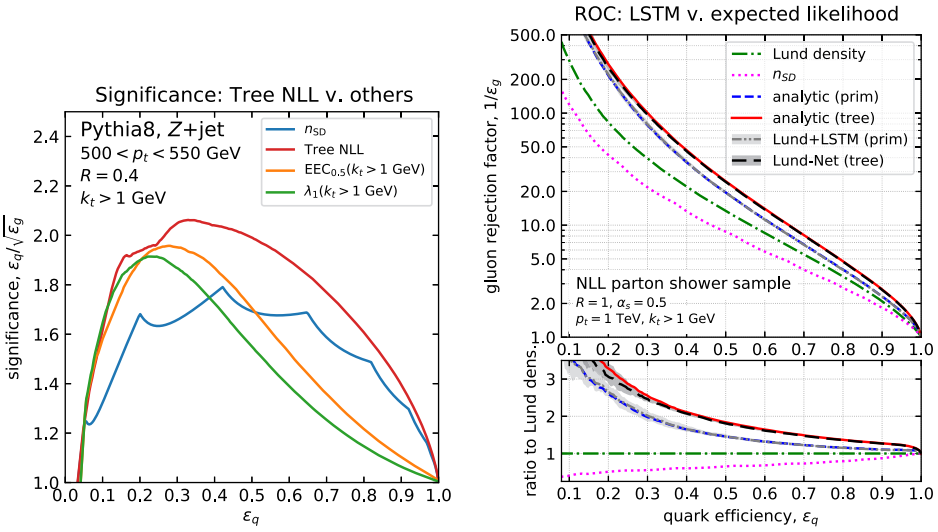


Figure 2.7: *Left:* The performance of our NLL tree discriminant compared to other analytic tagging methods. *Right:* The performance of different analytic and ML classifiers on strictly NLL jet shower samples. Reprinted from Ref. [36].

we used similar ingredients presented in Sec. 2.6.1 also see Ref. [38]. Our discriminant uses the exclusive probability of the jet reclustering tree. The full tree provides enough information to fix all NLL terms[11]. The likelihood ratio is

$$\mathbb{L}_{\text{tree}} = \frac{p_g(\mathcal{L}_{\text{tree}})}{p_q(\mathcal{L}_{\text{tree}})}\,, \tag{2.38}$$

Here $\mathcal{L}_{\text{tree}}$ denotes the binary clustering tree of a jet with the kinematics of each splitting. The $p_i(\mathcal{L}_{\text{tree}})$ is the probability of the jet tree to originate from i parton.[12] The probability is calculated as follows. Branchings are described with DGLAP splitting functions, and between each branching additional splittings are prohibited using Sudakov no-emission probability. Further complications in the algorithm are the degradation of energy and the possibility of quark/gluon flipping during the splitting process.

The left panel in Fig. 2.7 compares the performance of different analytic q/g classifiers, where ε_i is the tagging efficiency, and the higher the significance is better. Our NLL likelihood ratio gives the best signal significance and thus outperforms other analytic methods. In the test, we used Pythia8 [31] generated jet samples. The flavor is defined through the quark recoiled to Z boson.

[11]For example, the SoftDrop multiplicity (n_{SD}), or the primary declustering tree (Lund plane) would not give enough information for a complete NLL description. They give enough for a LL description.

[12]Other likelihood can be constructed by using, for example, the primary tree $\mathbb{L}_{\text{primary}}$, Lund plane density [39], or SoftDrop multiplicity.

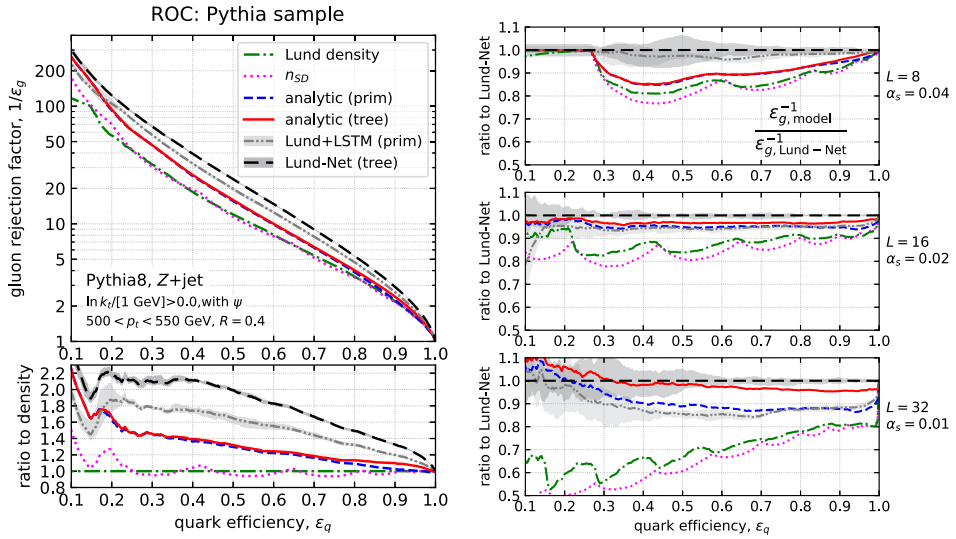


Figure 2.8: *Left:* The performance of different q/g classifiers trained and tested on Pythia8 jets. *Right:* Same as the ratio plot on the left side, for different steps in the numeric NLL limit. Reprinted from Ref. [36].

For the machine-learning tagger, we used the Lund-Net graph neural network (GNN) [40] and LundPlane long, short-term memory network (LSTM) [39]. We extended LundNet for the full clustering tree, matching the NLL requirements similar to the analytic NLL discriminant.[13] Firstly, we trained and tested the classifiers on pure NLL-generated jet samples. This sample matches the ingredients of our analytic calculation. After the successful training, the ML-based classifiers recover their analytic counterparts, as it is shown in the right panel of Fig. 2.7. The axes are the quark and gluon tagging efficiency (ROC) curve, the higher the better. Therefore, we proved our original hypothesis: the exclusive probability is the ideal discriminant at NLL accuracy, and not even NN can outperform it. Using the full tree over primary declustering trivially gives better performance as it has much more knowledge of the system.

In the next step, we retrained the NN classifiers on the more realistic Pythia8 samples. The ML classifiers outperform our analytic discriminates discriminating Pythia8 jets see the left panel in Fig. 2.8. The ordering in the performance remained as expected: n_{SD} < primary tree < full tree. What is the origin of the huge improvement of the ML classifiers?

To understand the gain in the performance, we performed numerous studies in Ref. [36],

[13]We also used the ML classifiers using only primary declustering (GNN and LSTM) and Lund plane density (LSTM).

among others, the state-of-the-art NLL test. In the strict NLL limit, our analytic tree discriminant is ideal. Improvements in the NN performance must come from fitting uncontrolled higher-order terms in the all-order expansion. For example, the DGLAP splitting function includes terms that contribute beyond NLL accuracy. Including the full DGLAP splitting function in the parton shower will not improve the accuracy, as they capture only parts beyond NLL terms. Beyond NLL terms are also present in the 2-loop coupling (see Eq. (2.5)). These uncontrolled higher-order terms and so hadronization were present in the Pythia8 samples during the training procedure that the NN could learn about.

To test this hypothesis, we generated jet samples for which $\alpha_s \to 0$ and $\alpha_s \ln(k_{t,\text{cut}}/Q) =$ fix, where $k_{t,\text{cut}}$ is the minimum transverse momentum we allow in a jet, and Q is the scale of the hard scattering. This NLL limit, therefore, numerically removes uncontrolled, higher-order terms. The right panel in Fig. 2.8 shows ML taggers converging to their ideal analytic counterpart in the numerical NLL limit. ML classifiers, therefore, not only learn NLL physical features, but on top, they also learn about higher-order terms which are uncontrolled. The gain in performance, therefore, must come with the worsening of resilience as we presented in Ref. [36].

2.7 Conclusions of vacuum physics

In this chapter, we introduced the underlying theory of jets, focusing on perturbative QCD at fixed, and to all orders. We presented the importance of soft and collinear emissions. Scattering amplitudes factorize in the soft and collinear limit making possible the all-order treatment of large logarithms and thus, the description of jets.

As an introduction to practical jet calculations, we performed fixed, and all-order resummation of dynamically groomed jet observables. We introduced different ingredients to achieve high accuracy that we applied to LHC phenomenology by performing predictions to ALICE measurements. Section 2.6.1 serves as an overview of methods and tools necessary for producing state-of-the-art analytics for most jet substructure observables. The practical knowledge that is accumulated here will benefit us when we turn to jet modification in heavy-ion collisions.

Finally, in Sec. 2.6.2, we formulated a quark-gluon jet classification task in perturbative QCD. We also introduced the corresponding Machine Learning framework. We showed that, in the strict NLL accuracy limit, the pQCD-based classifier sets an upper limit on the performance. A well-trained neural network can recover this performance. Using

the novel NLL accuracy limit of parton showers, we showed that any gain in the ML classifier's performance originates from fitting model-dependent and uncontrolled physics implemented in event generators. Section 2.6.2 serves as a benchmark for testing ML algorithms and defining the accuracy of NN classifiers.

Chapter 3

Resumming multiple emissions in heavy-ion collisions

In this chapter, we review the current status of jet modification in the quark-gluon plasma. We introduce the corresponding Lagrangian and show medium effects reduce to external currents. We list phenomenological models describing these background currents. These new currents modify the common perturbative QCD by introducing new "in-medium" Feynman rules by including scatterings with the background constituents.

We evaluate the in-medium Feynman diagrams following the footsteps of Ch. 2. We study the non-trivial infrared structure of medium-induced branchings and show the importance of their all-order treatment (or resummation). We review the factorization of multiple emissions in the medium and medium-induced cascades. We also overview the interference effects between vacuum and medium emissions, introducing the factorized picture of in-medium jet evolution. In this chapter, we review current developments and their possible impact on the factorized picture. We sketch the all-order expansion structure of quenched jet observables for the first time.

Finally, as applications of the novel jet quenching formalism, we will calculate jet observables, measured in heavy-ion collisions. First, we focus on the energy loss of jets, describing the nuclear modification factor of jets. Then we put the factorized picture of jet quenching to the test by promoting and calculating dynamically groomed jet substructure to test color coherence in the quark-gluon plasma. Moreover, during these calculations, we illustrate how the all-order resummation structure defined accuracy for quenched jet observables.

3.1 Jet particles in the quark-gluon plasma

In the previous chapter, we saw that when the scale of the hard scattering is large compared to the non-perturbative scale ($Q_{\text{hard}} \gg Q_{\text{np}}$), perturbation theory is applicable as the coupling is small $\alpha_s(Q_{\text{hard}}) \ll 1$. Furthermore, in this wide separation of scales, there is a big phase space for soft and collinear emissions that has to be resumed. This resummation leads to multi-parton production and the experimental observation of jets. In heavy-ion collisions, at least one further scale Q_{med} presents that characterizes the interaction among the plasma constituents. For high-energy jets, $Q_{\text{hard}} \gg Q_{\text{med}}$, and therefore the jet-medium coupling is expected to be suppressed by powers of $Q_{\text{med}}/Q_{\text{hard}}$ similar to the correction in Eq. (2.8). We introduce the theory corresponding to this separation in the following.[1]

3.1.1 The Lagrangian of heavy-ion collisions

As both jets and the quark-gluon plasma consist of colored partons, the QCD Lagrangian (Eq. (2.1)) governs all dynamics. By separating hard (or h) from soft medium (med or m) modes, one can divide the Lagrangian from Eq. (2.1) to

$$\mathcal{L} = \mathcal{L}_h[q_h, A_h] + \mathcal{L}_m[q_m, A_m] + \mathcal{L}_{h-m}[q_h, A_h, q_m, A_m]\,. \tag{3.1}$$

Here we used the separation of the quark and gluon fields, $q = q_h + q_m$, and $A = A_h + A_m$. The $\mathcal{L}_{\text{hard}}$ is equivalent with the QCD Lagrangian Eq. (2.1) that we investigated in detail in Ch. 2 resulting in the description of jets. In jet quenching terminology, this term is referred to as "vacuum physics". The $\mathcal{L}_{\text{med}}$ describes the dynamics of the medium, and it is unknown in general. If one assumes a plasma with very high temperature $T_{\text{med}} \gg \Lambda_{\text{QCD}}$, high-temperature perturbation theory can be used to describe medium dynamics [46, 47]. The high-temperature limit corresponds to the weakly coupled picture of the quark-gluon plasma as the running coupling is small $g(T_{\text{med}}) \ll 1$. Furthermore, heavy-ion collision experiments showed that the temperature of the created plasma is not very high ($T_{\text{med}} \sim \Lambda_{\text{QCD}}$), and therefore alternative descriptions based on strong coupling (e.g., hydrodynamics) is more favored [48, 49].[2] Finally, $\mathcal{L}_{\text{hard}-\text{med}}$ describes the interaction between jet and medium. Even if the medium is non-perturbative, its interaction with the jet still can be perturbative if the exchanged momentum is high

[1] The medium modification of jets should fit well in the framework of soft-collinear effective theory (SCET) [41, 42]. Until now, however, existing formulations are typically focused on a small or dilute medium which has limited applicability in LHC phenomenology [43–45]. The formalism presented here includes this physics.

[2] The description of the medium is not in the scope of this thesis work, even though it is an extremely rich and interesting field. Several exciting works are reviewed in Refs. [50–53].

enough.

As we neglect the dynamics of the background $\mathcal{L}_{\rm med}$ in this thesis, we integrate out the medium fields (q_m, A_m). This is a fundamental and not trivial step in constructing effective field theories.[3] Integrated out the medium fields, the corresponding part of the Lagrangian transforms into two external currents $\mathcal{L}_m + \mathcal{L}_{h-m} = -g\bar{q}_h \langle \hat{J} \rangle_{A_m} q_h - g A_h \langle \hat{J} \rangle_{q_m} A_h$. The Lagrangian no longer depends on the medium fields, as those transformed into external currents, which are complicated operators of hard fields, space time, spin, and color.

The sketched derivation above is the modern picture established by the series of works in Refs. [56, 57] and referred to as BDMPS-Z formalism. It was reformulated later with effective Feynman rules in Ref. [58]. This formalism simplifies the medium fields to an external classical gauge field $A_0(x)$ and performs a series of simplifications on $\langle \hat{J} \rangle_{A_m}$ as we will see. The Lagrangian in this formalism is

$$\mathcal{L} = \mathcal{L}_h + ig \sum_f \bar{q}_h^f(x) \not{A}_0(x) q_h^f(x) - g(\partial^\mu A_h^\nu(x)) A_{0,\mu}(x) A_{h,\nu}(h)\,, \tag{3.2}$$

where the first term is "vacuum" physics. The second and third terms are the interaction between hard partons and the external current, which are taken to be $\langle \hat{J} \rangle_{A_m} \approx -i\gamma^\mu t^a A^a_{0,\mu}(x)$, and $\langle \hat{J} \rangle_{q_m} \approx f^{abc} \partial^\mu A_h^{\nu,a}(x) A^b_{0,\mu}(x) A^c_{h,\nu}(x)$. The external field, $A_0(x)$, is assumed to be independent of the hard field and so it is a space-time dependent density. In general, A_m and A_0 are different, and Eq. (3.2) could include higher order terms, coming from the operator expansion of the currents $\langle \hat{J} \rangle$. These terms are usually neglected, and $A_0 = A_m$ is referred to as the "classical" approximation.[4]

Equation (3.2) is more than a natural formula in the weak coupling picture. Keeping explicitly the current, it is also valid for strong couplings. Let us point out that the coupling g in the interaction term in Eq. (3.2) is the same QCD coupling that presents in $\mathcal{L}_{\rm hard}$. The numerical value of it is different because it is evaluated at typically different energy scales. Furthermore, A_0 scales with a background density that is high. Therefore, the overall expansion parameter in a jet-medium perturbation theory is $\sim gA_0$, as we will see.

[3] The modern formulation of this step relies on density matrices, referred to as open quantum systems. This description has been applied to the quenching of heavy quarks and diluted medium [54, 55].

[4] In words, the classical approximation neglects the partonic dynamics of the external field. In our derivation, this corresponds to integrating out the medium fields. This approximation also neglects higher-order couplings with a medium background. Considering only the tree-level interaction with the external field is well justified for small couplings.

3.1.2 The background field

Before we arrive at the jet-medium Feynman rules, let us discuss first what we know about the classical current of the medium background field $\langle J_a^\mu(x)\rangle_{A_m}$ (or A_0). We will use light-cone (LC) coordinates and the corresponding gauge $A_m^{+,a} = 0$.[5] In the weakly coupled picture, the external current is sourced by the density of independent color charges $\langle J_a^\mu\rangle_{A_m} \approx u^\mu \rho_a(x)$, where ρ is the color charge density and $u^\mu(x)$ is their 4-momentum. The corresponding Euler–Lagrange equation can be solved, simplifying the index structure to $\langle J_a^\mu\rangle_{A_m} \approx \delta^{\mu-}\rho(x^+, \boldsymbol{x}, x^- \approx 0)$ [60]. The background field does not scatter on itself, and therefore, medium and hard fields are loosely connected. Practically, the highly boosted jet particles quickly propagate through the medium and experience mostly its $x^- \approx 0$ state.[6] This can also be seen by considering the Heisenberg uncertainty $x^- \sim \frac{1}{p^+}$, when p^+ is large, x^- becomes small. The remaining part of the Yang–Mills equation is $\boldsymbol{\partial}^2 A_0^{-,a} = -g\rho^a(x^+, \boldsymbol{x})$, resulting

$$A_0^{-,a}(x^+, \boldsymbol{x}) = g \int_{\boldsymbol{k}} \frac{e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}}{\boldsymbol{k}^2} \rho^a(x^+, \boldsymbol{k})\,. \tag{3.3}$$

Physically this means that the density of independent color charges sources the background current/field very similar to the Poisson equation for the electric field in the presence of external electric charge density.

The color charge density $\rho_a(x^+, \boldsymbol{x})$ is a statistical field as it varies in every collision. Assuming independent color charges, we treat ρ as a stochastic field. The density has a vanishing average as the plasma's total color charge is zero. Keeping only the variance of ρ, it will follow Gaussian statistics, $p(\rho) = \exp[-\int_x \rho(x)\cdot\rho(x)/(2n(x))]$, where n is the variance of ρ, or in words, the average of non-vanishing color-charge fluctuation. It results in trivial correlations among the non-interacting charges

$$\langle \rho_a(x^+, \boldsymbol{x})\rho_b(y^+, \boldsymbol{y})\rangle = n\delta_{ab}\delta(x^+ - y^+)\delta(\boldsymbol{x} - \boldsymbol{y})\,. \tag{3.4}$$

Applying Eq. (3.4) to the background fields, one gets

$$\langle A_{0,a}^-(x^+, \boldsymbol{x})A_{0,b}^-(y^+, \boldsymbol{y})\rangle = n\delta_{ab}\delta(x^+ - y^+)\gamma(\boldsymbol{x} - \boldsymbol{y})\,, \tag{3.5}$$

[5] Assuming a parton traversing in the z-direction, the LC coordinates are $(t, \vec{x}) \mapsto (x^+, x^-, \boldsymbol{x})$, where $x^\pm = (t \pm z)/\sqrt{2}$, and $\boldsymbol{x} = (x, y)$, also $x\cdot y = x^+y^- + x^-y^+ - \boldsymbol{x}\cdot\boldsymbol{y}$. In momentum space $p^\pm = (E \pm p_z)/\sqrt{2}$, and $(E, \vec{p}) \mapsto (p^+, p^-, \boldsymbol{p})$. For more about quantum field theories in LC coordinates, see Ref. [59].

[6] The above statement is well justified for static medium and pA collisions in the forward region. To extend this picture for a longitudinally expanding medium, one has to revisit this assumption in the future. As we said, we neglect the dynamics of the medium in this thesis.

where

$$\gamma(\boldsymbol{x}) = g^2 \int_{\boldsymbol{k}} \frac{e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}}{\boldsymbol{k}^4}\,, \tag{3.6}$$

that is the inverse Fourier transform of Coulomb scatterings. Physically, Eq. (3.4) means that the color charge density ρ, consisting of independent charges has no correlations. This current generates a background gauge field A_m with a non-trivial (Gaussian) correlation. Equation (3.6) is not well defined as it diverges in the infrared ($\boldsymbol{k} \to 0$) and needs to be regulated.

Interestingly, in the weak coupling/high-temperature medium assumption, one ends up with the same Eq. (3.5), where the correlations result from thermal fluctuations. In this thermal limit, the infrared divergence is naturally regulated by Debye screening, resulting in [61]

$$N_c n \gamma^{\mathrm{HTL}}(\boldsymbol{k}) = \frac{4\pi \hat{q}_0^{\mathrm{HTL}}}{\boldsymbol{k}^2(\boldsymbol{k}^2 + m_D^2)}\,, \tag{3.7}$$

where the average accumulated transverse momentum square by unit length is $\hat{q}_0^{\mathrm{HTL}} = \frac{N_c g^2 m_D^2 T}{4\pi}$. The Debye mass is a collective effect of color charges polarizing their surrounding in the plasma, and therefore, behaving as collective modes. The effective (or screened) mass of a color-charge is then $m_D \sim gT$, and plasma temperature is T. As we mentioned, the high-temperature assumption is not justified experimentally, and therefore, higher-order corrections can play an important role [62]. Efforts have been made to solve the finite-temperature formalism using non-perturbative techniques [62, 63]. An important recent result is the lattice determination of the scattering potential $\gamma(\boldsymbol{k})$. Based on Ref. [64], the LO and NLO HTL scattering potential correctly describes the Coulomb tail $\sim \boldsymbol{k}^{-4}$, however, non-perturbative effects dominate soft scatterings ($\boldsymbol{k}^2 \lesssim m_D^2$) for realistic temperatures $T_{\mathrm{med}} \approx 200 - 500$ GeV.

Furthermore, in the other weakly coupled picture, the Gyulassy–Wang model [65], scattering centers are described with Yukawa-like interactions that also regularize the infrared

$$N_c n \gamma^{\mathrm{GW}}(\boldsymbol{k}) = \frac{4\pi \hat{q}_0^{\mathrm{GW}}}{(\boldsymbol{k}^2 + \mu^2)^2}\,, \tag{3.8}$$

where $4\pi\hat{q}_0^{\mathrm{GW}} = N_c g^2 n$, and μ are Yukawa-like screening masses. This has the same Coulomb tail at high $\boldsymbol{k}$, and it results in a different soft scattering limit.

Finally, as we mentioned, current operators are more general than their perturbative expansion in the weak coupling limit. How would a non-perturbative, hydrodynamic system be coupled to hard fields? It is tempting, for example, to replace the current with the hydrodynamic current. Higher precision in the gradient expansion would constrain the accuracy of the jet modification then. We entered, however, the territory of the medium dynamics that we certainly do not want to worry about in this thesis.

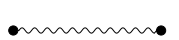
To summarize this section, we sketched an effective field theory in which soft medium modes are integrated out, resulting in background currents. Phenomenology constrains the background currents. Our picture is advantageous, as it is independent of the microscopic description of the background that is either not known, or non-perturbative. Our theory (as any EFTs), therefore, needs phenomenological input.

3.1.3 Effective Feynman rules in the medium

Most quantum field theories in light-cone coordinates are in Ref. [59]. Additionally, in the presence of the background, the propagators are no longer diagonal in momentum space, as multiple interactions with the medium can change the transverse momentum of particles. We also use mixed Fourier representation to keep track of "time" ($t \equiv x^+$): 4-vectors in this mixed space are $(p^+, p^- \mapsto x^+, \boldsymbol{p})$ in momentum space and $(x^+, x^- \mapsto p^+, \boldsymbol{x})$ in spatial space, where the different components connect through the Fourier phase $ip \cdot x = i(p^+x^- + p^-x^+ - \boldsymbol{p}\cdot\boldsymbol{x})$. After considering multiple scatterings in the medium, the dressed propagators read as [58, 66][7]

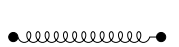
$(x_0^+, p_0^+, \boldsymbol{x}_0)$ $(x^+, p_0^+, \boldsymbol{x})$

(λ_0/s_0) (λ/s)



$$= \frac{1}{2p_0^+}(\boldsymbol{x}|\mathcal{G}_0(x^+, x_0^+; p_0^+)|\boldsymbol{x}_0)\delta^{\lambda\lambda_0}\,, \tag{3.10}$$

$$= \frac{1}{2p_0^+}(\boldsymbol{x}|\mathcal{G}_F^{ji}(x^+, x_0^+; p_0^+)|\boldsymbol{x}_0)\delta^{ss_0}\,, \tag{3.11}$$

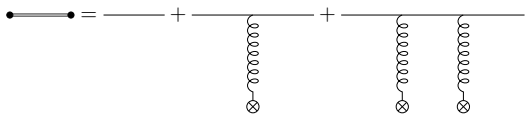


$$= \frac{1}{2p_0^+}(\boldsymbol{x}|\mathcal{G}_A^{ba}(x^+, x_0^+; p_0^+)|\boldsymbol{x}_0)\delta^{\lambda\lambda_0}\,, \tag{3.12}$$

where p_0^+, spin, and helicity are all conserved during propagation between two spacetime points (x_0, x). The free propagator is

$$(\boldsymbol{x}|\mathcal{G}_0(x^+, x_0^+; p_0^+)|\boldsymbol{x}_0) = \Theta(x^+ - x_0^+)\frac{p_0^+}{2\pi i(x^+ - x_0^+)}\exp\left[i\frac{p_0^+}{2}\frac{(\boldsymbol{x}-\boldsymbol{x}_0)^2}{x^+ - x_0^+}\right]\,, \tag{3.13}$$

[7] Propagators are referred to be "dressed" as they include infinitely many scatterings on the background. They could also be "resummed". We avoid this terminology here and keep the word resummation for emissions, and not scatterings. Driagammatically it is



(3.9)

As a simplification, we use single lines and omit the usual double-line notation of dressing.

and the propagators in the path integral representation are

$$(\boldsymbol{x}|\mathcal{G}_{A/F}(x^+,x_0^+;p_0^+)|\boldsymbol{x}_0)=\int\mathcal{D}\boldsymbol{r}\,e^{i\frac{p_0^+}{2}\int_{x_0^+}^{x^+}\mathrm{d}t\,\dot{\boldsymbol{r}}^2}U_{A/F}(x^+,x_0^+;\boldsymbol{r})\,, \tag{3.14}$$

where the Wilson line is

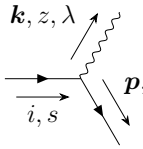
$$U_{A/F}(x^+,x_0^+;\boldsymbol{r})=\mathrm{T}\exp\left[ig\int_{x_0^+}^{x^+}\mathrm{d}t\,A_0^-(t,\boldsymbol{r}(t))\right]. \tag{3.15}$$

Here the Wilson line $U_{A/F}$ resums multiple interactions with the background field through the time-ordered path integral. For gluons (adjoint representation) $A_0^-=-if^{abc}A_{0,c}^-$, and for quarks (fundamental representation) $A_0^-=t_{ij}^aA_{0,a}^-$, while $[t^a,t^b]=if^{abc}t^c$ with f^{abc} structure constans of $SU(3)$. The form of the Wilson line in Eq. (3.15) is remarkable, as it has the same form for quarks and gluons(up to their different color representation. This simple form is the result of the choice that has been made in the currents in Eq. (3.2), and the background correlator in Eq. (3.5). The propagator in mixed momentum space is defined through the Fourier transform

$$(\boldsymbol{p}|\mathcal{G}(x^+,x_0^+;p_0^+)|\boldsymbol{p}_0)=\int_{\boldsymbol{x},\boldsymbol{x}_0}e^{-i\boldsymbol{p}\cdot\boldsymbol{x}+i\boldsymbol{p}_0\cdot\boldsymbol{x}_0}(\boldsymbol{x}|\mathcal{G}(x^+,x_0^+;p_0^+)|\boldsymbol{x}_0)\,, \tag{3.16}$$

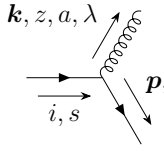
that is for the free propagator is $(\boldsymbol{p}|\mathcal{G}_0(x^+,x_0^+;p_0^+)|\boldsymbol{p}_0)=(2\pi)^2\delta(\boldsymbol{p}-\boldsymbol{p}_0)\exp[-i\frac{\boldsymbol{p}_0^2}{2p_0^+}(x^+-x_0^+)]$.

The vertices after combining with the incoming/outgoing spinor and helicity states are [66]



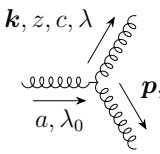
$$=V_{ss'\lambda}^{ij}(\boldsymbol{q},z)$$

$$=\frac{2ie\delta^{ij}}{\sqrt{z(1-z)}}\delta_{ss'}\left[\delta_{\lambda s}+(1-z)\delta_{\lambda-s'}\right]\boldsymbol{q}\cdot\boldsymbol{\varepsilon}_\lambda^*\,, \tag{3.17}$$



$$=V_{ss'\lambda}^{ija}(\boldsymbol{q},z)$$

$$=-2igt_{ij}^a\frac{1}{z\sqrt{1-z}}\delta_{s's}\left[\delta_{\lambda s}+(1-z)\delta_{\lambda-s}\right]\boldsymbol{q}\cdot\boldsymbol{\varepsilon}_\lambda^*\,, \tag{3.18}$$



$$= V^{abc}_{\lambda_0\lambda'\lambda}(\boldsymbol{q}; z)$$

$$= -2igf_{abc}\left[\frac{1}{z}(\boldsymbol{q}\cdot\boldsymbol{\varepsilon}^*_\lambda)\delta_{\lambda_0\lambda'} + \frac{1}{1-z}(\boldsymbol{q}\cdot\boldsymbol{\varepsilon}^*_{\lambda'})\delta_{\lambda_0\lambda} - (\boldsymbol{q}\cdot\boldsymbol{\varepsilon}_{\lambda_0})\delta_{\lambda'\lambda}\right], \tag{3.19}$$

where $\boldsymbol{q} = \boldsymbol{k} - z\boldsymbol{p}$, the transverse-, and the +-momentum is conserved, and all vertices comes with a time integral $\int \mathrm{d}t$. In spatial space, the momentum gets replaced by the spatial derivative, while momentum conservation joins the endpoints, thus $\boldsymbol{q} \mapsto -i\int_{\boldsymbol{w}}(\partial_{\boldsymbol{x}_k} - z\partial_{\boldsymbol{x}_p})_{\boldsymbol{x}_k=\boldsymbol{x}_p=\boldsymbol{w}}$. The spinor and helicity states are $\varepsilon(k)_\lambda/\varepsilon^*(k)_\lambda$ for gluons, and $u_s(p)/\bar{u}_s(p)$ for quarks, and $\bar{v}_s(p)/v_s(p)$ for antiquarks. It is useful to know the sum of gluon polarization $\sum_\lambda \varepsilon^\mu_\lambda(k)\varepsilon^{*,\mu}_\lambda(k) = g^{\mu\nu}$.

The in/outgoing lines are the same as the propagators, without the $\frac{1}{2p_0^+}$ factor. There are no additional spinor or helicity components, as these are included in the vertices as we will see later. Longitudinal gluons are usually neglected, as they only contribute to "instantaneous" 4-particle interactions. These terms were studied in detail in Ref. [67]. The relevant sum is, therefore, in the transverse polarization vectors $\sum_\lambda \varepsilon^i_\lambda(k)\varepsilon^{*,j}_\lambda(k) = \delta^{ij}$.

In summary, the EFT we introduced in the previous sections leads to new, in-medium Feynman rules. There is a new vertex between the hard partons and the background. As the background current has a relatively simple form, dressed hard propagators can easily account for multiple scatterings using Wilson lines. The in-medium Feynman rules then consist of dressed hard propagators and vertices among dressed external lines.

3.1.4 Momentum broadening in medium:

A simple application of the medium Feynman rules presented in Sec. 3.1.3 is the in-medium propagation of a gluon. We follow Ref. [58]. In the vacuum, this would be trivial as the gluon remains unchanged during its propagation. The matrix element in medium is

$$\mathcal{M}^b_\lambda(k) = \int_{\boldsymbol{k}_0} (\boldsymbol{k}|\mathcal{G}^{ba}_A(x^+, x_0^+; k^+)|\boldsymbol{k}_0)\, \boldsymbol{\varepsilon}^*_\lambda \cdot \boldsymbol{J}^a_\lambda(k_0, x_0^+) \tag{3.20}$$

where $\boldsymbol{J}$ is some arbitrary process that creates the hard gluon, and it is multiplied by the outgoing gluon leg. The complex conjugate is

$$\bar{\mathcal{M}}^b_\lambda(k) = \int_{\bar{\boldsymbol{k}}_0} (\bar{\boldsymbol{k}}_0|\mathcal{G}^{\dagger\bar{a}b}_A(x_0^+, x^+; k^+)|\boldsymbol{k})\, \boldsymbol{\varepsilon}_\lambda \cdot \boldsymbol{J}^{*\bar{a}}_\lambda(k_0, x_0^+)\,. \tag{3.21}$$

The amplitude square then is

$$\begin{aligned}\frac{1}{N_c^2-1}\sum_{\lambda,b}|\mathcal{M}^b_\lambda(k)|^2 &= \frac{1}{N_c^2-1}\int_{\boldsymbol{k}_0,\bar{\boldsymbol{k}}_0} \mathrm{Tr}(\boldsymbol{k}|\mathcal{G}_A(x^+, x_0^+; k^+)|\boldsymbol{k}_0)(\bar{\boldsymbol{k}}_0|\mathcal{G}^\dagger_A(x_0^+, x^+; k^+)|\boldsymbol{k}) \\ &\quad \cdot |\boldsymbol{J}_\lambda(k_0, x_0^+)|^2 \\ &= \int_{\boldsymbol{k}_0,\bar{\boldsymbol{k}}_0} (\boldsymbol{k};\boldsymbol{k}|S^{(2)}(x^+, x_0^+; k^+)|\boldsymbol{k}_0;\bar{\boldsymbol{k}}_0)\, |\boldsymbol{J}_\lambda(k_0, x_0^+)|^2\,, \end{aligned} \tag{3.22}$$

where we averaged for the incoming color. We used the summation property of helicity states and the fact that the color is the same in the matrix element and its complex conjugate at every time ($a = \bar{a}$) resulting in the color trace. This comes from the $\delta(x^+ - y^+)$ in the Gaussian correlator (Eq. (3.5)). In the last line, we introduced the 2-point function. One can evaluate the path integrals in the 2-point function in spatial space[8]

$$\begin{aligned}&(\boldsymbol{x};\boldsymbol{x}|S^{(2)}(x^+, x_0^+; k^+)|\boldsymbol{x}_0;\bar{\boldsymbol{x}}_0) \\ &= \int_{\boldsymbol{x}_0}^{\boldsymbol{x}} \mathcal{D}\boldsymbol{r} \int_{\bar{\boldsymbol{x}}_0}^{\boldsymbol{x}} \mathcal{D}\bar{\boldsymbol{r}} \exp\left[\frac{ik^+}{2}\int_{x_0^+}^{x^+} \mathrm{d}t\,(\dot{\boldsymbol{r}}^2 - \dot{\bar{\boldsymbol{r}}}^2) - \frac{N_c n}{2}\gamma(\boldsymbol{r}-\bar{\boldsymbol{r}})\right] \\ &= (\boldsymbol{x}|\mathcal{G}_0(x^+, x_0^+; k^+)|\boldsymbol{x}_0)(\bar{\boldsymbol{x}}_0|\mathcal{G}^\dagger_0(x_0^+, x^+; k^+)|\boldsymbol{x}) \exp\left[-\frac{N_c n}{2}\int_{x_0^+}^{x^+} \mathrm{d}t\,\gamma(\boldsymbol{u}_{\mathrm{cl}}(t))\right]\,, \end{aligned} \tag{3.23}$$

where the classical path is $\boldsymbol{u}_{\mathrm{cl}}(t) = \boldsymbol{x}_0 - \bar{\boldsymbol{x}}_0 - \frac{\boldsymbol{x}_0 - \bar{\boldsymbol{x}}_0}{x^+ - x_0^+}(t - x_0^+)$. After taking the Fourier transform, it is

$$(\boldsymbol{k};\boldsymbol{k}|S^{(2)}(x^+, x_0^+; k^+)|\boldsymbol{k}_0;\bar{\boldsymbol{k}}_0) = (2\pi)^2\delta(\boldsymbol{k}_0 - \bar{\boldsymbol{k}}_0)\mathcal{P}_{\mathrm{broad}}(\boldsymbol{k} - \boldsymbol{k}_0, x^+ - x_0^+)\,, \tag{3.24}$$

$$\mathcal{P}_{\mathrm{broad}}(\Delta\boldsymbol{k}, \Delta t) = \int_{\boldsymbol{u}} e^{-\boldsymbol{u}\cdot\Delta\boldsymbol{p}} \cdot \exp\left[-\frac{N_c n}{2}\gamma(\boldsymbol{u})\Delta t\right]\,. \tag{3.25}$$

Inserting this into Eq. (3.22), the differential cross-section of a single gluon propagating in the medium is simply the broadening

$$\frac{\mathrm{d}\sigma_g}{\mathrm{d}\Omega_k} = \int_{\boldsymbol{k}_0} \mathcal{P}_{\mathrm{broad}}(\boldsymbol{k} - \boldsymbol{k}_0, x^+ - x_0^+) \cdot \frac{\mathrm{d}\sigma_0}{\mathrm{d}\Omega_{k_0}}\,, \tag{3.26}$$

[8]The easiest way to evaluate the path integral is to read off the Lagrangian $\mathcal{L} = \frac{ik^+}{2}(\dot{\boldsymbol{r}}^2 - \dot{\bar{\boldsymbol{r}}}^2) - \frac{N_c n}{2}\gamma(\boldsymbol{r} - \bar{\boldsymbol{r}})$, and solve the corresponding equation of motion. Then, insert it back into the path integral.

where the incoming cross-section is $\mathrm{d}\sigma_0/\mathrm{d}\Omega_{k_0} = |\boldsymbol{J}^a(p_0)|$, and the phase space element is $\mathrm{d}\Omega_k = \frac{1}{(2\pi)^3}\frac{\mathrm{d}k^+\mathrm{d}^2\boldsymbol{k}}{2k^+}$. When the broadening is evaluated, one has to choose a background model for $\gamma(\boldsymbol{u})$ and evaluate the integrals. See Refs. [68] for an overview of the broadening of several medium models.

3.1.5 Medium-induced emissions

We apply the medium Feynman rules to a splitting in the medium. The $q \to qg$ matrix element is

$$\mathcal{M}^{ia}_{s\lambda}(p,k) = \frac{1}{N_c}\int_{\boldsymbol{k}',\boldsymbol{p}',\boldsymbol{p}_0}\int_0^\infty \mathrm{d}t\,\frac{1}{2p_0^+}(\boldsymbol{k}|\mathcal{G}_A^{ab}(t_L,t;zp_0^+)|\boldsymbol{k}') \\ \cdot(\boldsymbol{p}|\mathcal{G}_F^{ii'}(t_L,t;(1-z)p_0^+)|\boldsymbol{p}'-\boldsymbol{k}')V^{i'ib}_{s's\lambda}(\boldsymbol{k}'-z\boldsymbol{p}',z)(\boldsymbol{p}'|\mathcal{G}_F^{i'j}(t,t_0;p_0^+)|\boldsymbol{p}_0)\mathcal{M}^j_{s'}(p_0)\,. \quad (3.27)$$

By going backward, $\mathcal{M}(p_0)$ is the arbitrary creation of the initial quark, then a quark propagator $\frac{1}{2p_0^+}\mathcal{G}_F$, that is multiplied with the vertex V, and outgoing quark and gluon lines $\mathcal{G}_F\cdot\mathcal{G}_A$. All intermediate momenta and indices are summed over. After taking the amplitude square, evaluating multiple path integrals is highly not trivial. The color rotations in the Wilson lines complicate the evaluation, and therefore, we take the large-N_c approximation, leaving behind $\mathcal{O}(1/N_c^2)$ corrections [69]. An interesting property of medium-induced emissions is the lack of collinear singularity. This is clearly different in the vacuum, where the collinear singularity was presented in Sec. 2.3. It is possible to integrate out the transverse momentum of medium-induced emission resulting in the spectrum [56, 70–73] (our notation follows Refs. [74, 75]),

$$\omega\frac{\mathrm{d}I}{\mathrm{d}\omega} = \frac{2\alpha_s C_R}{\omega^2}\mathrm{Re}\int_0^\infty \mathrm{d}t_2\int_0^{t_2}\mathrm{d}t_1\,\partial_{\boldsymbol{x}}\cdot\partial_{\boldsymbol{y}}\big[\mathcal{K}(\boldsymbol{x},t_2;\boldsymbol{y},t_1)-\mathcal{K}_0(\boldsymbol{x},t_2;\boldsymbol{y},t_1)\big]_{\boldsymbol{x}=\boldsymbol{y}=0}\,, \quad (3.28)$$

where ω is the energy of the induced emission (rigorously, it is k^+).[9] Vacuum emissions have been subtracted with the second term, and the 3-point function $\mathcal{K}$ satisfies the 2+1 dimensional Schrödinger equation

$$\left[i\partial_t + \frac{\partial^2_{\boldsymbol{x}}}{2\omega} + iv(\boldsymbol{x},t)\right]\mathcal{K}(\boldsymbol{x},t;\boldsymbol{y},t_0) = i\delta(t-t_0)\delta(\boldsymbol{x}-\boldsymbol{y})\,, \quad (3.29)$$

where the imaginary potential is[10]

$$v(\boldsymbol{x},t) = \int_{\boldsymbol{q}}\sigma(\boldsymbol{q},t)\left(1-e^{i\boldsymbol{q}\cdot\boldsymbol{x}}\right)\,. \quad (3.30)$$

[9] In the soft limit $zp_0^+ \to \omega$, and $(1-z)p_0^+ \to p_0^+ = E$, in Eq. (3.27).

[10] We use the shorthand notation, $\int_p = \int\frac{\mathrm{d}^4p}{(2\pi)^4}$, $\int_{\boldsymbol{p}} = \int\frac{\mathrm{d}^2\boldsymbol{p}}{(2\pi)^2}$, and $\int_{\boldsymbol{x}} = \int\mathrm{d}^2\boldsymbol{x}$.

Here $\sigma(\boldsymbol{q},t) = N_c n(x)\gamma(\boldsymbol{k})$ is proportional to the in-medium elastic scattering cross section, where $n(x)$ is the density of scattering centers.

Equation (3.28) is a leading order result in α_s, but it includes interactions with the medium to all orders in $g^2 n$ which is hidden in Eq. (3.28). It is presented in the soft limit $\omega \ll E$, where E is the emitter's energy (rigorously p_0^+), keeping only the leading number of color terms (neglecting $\mathcal{O}(1/N_c^2)$). It is valid for any slowly varying potential $v(\boldsymbol{x},t)$, but here we focus on a homogeneous medium with fixed length L (rigorously x^+). The jet-medium scatterings are modeled through the GW model $\sigma(\boldsymbol{q}, t < L) = 4\pi\hat{q}_0/(\boldsymbol{k}^2+\mu^2)^2$, where the mean free path among scatterings is inversely proportional with the medium density and $\lambda = (\int_{\boldsymbol{q}} \sigma_{\rm tot})^{-1} = \mu^2/\hat{q}_0$.

Equation (3.28) has been extensively studied in the literature as it is the essential ingredient of any jet-quenching calculation. It is viewed as a master formula, and the available analytic solutions are approximations of this "full" solution in different limits. For example,

- The BDMPS-Z limit [56, 57] approximates the scattering potential with its soft limit $v(\boldsymbol{x}) \approx \frac{1}{4}\hat{q}_0\boldsymbol{x}^2$, allowing to solve the Schrödinger equation (and thus the path integrals) analytically. We refer to this as the harmonic oscillator (HO) approximation. The resulting spectrum recovers Eq. (3.28) in the soft limit for a big enough medium ($\omega \to 0$ and $L \gg \lambda$).

- The GLV-W formula [71, 76] instead of resumming medium interactions, it expands and truncates in the number of scatterings. We refer to this as opacity expansion (OE). It recovers Eq. (3.28) if the medium is small ($L \ll \lambda$), and if the induced emission is hard ($\omega \to \infty$).

More recently, numerical solutions of Eq. (3.28) became available, illustrating the interpolation between the two limiting analytic solutions [77–79]. The numerical analyses showed that the limiting cases are good approximations of the full formula, however, the phase space structure was not entirely understood. These findings motivated a series of recent efforts searching for a better analytic understanding of Eq. (3.28).

- In Refs. [74, 80], the authors constructed a perturbation theory, which combines the all-order HO resummation with a finite number of hard scatterings similar to the OE. They separated the scattering potential to harmonic oscillator plus corrections $v(\boldsymbol{x},t) = v_{\rm HO}(\boldsymbol{x},t) + \delta v(\boldsymbol{x},t)$. The HO part is solved analytically, while δv is treated as a perturbation. We refer to this as improved opacity expansion (IOE).

In Ref. [75], we constructed a framework of different analytical expansion schemes to solve Eq. (3.28) analytically to all orders (in the medium scatterings). By studying Eq. (3.28), we identified the phase space structure illustrated in Fig. 3.1. The emission phase space consists of the energy of the induced emission ω and the length of the medium t. Different regions in Fig. 3.1 show the dominant scattering process that induced the emission. Hard emissions with $\omega > \omega_c(t) = \frac{1}{2}\hat{q}t^2$ are triggered by a single hard scattering on the medium (the green region in the figure). The opacity expansion is a good description of this region. Surprisingly, this regime extends for large medium $L \gg \lambda$ (where $\lambda \sim \mu^2/\hat{q}_0$ is the mean free path of scatterings). Physically, this is possible because the high energy of the emission forces harder scatterings, effectively reducing the probability of multiple scatterings. See also that $\gamma(\boldsymbol{q})$ decays fast for large $\boldsymbol{q}^2$. In the region $\omega_{\text{BH}} = \frac{1}{2}\mu^2\lambda < \omega < \omega_c(t)$ emissions are induced by multiple soft scatterings in the medium, and therefore, the harmonic oscillator scatterings play the dominant physics (red region). These emissions typically have long formation times, and scatterings "add up" to a single coherent scattering during the formation. The IOE showed that corrections from rare, hard scatterings are also important and enhance the value of $\hat{q}_0 \mapsto \hat{q}(\omega)$. Finally, soft emissions $\omega < \omega_{\text{BH}}$ are still triggered by soft scatterings. The number of scattering is very few, even for large medium. This region is referred to as Bethe–Heitler emissions by the QED analog (blue region). The diagrammatic sketch of the induced emission and the corresponding scattering process is in Fig. 3.1. The shaded blob represents formation time.

- In Ref. [75], we constructed an all-order perturbation theory describing very soft emissions for the first time. It is similar to the opacity expansion as it truncates in the number of real scatterings (second term in Eq. (3.30)), while it resums virtual scatterings (first term in Eq. (3.30)) to all orders. Therefore it accounts for the no scattering probability to all orders (similar to the no emission probability Sudakov factor in jet physics). We refer to this formalism as resumed opacity expansion (ROE).

In Fig. 3.2, we compare our analytic description of Eq. (3.28) to the numerical solution from Ref. [78]. The three different panels correspond to three different medium sizes, where the mean free path is $\lambda = 0.06$ fm. The plot range includes very low energies to illustrate the separation between the different regions, even though pQCD should break down below $\lesssim 1$ GeV. Different colors correspond to different phase space regions from Fig. 3.1, while the dotted lines are different analytic limits (first-order opacity (or GLV limit), harmonic oscillator (BDMPS-Z limit), and first-order ROE (Bethe–Heitler limit)). As our framework combines different perturbation expansions, we truncated our formulas at first order in the plots ($N_r = 1$ in ROE, and HO+NHO in the IOE) even

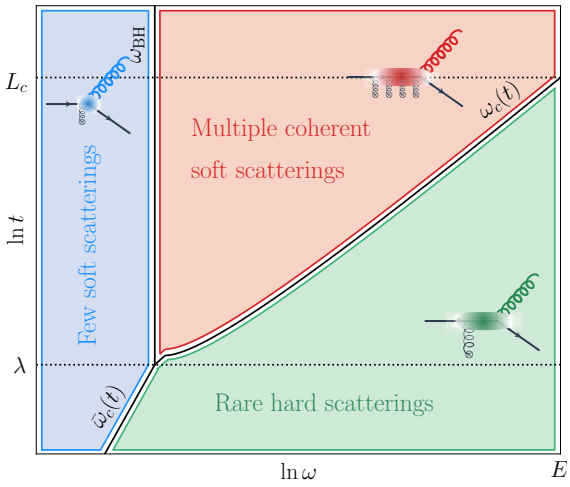


Figure 3.1: The emission phase-space of medium-induced emissions, and the corresponding dominant scattering process that triggers the emission. Reprinted from Ref. [75].

though nothing would prevent us from going to higher orders. This "first order" result already captures most of the full solution, resulting in 10% deviation from the numerical solution on the boundary of the expansion schemes that decreases as one goes to higher orders.

Equation (3.28) includes several approximations that could be relaxed. First, when we integrated the transverse momentum, we neglected the kinematic constraint, which restricts the angle $\theta < \pi$. Finite-z and flavor-changing corrections are available in Ref. [75]. The finite number of color corrections are available in Ref. [69]. Next-to-leading order α_s^2 corrections are available in the HO limit [81] and in the OE [82]. Finite quark mass

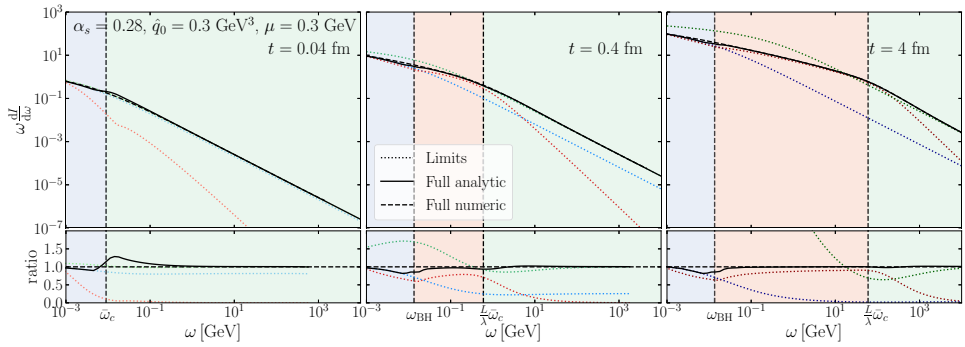


Figure 3.2: The medium-induced emission spectrum Eq. (3.28) for different medium lengths, and using different analytic approximations compared to the numerical solution.

corrections are available in the HO [83] and OE limits [84]. Regarding the medium potential $\gamma(\boldsymbol{k})$, solutions are available beyond the HTL and GW models, including non-perturbative effects [79, 85]. Finally, the dynamical medium was also studied, resulting in longitudinally expanding formulas [86, 87] and so plasma with transverse structure [88–90]. Considering these corrections, the simple physical interpretation presented in Fig. 3.1 has to be studied in the future.

As one can see, many improvements accumulated for medium-induced emission recently, however, these corrections are uncontrolled as no one could define a hierarchy among them. In the next section, we will sketch a hierarchy of importance among these corrections using the knowledge we accessed in Ch. 2.

3.2 Infrared divergences in the medium

Let us first summarize, the limiting behavior of the spectrum Eq. (3.28) in different regions of the phase space[11]

$$\omega\frac{\mathrm{d}I}{\mathrm{d}\omega}\bigg|_{L\gg\lambda} = \begin{cases} 2\bar{\alpha}\frac{L}{\lambda}\ln\left(\frac{\omega_{\mathrm{BH}}}{\omega}\right), & \text{for } \omega \ll \omega_{\mathrm{BH}}\,, \\ \bar{\alpha}\sqrt{\frac{2\omega_c}{\omega}}\,, & \text{for } \omega_{\mathrm{BH}} \ll \omega \ll \omega_c\,, \\ \frac{\pi}{2}\bar{\alpha}\frac{L}{\lambda}\frac{\bar{\omega}_c}{\omega}\,, & \text{for } \omega_c \ll \omega\,, \end{cases} \tag{3.31}$$

where $\bar{\alpha} = \frac{C_i\alpha_s}{\pi}$ $\omega_{\mathrm{BH}} = \frac{1}{2}\mu^2\lambda$, $\omega_c = \frac{1}{2}\hat{q}L^2$ $\bar{\omega}_c = \frac{1}{2}\mu^2 L$, and $\lambda = \frac{\mu^2}{\hat{q}_0}$. The different colors denote different regions in Fig. 3.1. We do not forget that medium-induced emissions lack collinear singularity that made it possible to integrate the transverse momentum of the emitted gluon. Various shapes of the soft divergence are present in Eq. (3.31) and therefore, we expect the necessity of resummation similar to Ch. 2. A simple way to see this is by integrating out ω that is proportional to the number of induced emissions. When this number is large $\gg 1$, corrections from multiple emissions have to be included,

$$\int_{\omega_{\min}}^{\infty}\mathrm{d}\omega\,\frac{\mathrm{d}I}{\mathrm{d}\omega} \approx \bar{\alpha}\frac{L}{\lambda}\ln^2\frac{\omega_{\mathrm{BH}}}{\omega_{\min}} + 2^{\frac{3}{2}}\bar{\alpha}\left(\sqrt{\frac{\omega_c}{\omega_{\mathrm{BH}}}} - \ln\frac{L}{\lambda}\right) + \pi\bar{\alpha}\left(1 - \frac{\hat{q}_0}{\hat{q}}\ln\frac{L}{\lambda}\right), \tag{3.32}$$

[11] As a reminder, there are also virtual emission diagrams in the medium that are usually neglected in most notations. These terms are crucial, however, to preserving unitarity. Luckily, virtual emissions come with the same functional form as real ones with an additional negative sign. For the LO induced spectrum it is $-\delta(\omega)\int\mathrm{d}\omega\frac{\mathrm{d}I}{\mathrm{d}\omega}$.

where we used the corresponding phase space boundaries $\omega_{\rm BH}$ and $\omega_c(L)$. In the soft limit $\omega < E \to \infty$ can be safely taken. Let us study the different contributions separately:[12]

- HO region: the contribution of this region can be rewritten as $\sim \bar{\alpha}(\frac{L}{\lambda} - \ln\frac{L}{\lambda} + \dots)$. For large medium ($L \gg \lambda$) the first term dominates and shows the importance of multiple emissions as it makes the multiplicity big even for $\bar{\alpha} \ll 1$. Even the second term is enhanced and therefore the HO region definitely requires resummation.

- OE region: the divergence is milder in this region. The first two terms are not enhanced for a large medium as $E \gg \frac{L}{\lambda}\bar{\omega}_c$. Surprisingly, the last term has a length logarithm similar to the HO region. This term, however, is suppressed compared to the HO region by $\hat{q}_0/\hat{q} \ll 1$ (this ratio is the result of including single hard corrections in the bare $\hat{q}_0$ parameter in the IOE formalism). The resummation of hard-scattering-induced emissions is needed, but it is subleading compared to the HO region.

- BH region: emissions from $\omega < \omega_{\rm BH}$ results both length and double logarithmic enhancement $\ln^2\frac{\omega_{\rm BH}}{\omega_{\rm min}}$. Clearly, if the medium is big $L \gg \lambda$ and/or $\omega_{\rm BH} \gg \omega_{\rm min}$ the integral becomes very large, and resummation is needed. This region is usually neglected in phenomenology studies as $\omega_{\rm BH}$ is already non-perturbatively small energy and so formally $\omega_{\rm BH} \approx \omega_{\rm min}$.

Including corrections to the induced spectrum will not necessarily change the pole structure of $\lim_{\omega\to 0}\frac{{\rm d}I}{{\rm d}\omega}$. These corrections will clearly not be enhanced in Eq. (3.32), and therefore they are not important in terms of resummation. Other corrections, e.g. finite-z, running α_s, etc., might change the pole-structure and therefore have a large contribution in Eq. (3.32). Their relative importance can be studied by reading off their contribution to the multiplicity. In Sec. 2.2, we went a bit further than the multiplicity and we showed how wold this happens on the level of an observable. We will extend this educated guess in the next sections.

[12] In the vacuum, the multiplicity would look like

$$\int_0^R {\rm d}\theta \int_0^1 {\rm d}z\, \widetilde{P}(z,\theta)\Theta(k_t = z\theta p_t > k_{t,\rm min}) \approx \bar{\alpha}\ln^2\left(\frac{Q_{\rm hard}}{Q_{\rm np}}\right) + \mathcal{O}(\alpha_s \log)\,, \tag{3.33}$$

where splitting function in the IRC limit is $\widetilde{P}(z,\theta) \approx \frac{2\bar{\alpha}}{z\theta} + ...$, and $\frac{Q_{\rm hard}}{Q_{\rm np}} = \frac{p_t R}{k_{t,\rm min}}$. The higher order terms in the multiplicity are coming from corrections to the splitting function beyond the IRC limit such as hard collinear emissions, running coupling, etc.

3.3 Factorization in the soft limit in the medium

To build on the analogy of vacuum resummation, we treat Eq. (3.28) as our LO_m $1 \to 2$ cross-section. This consists of the first order in α_s (from $\mathcal{L}_{\mathrm{hard}}$), but all orders in $\hat{q}_0$ (from $\mathcal{L}_{\mathrm{hard-med}}$). We saw that it has an interesting infrared structure (i) it lacks collinear divergence, and (ii) it has a soft divergence with a non-trivial shape. We identified regions of the phase space, where different divergences emerge. We showed how the different regions of the infrared pole contribute to power and logarithmic enhancements in the multiplicity, which requires all-order resummation.

To perform resummation, we have to evaluate multi-parton final states in the medium. We will follow in the footsteps of vacuum resummation. In Sec. 2.4, we showed that the $|\mathcal{M}_{m+1}|^2$ cross-section factorizes to $|\mathcal{M}_m|^2 \otimes D_{ij,k}$ in the soft and collinear limit. This calculation is waiting to be done in the medium. Two gluon emissions have been addressed, already, showing its factorization in Ref. [91]. The authors used the strongly ordered, double soft limit ($\omega_2 \ll \omega_1 \ll E$) in the harmonic oscillator approximation. In this limit, $|\mathcal{M}_{1\to3}|^2 \approx \sum_{\mathrm{conf}} |\mathcal{M}_{1\to2}|^2 |\mathcal{M}_{1\to2}|^2 + \mathcal{O}(\mathrm{corr.})$, where the summation runs over all possible emission configurations. The factorization of multiple legs was also argued in the same work. The resulting resummation picture was introduced and solved in several later works Refs. [92–95], and more recently see in Refs. [96–98].

Reference [81, 99–101] went further by extracting corrections to the factorization beyond the soft limit $\omega_2 \sim \omega_1 \sim E$ limit similarly in the HO approximation. The leading order goes as $\frac{\mathrm{d}I}{\mathrm{d}\omega} = \mathcal{O}(\omega_1^{-3/2}\omega_2^{-3/2})$, while the first correction comes is $\Delta\frac{\mathrm{d}I}{\mathrm{d}\omega} = \mathcal{O}(\omega_1^{-1}\omega_2^{-3/2})$. This correction should be absorbed in the LO_m splitting in the future. They also extracted NLO_m corrections to the splitting function, which leads to the running of the coupling α_s and additional running of the $\hat{q}$. These results are now awaiting their application in a resummation framework [102, 103]. An important limitation of these studies is that they neglect vacuum physics, consider a medium with infinite size, and stick to the harmonic oscillator approximation.

Ideally, 2-induced emission calculations should be done for the more general scattering potential $\gamma(\boldsymbol{k})$ to relax the harmonic oscillator approximation. Until that, one can use formation time arguments, showing that the most dominant contributions of the Bethe–Heitler and opacity expansion regions factorize in the strongly ordered, soft limit (see Sec. 4.2 in Ref. [75]). This factorization, however, is not essential as the resummation of these regions contributes to the multiplicity (accuracy) at much higher orders (see Eq. (3.32)).

3.4 Medium-induced cascade

In the previous sections, we explained it is necessary to consider medium-induced emissions to all orders and to resum many of them. In vacuum physics, this phenomenon results in jets, which are experimentally observable objects. Based on analogy, where are medium-induced jets in heavy-ion collisions? Do experiments reconstruct them? The answer is less trivial than one would think:

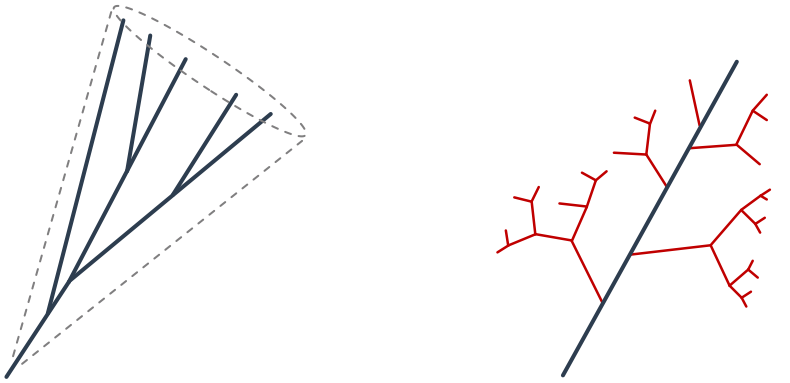
- Medium-induced emissions lack collinear singularity, and therefore they happen at wider angles than vacuum emissions. For the same reason, they are not collimated, and therefore, jet reconstruction algorithms will not cluster them to vacuum jets or together.

- When a vacuum jet is reconstructed with cone size $R = 0.2 - 0.4$, that angle is small compared to the typical angle where multiple medium-induced emissions happen. The typical angle for emissions that drives energy out of the cone is $\theta \sim \frac{\theta_c}{\alpha_s^2} \approx 1$, where $\theta_c \approx 0.01$, and will be introduced in the next section. Therefore, most medium-induced emissions are out of the jet cone [94, 98, 104]. Therefore, medium-induced emissions barely contaminate vacuum jets.

- Opening up the jet cone to $R = 1$, more medium-induced emissions will contribute to the jet. Multiple branchings of medium-induced emissions, however, degrading energy extremely efficiently to Q_{med} and below [93, 95]. Therefore, many induced emissions will disappear in background subtraction, driving us out of the context of this thesis.

Now we can understand why the heavy-ion community does not talk about medium-induced jets. Medium-induced cascade is a more correct terminology as it happens at wide angles and degrades energy rapidly. Figure 3.3 shows the sketches of a jet versus a cascade. The lines denote both quarks and gluons, while the red lines are medium-induced splittings. Vacuum branchings are collinear and thus form a jet, while medium-induced emissions are at wide angles and their rate rapidly accelerates.

The arguments above deal with absolutes: jets with infinitely high energies, medium with infinitely large lengths, etc. In collider experiments, the picture is not black and white, as the scales are not that separated. It is required to go beyond leading behavior arguments. In Ref. [75], we looked at how the finite length of the medium and not-so-soft induced emissions will change the resumed fragmentation function as a first step in addressing more realistic setups.



jet medium-induced cascade

Figure 3.3: The sketch of a jet and a medium-induced cascade.

Using the unified picture from Sec. 3.1.5 and the factorization argument from Sec. 3.3, one can write the energy distribution of induced-emissions as

$$\frac{\partial}{\partial t}D(x,t) = \int_x^1 \mathrm{d}z\, \mathcal{K}\left(z, \frac{x}{z}E, t\right) D\left(\frac{x}{z}, t\right) - \int_0^1 \mathrm{d}z \mathcal{K}(z, xE, t) D(x,t)\,, \tag{3.34}$$

where the fragmentation function is $D(x,t) = x\frac{\mathrm{d}N(t)}{\mathrm{d}x}$, with $x = \frac{\omega}{E}$, and E is the energy of the initiator. The branching kernel is the emission rate $\mathcal{K}(z,E,t) = 2\frac{\partial}{\partial t}\frac{\mathrm{d}I}{\mathrm{d}z}\big|_E$, and the incoming hard parton is $D(x, t=0) = \delta(1-x)$. We solved Eq. (3.34) in different approximations, which are in Fig. 3.4 in different time steps. At early times, $D(x)$ peaks close to $x \approx 1$, as the initial parton has not lost much energy. This peak reduces and disappears at later times. The full solution uses our unified medium-induced emission kernel, where the color shading corresponds to the regions in Fig. 3.1. The curve presented with D_0 contains the strict soft limit of $\frac{\mathrm{d}I}{\mathrm{d}\omega}$ and assumes infinite medium and, therefore, it only resums terms that are $\mathcal{O}(\alpha_s \frac{L}{\lambda})$ (see Eq. (3.32)). In a phase space picture, D_0 would assume the whole (ω, t) plane to be red. Compared to the full solution, it gives a good approximation if the propagation in the medium is long enough, and the emitted energy overlaps with the HO region (red region). The difference is also obvious if the medium is not large or one is interested in hard emissions $x \lesssim 1$. The curves HO and HO+NHO are the first two orders of the improved opacity expansion and include finite-z, finite length, and other corrections. They correctly capture the full solution in the red+green combined region.

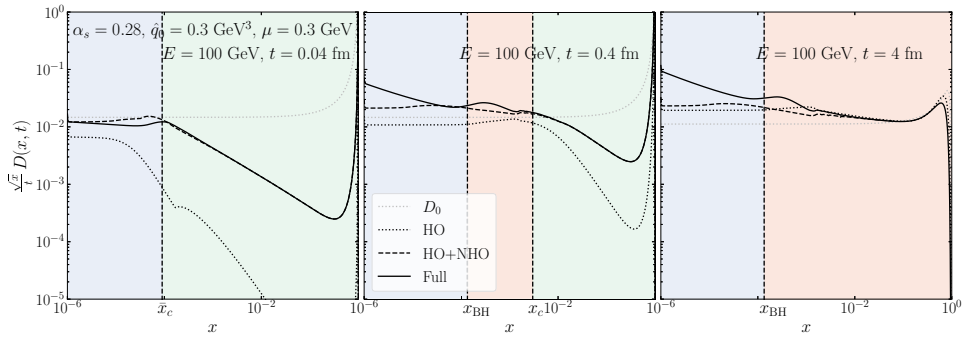


Figure 3.4: The fragmentation function of the medium-induced cascade initiated by a hard parton at different time-steps. Reprinted from Ref. [75].

Figure 3.4 shows how to improve energy loss models in the future, as most models include only the HO approximation or the opacity expansion (green region) exclusively. Based on Eq. (3.32) and Fig. 3.4, it becomes obvious, neglecting the HO region, one would miss leading effects in the energy loss. One could progress further, and show analytically the accuracy of $D(x,t)$ in the different approximations as we did in Ref. [75], which we skip for now.

3.5 Jets and their substructure in the medium

The previous chapter introduced how multiple collinear emissions in vacuum form jets. Jets are also observed in heavy-ion collisions, although, their distribution and substructure are modified [49, 105, 106]. These modifications have various physical origins, and here we will focus on changes related to medium-induced emissions.[13] As we mentioned, the cascade of medium-induced emissions lacks collinear singularity, and therefore they appear at a typically large angle outside of the jet cone. The fundamental effect of medium-induced cascades on jets is decreasing their energy.

At higher orders, medium-induced emissions can also modify jet substructure. Interference between vacuum and medium-induced emissions becomes important, and therefore, the transverse momentum of induced emissions has to be kept $\frac{\mathrm{d}^3 I}{\mathrm{d}\omega \mathrm{d}^2 \boldsymbol{k}}$. Most works are based on the harmonic oscillator approximation and the soft emission limit [107–111]. Works which studied these interference effects are referred to as color-coherence stud-

[13] In the formalism introduced in Ch. 3, we only consider a single medium scale $Q_{\rm med}$ and neglect the medium dynamics $\mathcal{L}_{\rm med}$. In a collider experiment, more scales are present, which vary dynamically. The evolution of the medium also contaminates jets in a non-trivial way. The formalism established here cannot be complete. Our formalism should be valid for narrow jets with very high energy, where non-perturbative effects are less important.

ies. They showed that the color connection of a color-dipole decoheres by the time $t_d = [4/(\hat{q}\theta_{q\bar{q}}^2)]^{1/3}$ due to multiple scatterings in the medium, where the initial angle of the dipole is $\theta_{q\bar{q}}$. Therefore, not every scattering in the medium changes colors right away.

This simple phenomenon has several important consequences:[14]

- Emissions formed much before the decoherence time $t_f \ll t_d$ are unaffected by the medium. They form through a vacuum process in the angular ordered fashion and are referred to as vacuum-like emissions. They appear in the red region in Fig. 3.5.
- Early vacuum-like emissions, after propagating t_d in the medium (for large medium $t_d \ll L$), will become color-independent of each other (decohere). They will source medium-induced emissions, and thus energy loss, independently. The energy loss of a jet will, therefore, correlate with the jet multiplicity in a non-trivial way [113]. See also Sec. 3.6.1.
- There exists a critical resolution angle θ_c when $t_d = L$ (vertical red line in Fig. 3.5). If the splitting is narrow ($\theta \ll \theta_c$), the medium will not resolve its constituents, and therefore, it loses less energy as a single color charge. Wide splittings, on the other hand, get resolved by the medium and trigger more energy loss.
- Medium-induced emissions happen in a very short time scale, which is quick enough that no vacuum process will occur during this time. As the branching process accelerates, the whole medium-induced cascade will finish without any vacuum interruption [93].
- Emissions formed slow $t_f \gg L$ are again unaffected by the medium in the angular ordered fashion (right side of the blue line in Fig. 3.5). As these emissions formed after their ancestors are decohered, the available phase space gets "reopened" [96, 114, 115].

The picture presented above is the factorized picture or 2-, or 3-stages of in-medium jet evolution, and it is the basis of any modern jet quenching model. The presented picture is also a sketch for the numeric interpretation, see in Refs. [96, 114]. In the future, it is necessary to include corrections to the factorized picture from the unified picture presented in Sec. 3.1.5.

Even though most models agree with the factorized picture, the exact shape of the

[14] An excellent summary of the vacuum-medium interference effects (including derivations) is in Ref. [112], or its shorter version is in Ref. [96].

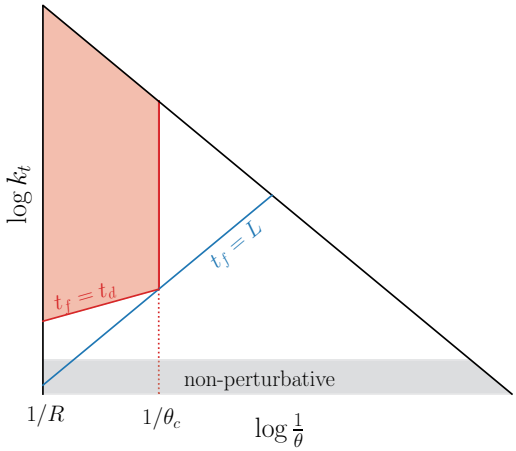


Figure 3.5: The sketch of the vacuum-like emission phase space after including its medium modification.

resolved phase space region and the medium-induced branching process vary between models. In Sec. 3.6.2, we construct an observable to distinguish between the models.

3.6 Applications

This section reviews two works that apply the formalism from Ch. 3 to perform jet quenching phenomenology. We will particularly focus on understanding the accuracy presented in the calculations.

3.6.1 Jet energy-loss and the quenched jet spectrum

A relatively simple application of medium-induced emissions is calculating the radiative energy loss of jets. In Ref. [116], we used a simplified version of the unified picture presented in Sec. 3.1.5 to compute corrections to the harmonic oscillator approximation by including hard scattering induced emissions. We applied this to calculate the nuclear modification factor R_{AA} in dijet and Z+jet events.

We write the jet spectrum in the medium as the vacuum spectrum shifted by lost energy,

$$\frac{\mathrm{d}\sigma_R^{\mathrm{med}}}{\mathrm{d}p_T} = \sum_{i=q,g} \int_0^\infty \mathrm{d}\varepsilon\, \mathcal{E}_i(\varepsilon, p'_T, R)\, \frac{\mathrm{d}\sigma_{R,i}^{\mathrm{vac}}}{\mathrm{d}p'_T}\bigg|_{p'_T=p_T+\varepsilon}, \tag{3.35}$$

where $\sigma^{\rm vac}_{R,i}$ is the vacuum spectrum of jets with cone size R and i initiator flavor. The $\mathcal{E}_i(\varepsilon, p_T', R)$ is the probability to lose energy ε. This model clearly does not say anything about the change in the jet substructure. The modification due to energy loss is then

$$R_{\rm med}(p_T) = \frac{\mathrm{d}\sigma^{\rm med}_R}{\mathrm{d}p_T} \Big/ \frac{\mathrm{d}\sigma^{\rm vac}_R}{\mathrm{d}p_T} \approx \sum_{i=q,g} \int_0^\infty \mathrm{d}\varepsilon\, \mathcal{E}_i(\varepsilon, p_T+\varepsilon, R) e^{-\frac{n}{p_T}\varepsilon}\,, \tag{3.36}$$

where in the second step we assumed a steeply falling vacuum spectrum ($\frac{\mathrm{d}\sigma^{\rm vac}}{\mathrm{d}p_T} \sim p_T^{-n}$, where $n \gg 1$), and therefore $\frac{\mathrm{d}\sigma^{\rm vac}}{\mathrm{d}(p_T+\varepsilon)} \approx \frac{\mathrm{d}\sigma^{\rm vac}}{\mathrm{d}p_T} e^{-\frac{n}{p_T}\varepsilon}$. Formally, Eq. (3.36) is a Laplace transform of the energy loss probability.

At LO$_m$, the radiative energy loss distribution consists of a single medium-induced emission is

$$\mathcal{E}_i(\varepsilon, p_T', R) \approx \frac{\mathrm{d}I_{>,i}}{\mathrm{d}\omega}\delta(\varepsilon-\omega) + \delta(\varepsilon)\left[1 - \int \mathrm{d}\omega\, \frac{\mathrm{d}I_{>,i}}{\mathrm{d}\omega}\right], \tag{3.37}$$

where the second term is from virtual terms, and the probability explicitly conserves. The out-of-cone emission spectrum is

$$\frac{\mathrm{d}I_{>,i}}{\mathrm{d}\omega} = \int_{\boldsymbol{k}} \frac{\mathrm{d}^3 I_i}{\mathrm{d}\omega \mathrm{d}^2\boldsymbol{k}}\, \Theta\left(\theta \approx \frac{|\boldsymbol{k}|}{\omega} > R\right). \tag{3.38}$$

The jet loses energy only if the emission goes out of the jet cone. The p_T' is the energy of the initial parton. The soft limit (see Eq. (3.31)) is independent of p_T'. Previously, in Eq. (3.32), we showed that the integral of the induced spectrum is large $\int \mathrm{d}I \gg 1$, and therefore, it is necessary to consider multiple emissions. This is also the case for Eq. (3.37) as in the $\varepsilon \to 0$ limit, the integral is unregulated, and thus multiple soft emissions will appear.

To account for multiple emissions, assume parton i to emit multiple independent induced gluons (following the factorization argument from Sec. 3.3), one can write the energy loss as [117, 118]

$$\mathcal{E}_i^{(0)}(\varepsilon, R) = \sum_{n=0}^\infty \frac{1}{n!}\left[\prod_{m=1}^n \int \mathrm{d}\omega_m \frac{\mathrm{d}I_{>,i}}{\mathrm{d}\omega_m}\right] \delta\left(\varepsilon - \sum_{m=1}^n \omega_m\right) e^{-\int \mathrm{d}\omega \frac{\mathrm{d}I_{>,i}}{\mathrm{d}\omega}}\,, \tag{3.39}$$

where n is the number of induced emissions (the $n!$ symmetry factor comes from indistinguishable gluons), and the radiated energy adds up to ε. In this limit, the p_T' dependence is negligible. The last exponential factor comes from virtual terms on which we already performed the summation. This formula should remind the reader of the coherent branching algorithm and Eq. (2.30). In Mellin space, the energy-loss distribution

simplifies as the sum in the delta function is

$$\delta\left(\varepsilon - \sum_{m=1}^{n} \omega_m\right) = \int_C \frac{\mathrm{d}\mu}{2\pi i}\, e^{\mu\varepsilon} \prod_{l=1}^{k} e^{-\mu\omega_l}\,, \tag{3.40}$$

resulting in the simple analytical expression referred to as quenching weight,

$$\mathcal{Q}_i^{(0)}(\nu, R) = \int_0^\infty \mathrm{d}\varepsilon\, \mathcal{E}_i^{(0)}(\varepsilon, R) e^{-\nu\varepsilon} = \exp\left[\int \mathrm{d}\omega \frac{\mathrm{d}I_{>,i}}{\mathrm{d}\omega}\left(e^{-\nu\omega} - 1\right)\right]. \tag{3.41}$$

Equation (3.41) and the nuclear modification Eq. (3.36) are very similar, one only has to take $\nu \mapsto \frac{n}{p_T}$.

As we showed in Eq. (3.32), the leading terms of $\int \mathrm{d}I = \mathcal{O}(\alpha_s \frac{L}{\lambda}) + \mathcal{O}(\alpha_s \ln \frac{L}{\lambda}) + \dots$. By resumming multiple independent emissions (neglecting the cone size for the moment), we roughly exponentiate the multiplicity and arrive at an expression that is $(\mathcal{O}(\alpha_s \frac{L}{\lambda}) + \mathcal{O}(\alpha_s \ln \frac{L}{\lambda}) + \dots)^n$. Of course, we did not capture all of these orders correctly! Taking the vacuum analogy: at LO, the integral of the splitting function is $\mathcal{O}(\alpha_s \log^2) + ...$ The argument of the logarithm is irrelevant for now. By resumming many independent emissions, one gets terms with $\mathcal{O}(\alpha_s^n)$, but we know emissions are independent only in the IRC limit. Therefore we correctly capture terms that are $\mathcal{O}(\alpha_s^n \log^{2n})$. When the logarithm is big (big phase space for IRC emissions $Q_{\mathrm{hard}} \gg Q_{\mathrm{np}}$), IRC-logarithm-enhanced terms dominate.

To understand which terms we captured correctly, we saw that, Ref. [91, 100] suggest independent emissions for large medium in the soft limit ($\omega \ll E$ and $L \ll \lambda$). This is valid for both HO terms presented in Eq. (3.32). Therefore, we expect Eq. (3.41) to capture correctly terms which are

$$\alpha_s^n \left(\frac{L}{\lambda}\right)^n, \alpha_s^n \left(\frac{L}{\lambda}\right)^{n-1} \ln \frac{L}{\lambda}, \dots, \alpha_s^n \ln^n \frac{L}{\lambda}\,, \tag{3.42}$$

where $n = 1, 2, \dots$, and $\lambda = \mu^2/\hat{q}_0$. We neglected higher-order terms, which do not necessarily factorize. Equation (3.42) is the first sketch of defining all-order accuracy for medium-induced observables. We expect factorization to hold for other BH and hard scattering too, however, these have been proven only through formation time arguments. We neglected them for the moment as the extension would be trivial. Moreover, using the pioneering work of Ref. [100], higher-order terms could also be included in the factorization, achieving higher accuracy.

Based on Sec. 3.5, we know that not every jet constituents source energy loss, only those resolved by the medium. We would like to include this resolution criterion in the

energy loss distribution. If the jet only consists of 2-particles (created through a vacuum splitting),

$$\begin{aligned}\mathcal{E}_i(\varepsilon, p_T', R) \approx & \int \mathrm{d}\varepsilon_0\, \mathcal{E}^{(0)}_{i,p_t',R}(\varepsilon_0) \int \mathrm{d}\widetilde{P}_i (1-\Theta_{\mathrm{res}})\delta(\varepsilon-\varepsilon_0) \\ & + \int \mathrm{d}\varepsilon_0\, \mathcal{E}^{(0)}_{i,(1-z)p_t',R}(\varepsilon_0) \int \mathrm{d}\widetilde{P}_i \Theta_{\mathrm{res}} \int \mathrm{d}\varepsilon_1 \mathcal{E}^{(0)}_{g,zp_t',R}(\varepsilon_1)\delta(\varepsilon-\varepsilon_0-\varepsilon_1) \\ & - \int \mathrm{d}\varepsilon_0\, \mathcal{E}^{(0)}_{i,p_t',R}(\varepsilon_0) \int \mathrm{d}\widetilde{P}_i \delta(\varepsilon-\varepsilon_0)\,, \end{aligned} \tag{3.43}$$

where $\int \mathrm{d}\widetilde{P}_i = \int \mathrm{d}\theta \int \mathrm{d}z \widetilde{P}_i^{\mathrm{vac}}(z,\theta)$ is the LO vacuum-like emission and its integral over all possible kinematics. The $\Theta_{\mathrm{res}} = \Theta(t_f < t_d < L)$ is the resolution criterion that corresponds to the red region in Fig. 3.5. The first line describes an unresolved emission. The initiator loses the energy. The second line describes a resolved emission, and thus both legs lose energy independently. The third line includes the virtual term, and it has no phase space constraint. By generalizing for an arbitrary number of vacuum-like emissions, one gets

$$\mathcal{E}_{i,p_t',R}(\varepsilon) = \int \mathrm{d}\varepsilon_0\, \mathcal{E}^{(0)}_{i,p_t',R}(\varepsilon_0) \sum_{n=0}^{\infty} \frac{1}{n!} \prod_{m=1}^{n} \int \mathrm{d}\widetilde{P}_m \mathrm{d}\varepsilon_m \Theta_{\mathrm{res}} \left[\mathcal{E}^{(0)}_{g,zp_t',R}(\varepsilon_m)\delta(\varepsilon-\varepsilon_0-\sum_m \varepsilon_m) - 1 \right]. \tag{3.44}$$

The physical interpretation of Eq. (3.45) is that $\mathcal{E}^{(0)}$ single parton energy loss gets correction from additional vacuum-like emissions inside the jet, which are in the resolved region. At double logarithmic accuracy, these emissions are always soft gluons and $(1-z)p_t' \approx p_t'$. In Mellin space, it takes the simple form

$$\mathcal{Q}_i(\nu, R) = \mathcal{Q}_i^{(0)}(\nu) \exp\left[\int \mathrm{d}\theta \mathrm{d}z \frac{\mathrm{d}\widetilde{P}}{\mathrm{d}\theta \mathrm{d}z} \Theta_{\mathrm{res}} \left(\mathcal{Q}_g^{(0)}\left(\frac{\nu}{z}\right) - 1 \right) \right]. \tag{3.45}$$

Sometimes Eq. (3.45) is referred to as resummed quenching weight, where the resummation refers to vacuum-like emissions (as in Eq. (3.41) multiple medium-induced emissions were already resummed).

The gluon quenching weight is shown in Fig. 3.6 for different $p_T = \frac{n}{\nu}$ (with constant spectrum slope) and for different jet cone sizes R. At high jet p_T, the lost energy is relatively small ($\mathcal{Q} \approx 1$), while at lower p_T, energy loss becomes more dominant. For $\mathcal{Q}^{(0)}$, it is clear that opening up the jet cone recovers some of the lost energy, and therefore $\mathcal{Q}^{(0)}$ gets closer to 1. For the resummed quenching weight $\mathcal{Q}$, two effects compete for 1) opening the cone recaptures some of the lost energy, and 2) a wider cone containing more resolved, vacuum-like emissions, sourcing more energy loss. The overall effect results in a mild cone size dependence.

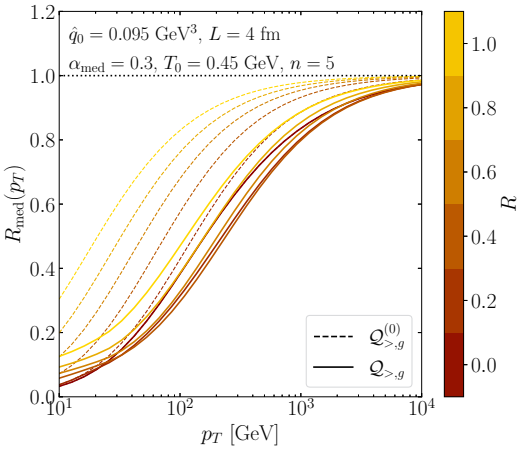


Figure 3.6: The p_T and jet cone size dependence of the single parton quenching weight $\mathcal{Q}^{(0)}$ from Eq. (3.41) and the resumed quenching weight $\mathcal{Q}$ from Eq. (3.45). Reprinted from Ref. [116].

Let us examine what kind of terms are resummed in Eq. (3.45). We have already identified the expansion of $\mathcal{Q}^{(0)}$ in Eq. (3.42). At leading resumed order it is $\mathcal{O}(\alpha_s \frac{L}{\lambda})^n$. The vacuum integral at leading double logarithmic accuracy is $\int \mathrm{d}\mathcal{P} = \mathcal{O}(\alpha_s \log^2)$. Therefore, Eq. (3.45) combines both vacuum and medium terms in a non-trivial way,

$$\left[\mathcal{O}(\alpha_s \log^2) + \mathcal{O}(\alpha_s \log) + \mathcal{O}(\alpha_s) + ...\right]^n \cdot \left[\mathcal{O}\left(\alpha_s \frac{L}{\lambda}\right) + \mathcal{O}\left(\alpha_s \ln \frac{L}{\lambda}\right) + ...\right]^m , \quad (3.46)$$

where $n, m = 1, 2, ...$ To understand which terms are captured correctly, one has to use Secs. 2.3-2.4, and 3.2-3.3 that we delay for a future work. As the medium resolution criterium results in the non-trivial combination of vacuum and medium emissions, color-coherence is the first, non-trivial interference in our jet modification theory. Moreover, Eq. (3.46) is the first sketch of the all-order expansion of a quenched jet observable.

Finally, to extend Eq. (3.45) to LHC phenomenology, several other effects are necessary to include: realistic vacuum spectrum, quark/gluon fraction, and nuclear modification of the PDF. Including all of these effects, in Ref. [116] (see also Ref. [119] where they focused more on medium fluctuations), we made predictions for dijet and Z+jet R_{AA}. It is shown on the left of Fig. 3.7. It is interesting comparing these seemingly different R_{AA} as they have the same medium in the collision. Therefore, the energy loss probability, nPDFs, medium recoil, and other background pollution are all the same. Differences originate in the vacuum spectrum $\sigma^{\rm vac}$, and therefore, they represent different biases (n) and different q/g fractions. Typically, the dijet spectrum falls steeper (favoring small energy loss), while it has more gluons which lose more energy due to their larger color

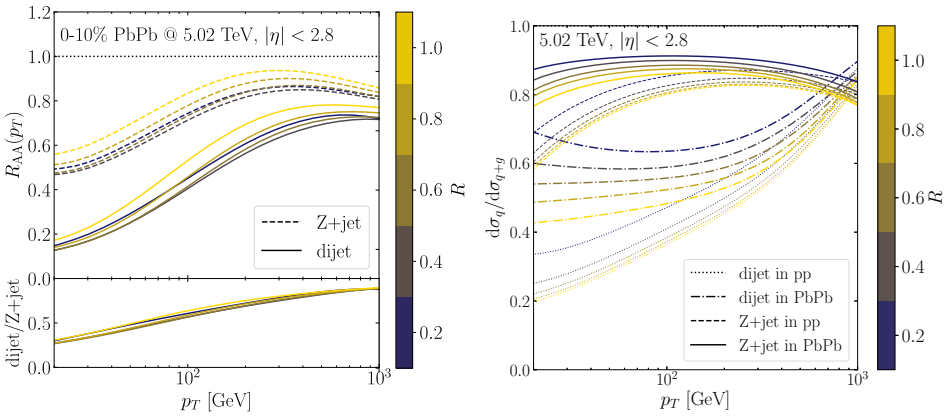


Figure 3.7: Using the quenching weight for LHC phenomenology. *Left:* predicting the nuclear modification factor for dijet and Z+jet events for different jet cone sizes. *Right:* quark initiated jet fraction before and after quenching. Reprinted from Ref. [116].

factor ($\frac{C_g}{C_q} = \frac{9}{4}$). Overall gluon quenching dominates, and dijets have more quenching. We illustrate this on the right side of Fig. 3.7, which shows the quark/gluon jet ratio in pp and after quenching. In both cases, the quark fraction increases as gluons jets are reduced due to the bias effect.

3.6.2 Dynamical grooming in heavy-ion collisions

As another application of the presented jet modification framework, we review Ref. [120]. In this work, we put the factorized picture of jet energy loss (Sec. 3.5) to the test. As mentioned, modern jet quenching models are based on a factorized picture. The exact shape of the modified phase space (red region in Fig. 3.5) and the induced emission kernel $\frac{\mathrm{d}I}{\mathrm{d}\omega}$ are different. We aimed to construct an observable, sensitive to the critical resolution angle θ_c. This angle is present in models which account for the interference between vacuum and medium emissions in the multiple soft scattering (HO) regime. This might sound peculiar, but as we saw in Sec. 3.2, effects from the HO region are the most dominant, and this interference is a lowest-order interference between vacuum and medium emissions.

We wanted an observable that is not sensitive to the medium background while it is sensitive to differences in splitting angles. Our requirement leads us to the angle distribution of the hardest emission inside the jet (dynamically groomed θ_g). We have already introduced dynamical grooming in Sec. 2.6.1. In brief, dynamical grooming loops through branchings in a jet and identifies the hardest branching, where the hardness is defined

through $\kappa^{(a)} = z\theta^a$, where a is the grooming parameter. DyG returns the hardest branching, one can study its kinematics, for example, the angle, transverse momentum, etc. As it is a groomer, it is not only less sensitive to non-perturbative physics but reduces contributions coming from the background as we will show. Based on Sec. 2.6.1, the angular distribution of the hardest splitting in the vacuum is then

$$\frac{1}{\sigma}\left.\frac{\mathrm{d}\sigma}{\mathrm{d}\theta_g}\right|_a = \int \mathrm{d}z_g\, \widetilde{P}_i(z_g,\theta_g)\Delta_i(\kappa = z_g\theta_g^a)\,, \tag{3.47}$$

$$\approx \frac{1}{\theta_g}\sqrt{\bar{\alpha}\pi a}\left[\mathrm{erf}\left(\sqrt{\bar{\alpha}a}\ln\theta_g\right)+1\right]\,, \tag{3.48}$$

where $\widetilde{P}$ is the $1\to 2$ branching probability, and $\Delta(\kappa)$ denotes the no harder emission than κ probability. In the second line, we evaluated the integral at double logarithmic accuracy. The θ_g distribution at DLA is shown in Fig. 3.9.

The medium modification of a groomed observable comes from several sources, most importantly, energy loss. The energy loss will depend on the jet substructure. When the hardest tagged emission is narrow ($\theta_g \ll \theta_c$), it is unresolved by the medium. The initiator loses energy only. When the tagged emission is wide ($\theta_g \gg \theta_c$), it gets resolved, and both legs will independently lose energy. Furthermore, softer, untagged, emissions will also contribute to energy loss, but only if they are resolved by the medium.

Extending our energy loss probability from Eq. (3.44), let us assume that we tagged the hardest splitting with (z_g,θ_g) that is $\kappa = z_g\theta_g^a$ (blue dot in Fig. 3.9). We are interested in the energy loss of this system by including a softer vacuum-like emission

$$\begin{aligned}\mathcal{E}_{i,p_t',R}(\varepsilon|\kappa) \approx &\int \mathrm{d}\varepsilon_0\, \mathcal{E}^{(0)}_{i,p_t',R}(\varepsilon_0)\int \mathrm{d}\widetilde{P}_1\Theta(\kappa - z_1\vartheta_1^a)(1-\Theta_{\mathrm{res}})\delta(\varepsilon-\varepsilon_0)\\ &+\int \mathrm{d}\varepsilon_0\, \mathcal{E}^{(0)}_{i,p_t',R}(\varepsilon_0)\int \mathrm{d}\widetilde{P}_1\Theta(\kappa - z_1\vartheta_1^a)\Theta_{\mathrm{res}}\int \mathrm{d}\varepsilon_1 \mathcal{E}^{(0)}_{g,z_1p_t,R}(\varepsilon_1)\delta(\varepsilon-\varepsilon_0-\varepsilon_1)\\ &-\int \mathrm{d}\varepsilon_0\, \mathcal{E}^{(0)}_{i,p_t',R}(\varepsilon_0)\int \mathrm{d}\widetilde{P}_1\delta(\varepsilon-\varepsilon_0)\,.\end{aligned} \tag{3.49}$$

Here $\Theta(\kappa - z\vartheta^a)$ sets the emission to be softer than κ. The first line is an unresolved emission by the medium, and it loses energy through the initiator. In the second line, the emission is resolved, and thus, both legs will lose energy. Here, we simplified $(1-z_1)p_t' \to p_t'$. The last line is the virtual term which has no phase space constraint. By taking the $\kappa \to 1$ limit, one reproduces the unconstrained probability in Eq. (3.44).[15] Generalizing

[15]The energy loss of the hardest emission also has to be included, that is $\mathcal{E}^{(0)}_{i,p_t,R}(\varepsilon_0)$ if $\theta_g < \theta_c$, and $\int \mathrm{d}\varepsilon_0''\, \mathcal{E}^{(0)}_{i,p_t,R}(\varepsilon_0')\mathcal{E}^{(0)}_{g,p_t,R}(\varepsilon_0'')\delta(\varepsilon_0-\varepsilon_0'-\varepsilon_0'')$ it $\theta_g > \theta_c$.

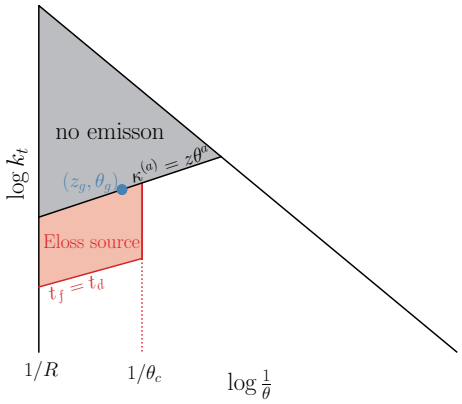


Figure 3.8: The emission phase space of dynamically groomed vacuum-like emissions, modified by the medium.

for an arbitrary number of independent emissions,

$$\mathcal{E}_{i,p_t,R}(\varepsilon|\kappa) = \int \mathrm{d}\varepsilon_0\, \mathcal{E}^{(0)}_{i,p_t,R}(\varepsilon_0) \sum_{n=0}^{\infty} \frac{1}{n!} \prod_{m=1}^{n} \int \mathrm{d}\varepsilon_m \mathrm{d}\widetilde{P}_m \Delta(\kappa)$$
$$\cdot\, \Theta(\kappa - z_m \vartheta_m^a) \Theta_{\mathrm{res}} \left[\mathcal{E}^{(0)}_{g,z_m p_t,R}(\varepsilon_m) \delta(\varepsilon - \varepsilon_0 - \sum_m \varepsilon_m) - 1 \right] . \tag{3.50}$$

The interpretation is simple: emissions softer than κ AND resolved source further energy loss. These sources are in the red region in Fig. 3.8. Its Laplace transformation is

$$\mathcal{Q}_i(\nu|\kappa) = \mathcal{Q}_i^{(0)}(\nu) \Delta(\kappa) \exp\left[\int \mathrm{d}\theta \mathrm{d}z \frac{\mathrm{d}\mathcal{P}}{\mathrm{d}\theta \mathrm{d}z} \Theta_{\mathrm{res}} \Theta(\kappa - z\vartheta^a) \left(\mathcal{Q}_g^{(0)}\left(\frac{\nu}{z}\right) - 1 \right) \right] . \tag{3.51}$$

The $\Delta(\kappa)$ no emission probability makes sure there is no harder emission inside the jet than κ.

By comparing our substructure independent energy loss from Eq. (3.44) with Eq. (3.51), one sees that the main difference relies on the $\Theta(\kappa - z\theta^a)$ condition. The description of $\mathcal{Q}^{(0)}$ and $\widetilde{P}(z,\theta)$ are unchanged, and therefore the only difference in the accuracy comes from the reduction of the vacuum-like emission phase space. This will change leading vacuum terms $\mathcal{O}(\alpha_s \log^2)$ in Eq. (3.46).

The tagged hardest branching can be medium-induced. This is subleading when $\kappa^{(a)} \sim 1$ because hard medium-induced emissions are rare. Capturing these subleading correc-

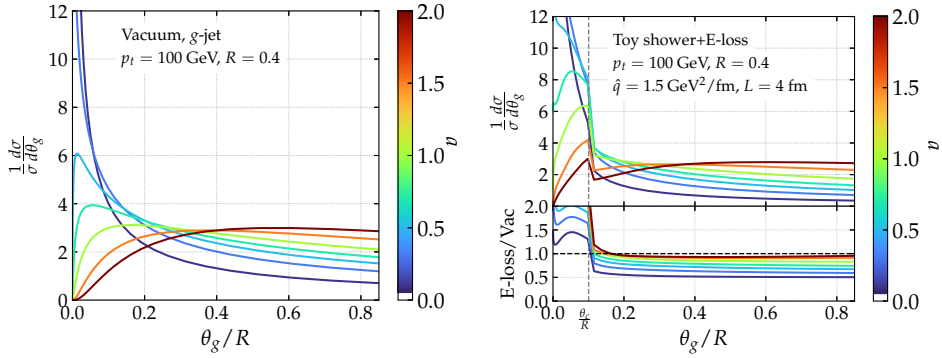


Figure 3.9: *Left:* The angular distribution of the hardest branching at double-logarithmic accuracy (Eq. (3.47)) for different hardness definitions $z\theta^a$. *Right:* the medium modification of the angular distribution (Eq. (3.55) for different grooming parameters a and its ratio to vacuum. Reprinted from Ref. [120].

tions, we follow similar steps as in Eq. (3.47). The medium-induced spectrum is

$$\frac{\mathrm{d}^2 I}{\mathrm{d}z\mathrm{d}\theta} \approx \bar{\alpha}\sqrt{\frac{2\omega_c}{z^3 p_t}} \cdot 2\theta \frac{z^2 p_t^2}{\hat{q}L}\Gamma\left(0, \frac{z^2 p_t^2 \theta^2}{\hat{q}L}\right) \cdot \Theta(\omega_c - z p_t)\,, \tag{3.52}$$

which is valid in the soft small angle ($z \ll 1$ and $\boldsymbol{k}^2 = z^2\theta^2 p_t^2 \ll \hat{q}L$) limit. We can use the small angle approximation for $\mathrm{d}I$ here, as the jet cone is typically narrow. The corresponding no-emission probability is

$$\Delta^{\mathrm{mie}}(\kappa) = \exp\left[-\int \mathrm{d}\theta \int \mathrm{d}z \frac{\mathrm{d}^2 I}{\mathrm{d}z\mathrm{d}\theta}\Theta\left(z\frac{\theta^a}{R^a} - \kappa\right)\right]. \tag{3.53}$$

Therefore the overall probability of tagging the hardest splitting is

$$P_i^{\mathrm{med}}(z_g, \theta_g|a) = \left(\widetilde{P}_i^{\mathrm{vac}}(z_g, \theta_g|a) + I^{\mathrm{mie}}(z_g, \theta_g)\right)\Delta_i^{\mathrm{vac}}(\kappa)\Delta^{\mathrm{mie}}(\kappa)\,. \tag{3.54}$$

Finally, including energy loss results in

$$\left.\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\theta_g}\right|_a = \sum_{i=q,g}\int \mathrm{d}\varepsilon \frac{\mathrm{d}\sigma_i}{\mathrm{d}(p_t+\varepsilon)}\int \mathrm{d}z_g P_i^{\mathrm{med}}(z_g, \theta_g|a)\mathcal{E}_{i,p_t,R}(\varepsilon|z_g\theta_g^a)\,. \tag{3.55}$$

The resulted θ_g distribution is showed in Fig. 3.9.

The simple quenching model presented in Eq. (3.55) correctly captures asymptotic properties of the θ_g observable if the jet energy is very high and the medium size is large. Corrections coming from finite jet-p_t, finite medium length, fluctuating medium, and pollution from the background particles will not vanish out the nice peak in Fig. 3.9. To show this, we studied different Monte Carlo (MC) event generators, which include

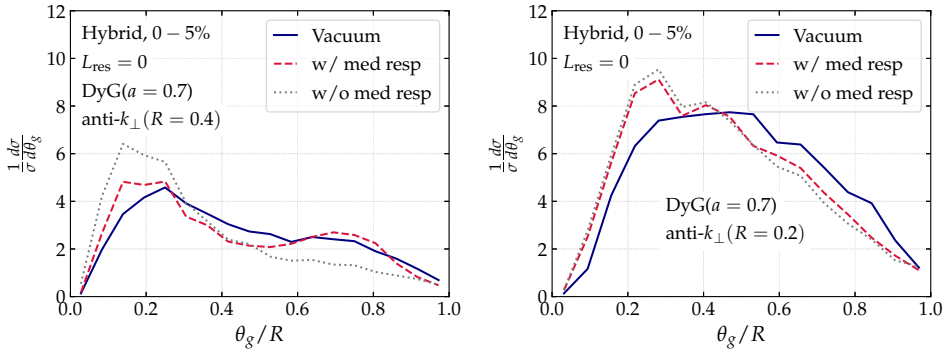


Figure 3.10: The angular distribution of the hardest branching in the jet using the Hybrid model for two different jet cone sizes, with and without medium response. Reprinted from Ref. [120].

non-perturbative effects. These MCs include, e.g., hadronization, dynamical and fluctuating medium, medium recoil, and medium pollution of jets, and therefore, they are ideal for testing our observable. The Hybrid model, for example, implements a similar factorization picture presented in Sec. 3.5, but the energy-loss is modeled through AdS/CFT correspondence [114, 121]. The lost energy is deposited to hydrodynamics as external source terms [122]. The results of these MC analyses are shown in Fig. 3.10. On the left, one can see the enhancement of wide-angle splittings around ~ 0.7, coming from the medium response. Therefore, the grooming was not enough to reduce medium background effects. On the right panel, we decreased the jet cone size to $R = 0.2$. The resulting θ_g distribution is now free of medium recoil. Dynamical grooming did its job by making the observable less sensitive to contamination from the medium background. There is a clear enhancement of small, and suppression of wide-angle emissions similar to our prediction in Fig. 3.9.

3.7 Conclusion of medium physics

In this chapter, we introduced the underlying theory of jet modification in the quark-gluon plasma. We focused on constructing our framework by relying the least on the dynamics of the medium. The created effective theory is based on perturbative QCD in a colored background that provides in-medium Feynman diagrams. We evaluated these diagrams showing that the medium induces extra emissions.

Following the guidelines of Ch. 2, we studied the infrared property of the emissions identifying the importance of multiple soft emissions. We acknowledge, similar to vacuum,

medium-induced emissions factorize in the strongly ordered soft limit. This makes possible the all-order resummation of medium-induced emissions and thus the description of their cascade. We argued why this cascade is fundamentally different from jets and therefore they are the main source of jet energy loss in the medium. We reviewed interference effects between vacuum and medium emissions that are necessary for jet quenching phenomenology. Finally, we presented the factorized picture of in-medium jet modification that is the basis of all modern jet quenching models.

As an application of the presented formalism, we first calculated the jet energy loss, which is essential for making predictions about the jet nuclear modification factor. We introduced the bias effect, and we showed how the bias effect change between dijet and Z+jet events. Finally, we reviewed a dedicated measurement to test the factorized picture of jet quenching and to observe the critical resolution angle that is a fundamental property of the quark-gluon plasma. We showed using dynamic grooming one is able to reduce the non-perturbative effect especially medium response which is an important effect in decreasing the signal strength.

During the construction of the in-medium jet theory, we tried to illustrate what is the accuracy of the medium, and how it looks in an all-order expansion. We reviewed many recent progressions in the theory of jet modification, and we organized these results in terms of their contribution to accuracy. As an application of the introduced formalism, we calculated jet energy loss, and we introduced for the first time how the accuracy manifests in quenched jet observables such as the nuclear modification factor and substructure observables.

Chapter 4

Outlook

As we have already summarized the two main chapters of this thesis in Secs. 2.7 and 3.7, we refer back to these sections, and here we give an outlook instead. We give a personal overview of which questions are important to address in the future. Several points from here were already discussed in this thesis, but we did not address them in detail.

In the BDMPS-Z formalism that is presented in this thesis, the medium is simplified to a homogeneous brick. The derivations presented in Ch. 3, however, is obviously more general than that, and therefore, more realistic plasma was studied by several works [86, 88–90, 123, 124]. These works lack the possible definition of a systematic all-order expansion and thus all-order accuracy. They use the phenomenologic energy-momentum tensor $T^{\mu\nu}(x)$ of the medium as an input that is unconstrained by theory. To overcome this issue, one has to understand how to define the accuracy of hydrodynamic fields and their evolution. Do hydrodynamic fields converge in the gradient expansion? Is it easier to address this question in the weak coupling picture?

An obvious extension of the accuracy sketch in Ch. 3 involves the introduction of infrared power counting to in-medium diagrams. One expects $m+1$ diagrams to factorize, similar to vacuum physics. Borrowing techniques from advanced jet physics should make it possible to include higher-order corrections to the factorization of medium-induced emissions. Going to higher accuracy is necessary to study beyond leading effects, such as the quenching of heavy-quark jets [83, 125, 126], or jet quenching in a smaller plasma [127–129].

The current implementation of medium-induced cascades is on the level of leading logarithms founded in the 90'. There are several developments in parton shower simulations to which energy loss codes have to catch up, starting with implementing the dipole picture, global recoil, fixed order matching, and finite-N_c corrections.

Once the all-order expansion structure is well defined, one can use it for testing different jet quenching models. In the $E, L \to \infty$ and $\alpha_s \to 0$ numerical limit, one can fix different terms in Eq. (3.32) (e.g. $\alpha_s \frac{L}{\lambda}$, or $\alpha_s \ln \frac{L}{\lambda}$) and see if models correctly reproduce these limits. This numeric accuracy test is a recent rapidly spreading technique in jet physics, from which jet quenching modeling could benefit [36, 37, 130, 131].

Last but not least, the physics involving electron-ion collisions (EIC) is very similar to jet quenching where instead of collinear logarithms, the focus is on the soft ones. Jet modification in a nuclear deep-inelastic scattering uses the same formalism presented in Ch. 3. Most developments presented in this thesis can be directly applied there.

Bibliography

[1] S. Marzani, G. Soyez, and M. Spannowsky, *Looking inside jets: an introduction to jet substructure and boosted-object phenomenology*, vol. 958. Springer, 2019.

[2] O. Saarimäki, *Dijet studies with ALICE at the LHC*, Master's thesis, Jyvaskyla U., 2018.

[3] U. Heinz and R. Snellings, *Collective flow and viscosity in relativistic heavy-ion collisions*, *Ann. Rev. Nucl. Part. Sci.* **63** (2013) 123–151, [arXiv:1301.2826].

[4] M. D. Schwartz, *Quantum Field Theory and the Standard Model*. Cambridge University Press, 2013.

[5] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and collider physics*. Cambridge monographs on particle physics, nuclear physics, and cosmology. Cambridge University Press, Cambridge, 2003. Photography by S. Vascotto.

[6] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043, [arXiv:1002.2581].

[7] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079, [arXiv:1405.0301].

[8] A. Banfi, G. P. Salam, and G. Zanderighi, *Principles of general final-state resummation and automated implementation*, *JHEP* **03** (2005) 073, [hep-ph/0407286].

[9] A. Banfi, H. McAslan, P. F. Monni, and G. Zanderighi, *A general method for the resummation of event-shape distributions in e^+e^- annihilation*, *JHEP* **05** (2015) 102, [arXiv:1412.2126].

[10] A. Banfi, F. A. Dreyer, and P. F. Monni, *Higher-order non-global logarithms from jet calculus*, *JHEP* **03** (2022) 135, [arXiv:2111.02413].

[11] A. J. Larkoski, I. Moult, and B. Nachman, *Jet Substructure at the Large Hadron Collider: A Review of Recent Advances in Theory and Machine Learning*, *Phys. Rept.* **841** (2020) 1–63, [arXiv:1709.04464].

[12] G. Luisoni and S. Marzani, *QCD resummation for hadronic final states*, *J. Phys. G* **42** (2015), no. 10 103101, [arXiv:1505.04084].

[13] Y. L. Dokshitzer, V. A. Khoze, A. H. Mueller, and S. I. Troian, *Basics of perturbative QCD*. Editions Frontieres, 1991.

[14] S. Catani, B. R. Webber, and G. Marchesini, *QCD coherent branching and semiinclusive processes at large x*, *Nucl. Phys. B* **349** (1991) 635–654.

[15] S. Catani and M. H. Seymour, *A General algorithm for calculating jet cross-sections in NLO QCD*, *Nucl. Phys. B* **485** (1997) 291–419, [hep-ph/9605323]. [Erratum: Nucl.Phys.B 510, 503–504 (1998)].

[16] G. P. Salam, *Towards Jetography*, *Eur. Phys. J. C* **67** (2010) 637–686, [arXiv:0906.1833].

[17] R. Kogler et al., *Jet Substructure at the Large Hadron Collider: Experimental Review*, *Rev. Mod. Phys.* **91** (2019), no. 4 045003, [arXiv:1803.06991].

[18] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet User Manual*, *Eur. Phys. J. C* **72** (2012) 1896, [arXiv:1111.6097].

[19] M. Dasgupta and G. P. Salam, *Resummation of nonglobal QCD observables*, *Phys. Lett. B* **512** (2001) 323–330, [hep-ph/0104277].

[20] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, *Soft Drop*, *JHEP* **05** (2014) 146, [arXiv:1402.2657].

[21] P. Caucal, A. Soto-Ontoso, and A. Takacs, *Dynamical Grooming meets LHC data*, *JHEP* **07** (2021) 020, [arXiv:2103.06566].

[22] Y. Mehtar-Tani, A. Soto-Ontoso, and K. Tywoniuk, *Dynamical grooming of QCD jets*, *Phys. Rev. D* **101** (2020), no. 3 034004, [arXiv:1911.00375].

[23] Y. Mehtar-Tani, A. Soto-Ontoso, and K. Tywoniuk, *Tagging boosted hadronic objects with dynamical grooming*, *Phys. Rev. D* **102** (2020) 114013, [arXiv:2005.07584].

[24] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, *JHEP* **04** (2008) 063, [arXiv:0802.1189].

[25] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, *Better jet clustering algorithms*, *JHEP* **08** (1997) 001, [hep-ph/9707323].

[26] M. Wobisch and T. Wengler, *Hadronization corrections to jet cross-sections in deep inelastic scattering*, in *Workshop on Monte Carlo Generators for HERA Physics (Plenary Starting Meeting)*, pp. 270–279, 4, 1998. hep-ph/9907280.

[27] A. J. Larkoski and J. Thaler, *Unsafe but Calculable: Ratios of Angularities in Perturbative QCD*, *JHEP* **09** (2013) 137, [arXiv:1307.1699].

[28] A. J. Larkoski, S. Marzani, and J. Thaler, *Sudakov Safety in Perturbative QCD*, *Phys. Rev. D* **91** (2015), no. 11 111501, [arXiv:1502.01719].

[29] A. Banfi, G. P. Salam, and G. Zanderighi, *Phenomenology of event shapes at hadron colliders*, *JHEP* **06** (2010) 038, [arXiv:1001.4082].

[30] **ALICE** Collaboration, *Measurements of the groomed jet radius and momentum splitting fraction with the soft drop and dynamical grooming algorithms in pp collisions at* $\sqrt{s} = 5.02$ *TeV*, arXiv:2204.10246.

[31] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159–177, [arXiv:1410.3012].

[32] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, *Eur. Phys. J. C* **76** (2016), no. 4 196, [arXiv:1512.01178].

[33] A. Banfi, G. P. Salam, and G. Zanderighi, *Infrared safe definition of jet flavor*, *Eur. Phys. J. C* **47** (2006) 113–124, [hep-ph/0601139].

[34] A. J. Larkoski and E. M. Metodiev, *A Theory of Quark vs. Gluon Discrimination*, *JHEP* **10** (2019) 014, [arXiv:1906.01639].

[35] S. Caletti, A. J. Larkoski, S. Marzani, and D. Reichelt, *A fragmentation approach to jet flavor*, *JHEP* **10** (2022) 158, [arXiv:2205.01117].

[36] F. A. Dreyer, G. Soyez, and A. Takacs, *Quarks and gluons in the Lund plane*, *JHEP* **08** (2022) 177, [arXiv:2112.09140].

[37] M. Dasgupta, F. A. Dreyer, K. Hamilton, P. F. Monni, G. P. Salam, and G. Soyez, *Parton showers beyond leading logarithmic accuracy*, *Phys. Rev. Lett.* **125** (2020), no. 5 052002, [arXiv:2002.11114].

[38] A. Lifson, G. P. Salam, and G. Soyez, *Calculating the primary Lund Jet Plane density*, *JHEP* **10** (2020) 170, [arXiv:2007.06578].

[39] F. A. Dreyer, G. P. Salam, and G. Soyez, *The Lund Jet Plane*, *JHEP* **12** (2018) 064, [arXiv:1807.04758].

[40] F. A. Dreyer and H. Qu, *Jet tagging in the Lund plane with graph networks*, *JHEP* **03** (2021) 052, [arXiv:2012.08526].

[41] T. Becher, A. Broggio, and A. Ferroglia, *Introduction to Soft-Collinear Effective Theory*, vol. 896. Springer, 2015.

[42] I. Z. Rothstein and I. W. Stewart, *An Effective Field Theory for Forward Scattering and Factorization Violation*, *JHEP* **08** (2016) 025, [arXiv:1601.04695].

[43] A. Idilbi and A. Majumder, *Extending Soft-Collinear-Effective-Theory to describe hard jets in dense QCD media*, *Phys. Rev. D* **80** (2009) 054022, [arXiv:0808.1087].

[44] G. Ovanesyan and I. Vitev, *An effective theory for jet propagation in dense QCD matter: jet broadening and medium-induced bremsstrahlung*, *JHEP* **06** (2011) 080, [arXiv:1103.1074].

[45] V. Vaidya, *Effective Field Theory for jet substructure in heavy ion collisions*, *JHEP* **11** (2021) 064, [arXiv:2010.00028].

[46] M. L. Bellac, *Thermal Field Theory*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 3, 2011.

[47] M. Laine and A. Vuorinen, *Basics of Thermal Field Theory*, vol. 925. Springer, 2016.

[48] C. Gale, S. Jeon, and B. Schenke, *Hydrodynamic Modeling of Heavy-Ion Collisions*, *Int. J. Mod. Phys. A* **28** (2013) 1340011, [arXiv:1301.5893].

[49] N. Armesto and E. Scomparin, *Heavy-ion collisions at the Large Hadron Collider: a review of the results from Run 1*, *Eur. Phys. J. Plus* **131** (2016), no. 3 52, [arXiv:1511.02151].

[50] P. Romatschke and U. Romatschke, *Relativistic Fluid Dynamics In and Out of Equilibrium*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 5, 2019.

[51] J. Ghiglieri and D. Teaney, *Parton energy loss and momentum broadening at NLO in high temperature QCD plasmas*, *Int. J. Mod. Phys. E* **24** (2015), no. 11 1530013, [arXiv:1502.03730].

[52] S. Schlichting and D. Teaney, *The First fm/c of Heavy-Ion Collisions*, *Ann. Rev. Nucl. Part. Sci.* **69** (2019) 447–476, [arXiv:1908.02113].

[53] J. Berges, M. P. Heller, A. Mazeliauskas, and R. Venugopalan, *QCD thermalization: Ab initio approaches and interdisciplinary connections*, *Rev. Mod. Phys.* **93** (2021), no. 3 035003, [arXiv:2005.12299].

[54] J.-P. Blaizot and M. A. Escobedo, *Quantum and classical dynamics of heavy quarks in a quark-gluon plasma*, *JHEP* **06** (2018) 034, [arXiv:1711.10812].

[55] V. Vaidya and X. Yao, *Transverse momentum broadening of a jet in quark-gluon plasma: an open quantum system EFT*, *JHEP* **10** (2020) 024, [arXiv:2004.11403].

[56] R. Baier, Y. L. Dokshitzer, A. H. Mueller, S. Peigne, and D. Schiff, *Radiative energy loss of high-energy quarks and gluons in a finite volume quark - gluon plasma*, *Nucl. Phys. B* **483** (1997) 291–320, [hep-ph/9607355].

[57] B. Zakharov, *Fully quantum treatment of the Landau-Pomeranchuk-Migdal effect in QED and QCD*, *JETP Lett.* **63** (1996) 952–957, [hep-ph/9607440].

[58] J.-P. Blaizot, F. Dominguez, E. Iancu, and Y. Mehtar-Tani, *Medium-induced gluon branching*, *JHEP* **01** (2013) 143, [arXiv:1209.4585].

[59] S. J. Brodsky, H.-C. Pauli, and S. S. Pinsky, *Quantum chromodynamics and other field theories on the light cone*, *Phys. Rept.* **301** (1998) 299–486, [hep-ph/9705477].

[60] F. Gelis and Y. Mehtar-Tani, *Gluon propagation inside a high-energy nucleus*, *Phys. Rev. D* **73** (2006) 034019, [hep-ph/0512079].

[61] P. Aurenche, F. Gelis, and H. Zaraket, *A Simple sum rule for the thermal gluon spectral function and applications*, *JHEP* **05** (2002) 043, [hep-ph/0204146].

[62] S. Caron-Huot, *O(g) plasma effects in jet quenching*, *Phys. Rev. D* **79** (2009) 065039, [arXiv:0811.1603].

[63] G. D. Moore and N. Schlusser, *The nonperturbative contribution to asymptotic masses*, *Phys. Rev. D* **102** (2020), no. 9 094512, [arXiv:2009.06614].

[64] G. D. Moore, S. Schlichting, N. Schlusser, and I. Soudi, *Non-perturbative determination of collisional broadening and medium induced radiation in QCD plasmas*, *JHEP* **10** (2021) 059, [arXiv:2105.01679].

[65] M. Gyulassy and X.-n. Wang, *Multiple collisions and induced gluon Bremsstrahlung in QCD*, *Nucl. Phys. B* **420** (1994) 583–614, [nucl-th/9306003].

[66] Y. Mehtar-Tani and K. Tywoniuk, *Radiative energy loss of neighboring subjets*, *Nucl. Phys. A* **979** (2018) 165–203, [arXiv:1706.06047].

[67] P. Arnold, T. Gorda, and S. Iqbal, *The LPM effect in sequential bremsstrahlung: incorporation of "instantaneous" interactions for QCD*, *JHEP* **11** (2022) 130, [arXiv:2209.03971].

[68] J. a. Barata, Y. Mehtar-Tani, A. Soto-Ontoso, and K. Tywoniuk, *Revisiting transverse momentum broadening in dense QCD media*, *Phys. Rev. D* **104** (2021), no. 5 054047, [arXiv:2009.13667].

[69] J. H. Isaksen and K. Tywoniuk, *Wilson line correlators beyond the large-N_c*, *JHEP* **21** (2020) 125, [arXiv:2107.02542].

[70] B. G. Zakharov, *Radiative energy loss of high-energy quarks in finite size nuclear matter and quark - gluon plasma*, *JETP Lett.* **65** (1997) 615–620, [hep-ph/9704255].

[71] U. A. Wiedemann, *Gluon radiation off hard quarks in a nuclear environment: Opacity expansion*, *Nucl. Phys. B* **588** (2000) 303–344, [hep-ph/0005129].

[72] P. B. Arnold, G. D. Moore, and L. G. Yaffe, *Photon and gluon emission in relativistic plasmas*, *JHEP* **06** (2002) 030, [hep-ph/0204343].

[73] Y. Mehtar-Tani and K. Tywoniuk, *Improved opacity expansion for medium-induced parton splitting*, *JHEP* **06** (2020) 187, [arXiv:1910.02032].

[74] J. a. Barata, Y. Mehtar-Tani, A. Soto-Ontoso, and K. Tywoniuk, *Medium-induced radiative kernel with the Improved Opacity Expansion*, *JHEP* **09** (2021) 153, [arXiv:2106.07402].

[75] J. H. Isaksen, A. Takacs, and K. Tywoniuk, *A unified picture of medium-induced radiation*, *JHEP* **02** (2023) 156, [arXiv:2206.02811].

[76] M. Gyulassy, P. Levai, and I. Vitev, *Reaction operator approach to nonAbelian energy loss*, *Nucl. Phys. B* **594** (2001) 371–419, [nucl-th/0006010].

[77] X. Feal and R. Vazquez, *Intensity of gluon bremsstrahlung in a finite plasma*, *Phys. Rev. D* **98** (2018), no. 7 074029, [arXiv:1811.01591].

[78] C. Andres, L. Apolinário, and F. Dominguez, *Medium-induced gluon radiation with full resummation of multiple scatterings for realistic parton-medium interactions*, *JHEP* **07** (2020) 114, [arXiv:2002.01517].

[79] S. Schlichting and I. Soudi, *Splitting rates in QCD plasmas from a nonperturbative determination of the momentum broadening kernel C(q⊥)*, *Phys. Rev. D* **105** (2022), no. 7 076002, [arXiv:2111.13731].

[80] Y. Mehtar-Tani, *Gluon bremsstrahlung in finite media beyond multiple soft scattering approximation*, *JHEP* **07** (2019) 057, [arXiv:1903.00506].

[81] P. Arnold, T. Gorda, and S. Iqbal, *The LPM effect in sequential bremsstrahlung: nearly complete results for QCD*, *JHEP* **11** (2020) 053, [arXiv:2007.15018].

[82] M. D. Sievert, I. Vitev, and B. Yoon, *A complete set of in-medium splitting functions to any order in opacity*, *Phys. Lett. B* **795** (2019) 502–510, [arXiv:1903.06170].

[83] M. Attems, J. Brewer, G. M. Innocenti, A. Mazeliauskas, S. Park, W. van der Schee, and U. A. Wiedemann, *The medium-modified $g \to c\bar{c}$ splitting function in the BDMPS-Z formalism*, *JHEP* **01** (2023) 080, [arXiv:2203.11241].

[84] M. Djordjevic and M. Gyulassy, *Heavy quark radiative energy loss in QCD matter*, *Nucl. Phys. A* **733** (2004) 265–298, [nucl-th/0310076].

[85] R. M. Yazdi, S. Shi, C. Gale, and S. Jeon, *Leading order, next-to-leading order, and nonperturbative parton collision kernels: Effects in static and evolving media*, *Phys. Rev. C* **106** (2022), no. 6 064902, [arXiv:2206.05855].

[86] P. Caucal, E. Iancu, A. H. Mueller, and G. Soyez, *Nuclear modification factors for jet fragmentation*, *JHEP* **10** (2020) 204, [arXiv:2005.05852].

[87] S. P. Adhya, C. A. Salgado, M. Spousta, and K. Tywoniuk, *Medium-induced cascade in expanding media*, *JHEP* **07** (2020) 150, [arXiv:1911.12193].

[88] A. V. Sadofyev, M. D. Sievert, and I. Vitev, *Ab initio coupling of jets to collective flow in the opacity expansion approach*, *Phys. Rev. D* **104** (2021), no. 9 094044, [arXiv:2104.09513].

[89] J. a. Barata, A. V. Sadofyev, and C. A. Salgado, *Jet broadening in dense inhomogeneous matter*, *Phys. Rev. D* **105** (2022), no. 11 114010, [arXiv:2202.08847].

[90] C. Andres, F. Dominguez, A. V. Sadofyev, and C. A. Salgado, *Jet broadening in flowing matter: Resummation*, *Phys. Rev. D* **106** (2022), no. 7 074023, [arXiv:2207.07141].

[91] J.-P. Blaizot and Y. Mehtar-Tani, *Renormalization of the jet-quenching parameter*, *Nucl. Phys. A* **929** (2014) 202–229, [arXiv:1403.2323].

[92] J.-P. Blaizot, F. Dominguez, E. Iancu, and Y. Mehtar-Tani, *Probabilistic picture for medium-induced jet evolution*, *JHEP* **06** (2014) 075, [arXiv:1311.5823].

[93] J.-P. Blaizot, E. Iancu, and Y. Mehtar-Tani, *Medium-induced QCD cascade: democratic branching and wave turbulence*, *Phys. Rev. Lett.* **111** (2013) 052001, [arXiv:1301.6102].

[94] J.-P. Blaizot, Y. Mehtar-Tani, and M. A. C. Torres, *Angular structure of the in-medium QCD cascade*, *Phys. Rev. Lett.* **114** (2015), no. 22 222002, [arXiv:1407.0326].

[95] J.-P. Blaizot and Y. Mehtar-Tani, *Energy flow along the medium-induced parton cascade*, *Annals Phys.* **368** (2016) 148–176, [arXiv:1501.03443].

[96] P. Caucal, E. Iancu, A. H. Mueller, and G. Soyez, *Vacuum-like jet fragmentation in a dense QCD medium*, *Phys. Rev. Lett.* **120** (2018) 232001, [arXiv:1801.09703].

[97] K. Kutak, W. Placzek, and R. Straka, *Solutions of evolution equations for medium-induced QCD cascades*, *Eur. Phys. J. C* **79** (2019), no. 4 317, [arXiv:1811.06390].

[98] E. Blanco, K. Kutak, W. Placzek, M. Rohrmoser, and R. Straka, *Medium induced QCD cascades: broadening and rescattering during branching*, *JHEP* **04** (2021) 014, [arXiv:2009.03876].

[99] P. Arnold and S. Iqbal, *The LPM effect in sequential bremsstrahlung*, *JHEP* **04** (2015) 070, [arXiv:1501.04964]. [Erratum: JHEP 09, 072 (2016)].

[100] P. Arnold, H.-C. Chang, and S. Iqbal, *The LPM effect in sequential bremsstrahlung 2: factorization*, *JHEP* **09** (2016) 078, [arXiv:1605.07624].

[101] P. Arnold, H.-C. Chang, and S. Iqbal, *The LPM effect in sequential bremsstrahlung: 4-gluon vertices*, *JHEP* **10** (2016) 124, [arXiv:1608.05718].

[102] P. Arnold, O. Elgedawy, and S. Iqbal, *Are gluon showers inside a quark-gluon plasma strongly coupled? a theorist's test*, arXiv:2212.08086.

[103] P. Arnold, O. Elgedawy, and S. Iqbal, *The LPM effect in sequential bremsstrahlung: gluon shower development*, arXiv:2302.10215.

[104] Y. Mehtar-Tani, S. Schlichting, and I. Soudi, *Jet thermalization in QCD kinetic theory*, arXiv:2209.10569.

[105] M. Connors, C. Nattrass, R. Reed, and S. Salur, *Jet measurements in heavy ion physics*, *Rev. Mod. Phys.* **90** (2018) 025005, [arXiv:1705.01974].

[106] L. Cunqueiro and A. M. Sickles, *Studying the QGP with Jets at the LHC and RHIC*, *Prog. Part. Nucl. Phys.* **124** (2022) 103940, [arXiv:2110.14490].

[107] Y. Mehtar-Tani, C. A. Salgado, and K. Tywoniuk, *Anti-angular ordering of gluon radiation in QCD media*, *Phys. Rev. Lett.* **106** (2011) 122002, [arXiv:1009.2965].

[108] Y. Mehtar-Tani, C. A. Salgado, and K. Tywoniuk, *Jets in QCD Media: From Color Coherence to Decoherence*, *Phys. Lett. B* **707** (2012) 156–159, [arXiv:1102.4317].

[109] Y. Mehtar-Tani, C. A. Salgado, and K. Tywoniuk, *The Radiation pattern of a QCD antenna in a dense medium*, *JHEP* **10** (2012) 197, [arXiv:1205.5739].

[110] J. Casalderrey-Solana and E. Iancu, *Interference effects in medium-induced gluon radiation*, *JHEP* **08** (2011) 015, [arXiv:1105.1760].

[111] J. Casalderrey-Solana, Y. Mehtar-Tani, C. A. Salgado, and K. Tywoniuk, *New picture of jet quenching dictated by color coherence*, *Phys. Lett. B* **725** (2013) 357–360, [arXiv:1210.7765].

[112] P. Caucal, *Jet evolution in a dense QCD medium*. PhD thesis, Saclay, 9, 2020. arXiv:2010.02874.

[113] Y. Mehtar-Tani and K. Tywoniuk, *Sudakov suppression of jets in QCD media*, *Phys. Rev. D* **98** (2018), no. 5 051501, [arXiv:1707.07361].

[114] J. Casalderrey-Solana, D. C. Gulhan, J. G. Milhano, D. Pablos, and K. Rajagopal, *A Hybrid Strong/Weak Coupling Approach to Jet Quenching*, *JHEP* **10** (2014) 019, [arXiv:1405.3864]. [Erratum: JHEP 09, 175 (2015)].

[115] J. H. Putschke et al., *The JETSCAPE framework*, arXiv:1903.07706.

[116] A. Takacs and K. Tywoniuk, *Quenching effects in the cumulative jet spectrum*, *JHEP* **10** (2021) 038, [arXiv:2103.14676].

[117] R. Baier, Y. L. Dokshitzer, A. H. Mueller, and D. Schiff, *Quenching of hadron spectra in media*, *JHEP* **09** (2001) 033, [hep-ph/0106347].

[118] C. A. Salgado and U. A. Wiedemann, *Calculating quenching weights*, *Phys. Rev. D* **68** (2003) 014008, [hep-ph/0302184].

[119] Y. Mehtar-Tani, D. Pablos, and K. Tywoniuk, *Cone-Size Dependence of Jet Suppression in Heavy-Ion Collisions*, *Phys. Rev. Lett.* **127** (2021), no. 25 252301, [arXiv:2101.01742].

[120] P. Caucal, A. Soto-Ontoso, and A. Takacs, *Dynamically groomed jet radius in heavy-ion collisions*, *Phys. Rev. D* **105** (2022), no. 11 114046, [arXiv:2111.14768].

[121] J. Casalderrey-Solana, G. Milhano, D. Pablos, and K. Rajagopal, *Modification of Jet Substructure in Heavy Ion Collisions as a Probe of the Resolution Length of Quark-Gluon Plasma*, *JHEP* **01** (2020) 044, [arXiv:1907.11248].

[122] J. Casalderrey-Solana, J. G. Milhano, D. Pablos, K. Rajagopal, and X. Yao, *Jet Wake from Linearized Hydrodynamics*, *JHEP* **05** (2021) 230, [arXiv:2010.01140].

[123] S. P. Adhya, K. Kutak, W. Płaczek, M. Rohrmoser, and K. Tywoniuk, *Transverse momentum broadening of medium-induced cascades in expanding media*, arXiv:2211.15803.

[124] Y. Fu, J. Casalderrey-Solana, and X.-N. Wang, *Asymmetric transverse momentum broadening in an inhomogeneous medium*, arXiv:2204.05323.

[125] M. Attems, J. Brewer, G. M. Innocenti, A. Mazeliauskas, S. Park, W. van der Schee, and U. Wiedemann, *Medium-enhanced $c\bar{c}$ radiation*, arXiv:2209.13600.

[126] L. Cunqueiro, D. Napoletano, and A. Soto-Ontoso, *Dead-cone searches in heavy-ion collisions using the jet tree*, arXiv:2211.11789.

[127] **ATLAS** Collaboration, G. Aad et al., *Transverse momentum and process dependent azimuthal anisotropies in $\sqrt{s_{\rm NN}} = 8.16$ TeV p+Pb collisions with the ATLAS detector*, *Eur. Phys. J. C* **80** (2020), no. 1 73, [arXiv:1910.13978].

[128] **ALICE** Collaboration, *Azimuthal anisotropy of jet particles in p-Pb and Pb-Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV*, arXiv:2212.12609.

[129] W. Ke and I. Vitev, *Understanding parton evolution in matter from renormalization group analysis*, arXiv:2301.11940.

[130] M. van Beekveld, S. Ferrario Ravasio, K. Hamilton, G. P. Salam, A. Soto-Ontoso, G. Soyez, and R. Verheyen, *PanScales showers for hadron collisions: all-order validation*, *JHEP* **11** (2022) 020, [arXiv:2207.09467].

[131] F. Herren, S. Höche, F. Krauss, D. Reichelt, and M. Schoenherr, *A new approach to color-coherent parton evolution*, arXiv:2208.06057.

Chapter 5

Scientific results

Article I.

5.1 Dynamical Grooming meets LHC data

Paul Caucal, Alba Soto-Ontoso and Adam Takacs

JHEP, **07**, 020 (2021)

Paper about jet substructure in pp collision. Keywords: jets, substructure, dynamical grooming, logarithmic counting, matching

Published for SISSA by Springer

Received: March 17, 2021
Accepted: June 16, 2021
Published: July 6, 2021

Dynamical Grooming meets LHC data

Paul Caucal,[a,b] **Alba Soto-Ontoso**[a] **and Adam Takacs**[c,d]

[a] *Institut de Physique Théorique, Université Paris-Saclay, CNRS, CEA, F-91191, Gif-sur-Yvette, France*

[b] *Physics Department, Brookhaven National Laboratory, Upton, NY 11973, U.S.A.*

[c] *Department of Physics and Technology, University of Bergen, Bergen 5020, Norway*

[d] *Department of Astronomy and Theoretical Physics, Lund University, S-223 62 Lund, Sweden*

E-mail: pcaucal@bnl.gov, alba.soto@ipht.fr, adam.takacs@uib.no

Abstract: In this work, we analyse the all-orders resummation structure of the momentum sharing fraction, z_g, opening angle, θ_g, and relative transverse momentum, $k_{t,g}$, of the splitting tagged by the Dynamical Grooming procedure in hadronic collisions. We demonstrate that their resummation does non-exponentiate and it is free of clustering logarithms. Then, we analytically compute the probability distributions of $(z_g, \theta_g, k_{t,g})$ up to next-to-next-to-double logarithm accuracy ($\mathrm{N^2DL}$) in the narrow jet limit, including a matching to leading order in α_s. On the phenomenological side, we perform an analytic-to-parton level comparison with Pythia and Herwig. We find that differences between the analytic and the Monte-Carlo results are dominated by the infra-red regulator of the parton shower. Further, we present the first analytic comparison to preliminary ALICE data and highlight the role of non-perturbative corrections in such low-p_t regime. Once the analytic result is corrected by a phenomenologically determined non-perturbative factor, we find very good agreement with the data.

Keywords: Perturbative QCD, Resummation

ArXiv ePrint: 2103.06566

https://doi.org/10.1007/JHEP07(2021)020

JHEP07(2021)020

Contents

1 Introduction 1

2 Theoretical analysis of dynamically groomed observables 3
- 2.1 Double-logarithmic estimation and basic properties 4
- 2.2 Revisiting the meaning of accuracy: from IRC to Sudakov safe observables 7
- 2.3 $k_{t,g}$ at LO+N^{2}DL accuracy 10
  - 2.3.1 Resummation 10
  - 2.3.2 Matching to fixed-order 15
  - 2.3.3 Results 17
- 2.4 z_g, θ_g at LO+N^{2}DL accuracy 19
  - 2.4.1 Boundary logarithms for the θ_g distribution 19
  - 2.4.2 Comparison between the resummation structure of Soft Drop and Dynamical Grooming 21
  - 2.4.3 Matching to fixed-order 22
  - 2.4.4 Results 25

3 Phenomenology at LHC energies 26
- 3.1 Analytics vs. Monte-Carlo parton level 26
- 3.2 Comparison to preliminary ALICE data 29

4 Conclusions and outlook 32

A Details of analytic calculations at N^{2}DL' 34

B The size of finite z corrections in the definition of κ 37

C Impact of jet clustering algorithms 39

D Non-perturbative corrections with Pythia and Herwig 39

E Monte-Carlo description of $(z_g, \theta_g, k_{t,g})$ data 42

1 Introduction

Jet physics aims at pinning down the microscopic properties of Quantum Chromodynamics (QCD [1]. In the context of heavy-ion physics, the modification of jets with respect to their vacuum counterparts is regarded as an experimental evidence for the formation of a dense, thermal medium, namely the Quark-Gluon plasma [2].

JHEP07(2021)020

Nowadays, most of the efforts in the field from a theoretical point of view, both from an analytic perspective and with machine learning tools (see ref. [3] for a review), are directed towards studying the space-time structure of a jet by characterising its radiation pattern through jet substructure observables, i.e. constructed from one (e.g. z_g [4, 5]), or a few branchings (e.g. N-subjettiness [6] or the Lund jet plane [7, 8]) at most. In this paper we focus on the former category where typically the one branching that defines the observable is selected in a region of phase space where perturbative QCD calculations are applicable, that is, far away from the soft and wide angle sector. This tagging task is handled by so-called 'grooming methods' through which the hard and collinear core of the jet is isolated. The way this general goal is achieved differs from one groomer to the other, e.g. Modified Mass Drop Tagger(mMDT) [9] or its extension Soft Drop (SD) [4] selects the splitting whose momentum sharing fraction obeys $z > z_{\rm cut}\theta^{\beta}$, while trimming [10] first reclusters a jet into subjets with a smaller radius $R_{\rm sub}$ and then keep only those subjets whose $p_t^{\rm subjet} > z_{\rm cut} p_t^{\rm jet}$. In the previous expressions ($z_{\rm cut}, \beta$ and $R_{\rm sub}$) are free parameters that need to be tuned with Monte-Carlo simulations to achieve an optimal performance [11]. These methodological differences leave their imprint into the analytic behavior of the observables that they define [12, 13]. For example, as we shall see in more detail in what follows, due to the presence of an explicit $z_{\rm cut}$ in the Soft Drop grooming condition this method is free of non-global logarithms in the resummation function. This fact has enable to push the accuracy of the calculation of Soft Drop groomed observables up to next-to-leading log accuracy in $p+p$ [14–21] and even next-to-next-to-leading log in $e^+ + e^-$ [22–24]. It is then clear that the usefulness of a given grooming method should not be judged only on the basis of its resilience to non-perturbative physics, but also on its analytic structure from a pQCD point of view. This paper aims at deepening our analytic understanding of jet substructure observables as defined by a novel grooming technique that has been recently introduced and dubbed 'Dynamical Grooming' (DyG) [25–27].

The Dynamical Grooming method consists in identifying the 'hardest' branching in a jet tree as a proxy for the physical jet scale. The hardness measure is given by

$$\kappa^{(a)} = \frac{1}{p_{t,\rm jet}} z(1-z) p_t \theta^a \tag{1.1}$$

where a is a continuous free parameter that has to be larger than zero in order to guarantee collinear safety. For certain values of a in eq. (1.1), the hardness measure translates into familiar kinematical quantities, e.g $\kappa^{(1)} = k_t$, where k_t is the transverse momentum of the splitting, or $\kappa^{(2)} = m^2$, with m being the branching mass. In addition, we define $\theta = \Delta R/R$ where ΔR is the angular separation between the sub-jets and R corresponds to the cone size. The hardest splitting is obtained after re-clustering the jet sample with Cambridge/Aachen algorithm [28] and finding the node with the largest κ in the clustering sequence. In fact, it can be proven through analytical arguments [29] that it is sufficient to look for the hardest splitting along the primary Lund plane of the jet, i.e. following the branch with the larger transverse momentum at each de-clustering step.[1] First steps towards

[1] We have numerically checked that our results are robust if we look for the hardest splitting in the whole tree and not only on the primary branch.

JHEP07(2021)020

the calculation from first-principles in perturbative QCD of the probability distribution of the momentum sharing fraction, z_g, the mass and the relative transverse momentum, $k_{t,g}$, of the hardest splitting were presented in the original Dynamical Grooming paper in the resummation region, i.e. when $z_g(k_{t,g}) \ll 1$ [25]. Interestingly, it was found that similarly to the Soft Drop case, the z_g distribution pertains to a special class of jet observables known as Sudakov safe [5, 30]. Together with the modified leading-log calculation of DyG jet substructure observables, a Monte-Carlo study of the impact of non-perturbative physics was presented in ref. [25]. An overall similar performance than Soft Drop was shown, but with a remarkable resilience to hadronization in some cases like the z_g distribution as tagged by $a=1$. This novel idea has triggered the interest of the ALICE collaboration that has recently conducted some preliminary measurements on the z_g, θ_g [31], and $k_{t,g}$ [32] distributions at $\sqrt{s} = 5.02\,\text{TeV}$ in the jet transverse momentum bin of $60 < p_t^{ch} < 80\,\text{GeV}$. As we will see, the low p_t reach of the ALICE detector challenges the analytic description of such data set given that non-perturbative effects are sizeable. In addition, first steps towards the experimental use of DyG in heavy-ion collisions were reported in ref. [33].

From an analytic point of view the purposes of this paper are multifold: (i) understand the resummation structure of Dynamical Grooming observables and propose a definition for their logarithmic accuracy which circumvents their non-exponentiating nature (double log, next-to-double log, etc) (ii) advance the resummation of z_g and $k_{t,g}$ from modified-leading logarithm [25] to next-to-next-to double logarithmic accuracy,[2] as well as presenting for the first time the resummation of θ_g, (iii) highlight the absence of clustering logarithms in dynamically groomed observables, (iv) perform a fixed-order matching for all three dynamically groomed observables. This last point is not trivial for the pair of Sudakov safe observables and we propose a novel method to match the resummed and fixed-order distributions. All these ingredients are contained in section 2. After a few sanity checks on the analytic side, in section 3.1 we compare our results to Monte-Carlo simulations at parton level in a high-p_t setup, where non-perturbative effects are mild. Next, in section 3.2, we present the first comparison between an analytic calculation and the preliminary ALICE data for (z_g, θ_g and $k_{t,g}$). In addition, Monte-Carlo studies with different general purpose event generators are performed showing the impact of different details in the definition of the Dynamical Grooming method in appendices B, C. The discriminating power of these type of jet substructure observables with respect to different hadronization models and parton showers are shown in appendix E.

2 Theoretical analysis of dynamically groomed observables

In this section, we present the all-order perturbative calculation of dynamically groomed observables in the $\kappa^{(a)} \ll 1$ region[3] and their matching to fixed order results applicable when $\kappa^{(a)} \sim 1$. In the soft limit, $z \ll 1$ and the $(1-z)$ factor can be removed from eq. (1.1). Furthermore, as the hardest splitting takes place along the primary branch, we neglect

[2] For a precise definition of N^{p}DL accuracy, see eq. (2.26) and the discussion below.

[3] At low enough values of κ the calculation is dominated by non-perturbative effects. Therefore, strictly speaking our resummation is valid when $\kappa_{\text{NP}} \ll \kappa \ll 1$.

JHEP07(2021)020

momentum degradation such that $p_t = p_{t,\text{jet}}$. Therefore, instead of eq. (1.1), the definition $\kappa^{(a)} = z\theta^a$ is adopted throughout this section, and we sometimes omit the a superscript to lighter the notation.[4]

2.1 Double-logarithmic estimation and basic properties

We shortly revisit the baseline calculation performed in ref. [25]. In the $\kappa \ll 1$ limit, the two-dimensional probability distribution of a splitting, with kinematic variables (z, θ), to be hardest in the clustering sequence is given by

$$\frac{\mathrm{d}^2\mathcal{P}_i(z,\theta|a)}{\mathrm{d}\theta \mathrm{d}z} = \widetilde{P}_i(z,\theta)\Delta_i(\kappa|a)\,, \tag{2.1}$$

where i indicates the flavor of the jet initiating parton. The two ingredients entering the right-hand side of the previous equation are actually connected through

$$\ln \Delta_i(\kappa|a) = -\int_0^1 \mathrm{d}z' \int_0^1 \mathrm{d}\theta' \widetilde{P}_i(z',\theta')\Theta\left(z'\theta'^a - \kappa^{(a)}\right)\,. \tag{2.2}$$

In physical terms, the branching kernel, $\widetilde{P}(z,\theta)$, represents the probability of a splitting with (z,θ) to occur, while $\Delta(\kappa|a)$ is the so-called Sudakov form factor and vetoes all harder emissions, i.e. those with $\kappa' > \kappa$. From eq. (2.2), it is easy to see that $a > 0$ is required to regulate the collinear singularity. The normalised probability distribution to measure an observable $\kappa^{(b,c)} = z^b\theta^c$ on the $\kappa^{(a)}$ tagged splitting is given by

$$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\kappa^{(b,c)}}\bigg|_a = \int_0^1 \mathrm{d}\theta \int_0^1 \mathrm{d}z\, \mathcal{P}_i(z,\theta|a)\delta(z^b\theta^c - \kappa^{(b,c)})\,, \tag{2.3}$$

where a sum over flavors including the proper quark/gluon fraction is implicit. The observables that we focus on are obtained from eq. (2.3) by setting: $(b=1, c=0)$ for z_g, $(b=0, c=1)$ for θ_g, and $(b=1, c=1)$ for $k_{t,g}$.

We start by considering branchings in the soft-collinear limit ($z \ll 1$ and $\theta \ll 1$) that generate terms with powers of $\alpha_s \ln^2(\kappa^{(b,c)})$ in eq. (2.1). That is, we achieve double logarithmic accuracy (DLA) in the language of logarithmic resummation, as we will see below. The soft-collinear limit of the branching kernel reads

$$\widetilde{P}_i(z,\theta) = \frac{\alpha_s}{\theta\pi}P_i(z)\,, \tag{2.4}$$

where P_i is the leading-order Altarelli-Parisi splitting function that, in this approximation, is given by

$$P_i(z) = \frac{2C_i}{z}\,, \tag{2.5}$$

with C_i being the color factor of the jet initiator parton; $C_i = C_A$ for gluons, and C_F for quarks. The running of the strong coupling is beyond DLA and therefore we fix to its value at the jet scale, i.e. $\alpha_s \equiv \alpha_s(p_{t,\text{jet}}R)$. In this limit, the Sudakov reduces to

$$\ln \Delta_i(\kappa|a) = -\int_\kappa^1 \mathrm{d}z' \int_{(\kappa/z')^{1/a}}^1 \frac{\mathrm{d}\theta'}{\theta'}\frac{\alpha_s}{\pi}\frac{2C_i}{z} = -\frac{\bar{\alpha}}{a}\ln^2\kappa\,, \tag{2.6}$$

[4] One can rigorously prove that $1-z$ corrections are beyond our targeted accuracy (see appendix B).

JHEP07(2021)020

where $\bar{\alpha}=C_i\alpha_s/\pi$. By plugging eq. (2.6) into eq. (2.3), we obtain the momentum sharing fraction of the tagged splitting

$$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}z_g}=\frac{1}{z_g}\sqrt{\frac{\bar{\alpha}\pi}{a}}\left[\mathrm{erf}\left(\sqrt{\frac{\bar{\alpha}}{a}}\ln z_g\right)+1\right], \tag{2.7}$$

its opening angle

$$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\theta_g}=\frac{1}{\theta_g}\sqrt{\bar{\alpha}\pi a}\left[\mathrm{erf}\left(\sqrt{\bar{\alpha}a}\ln(\theta_g)\right)+1\right], \tag{2.8}$$

and its relative transverse momentum

$$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}k_{t,g}}=\frac{1}{k_{t,g}}\frac{\sqrt{\bar{\alpha}\pi a}}{a-1}\left[\mathrm{erf}\left(\sqrt{\frac{\bar{\alpha}}{a}}\ln(k_{t,g})\right)-\mathrm{erf}\left(\sqrt{\bar{\alpha}a}\ln(k_{t,g})\right)\right]. \tag{2.9}$$

Location of the peak. An important feature of eqs. (2.7)–(2.9) is the value at which they are cut off. Its location can be obtained by taking the derivative of e.g. eq. (2.8)

$$\frac{\mathrm{d}}{\mathrm{d}\theta_g}\left(\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\theta_g}\right)=\sqrt{\bar{\alpha}a}\frac{1}{\theta_g^2}\left[-\sqrt{\pi}\left(1-\mathrm{erf}(\sqrt{x})\right)+\frac{2\sqrt{x}\exp(-x)}{\ln(1/\theta_g)}\right], \tag{2.10}$$

where $x\equiv\bar{\alpha}a\ln^2(1/\theta_g)$. Then, the maximum value of the distribution, $\theta_{\max}$, satisfies the implicit equation

$$\frac{2\sqrt{x}\exp(-x)}{\sqrt{\pi}(1-\mathrm{erf}(\sqrt{x}))}=\ln\left(\frac{1}{\theta_{\max}}\right). \tag{2.11}$$

If $\theta_{\max}\ll 1$, the left hand side can be approximated by its asymptotic behaviour ($x\to\infty$)

$$\frac{2\sqrt{x}\exp(-x)}{\sqrt{\pi}(1-\mathrm{erf}(\sqrt{x}))}\simeq 2x\,, \tag{2.12}$$

such that

$$\ln\left(\frac{1}{\theta_{\max}}\right)=\frac{1}{2a\bar{\alpha}}+\mathcal{O}(1)\,. \tag{2.13}$$

This equation indicates that the smaller the value of a, the deeper the tagged splitting is on the angular ordered shower, i.e. at smaller angles. In other words, at fixed θ_g, larger values of a lead to a bigger Sudakov suppression. Therefore, the distribution shifts to larger θ_g for larger a. Thus, eq. (2.13) confirms and provides an analytic explanation for the result reported in ref. [25] on the location of the tagged branching in the jet tree using Pythia [34] simulations. Notice that, in order to solve the implicit equation for the peak position, we have assumed that $\theta_{\max}\ll 1$. This approximation holds for not too large values of a. Otherwise, the smallness of $\bar{\alpha}$ can be compensated by a in the product $a\bar{\alpha}$ appearing in eq. (2.13) and $\theta_{\max}\sim 1$.

Following similar steps for the maximum of the momentum sharing fraction, we obtain[5]

$$\ln\left(\frac{1}{z_{\max}}\right)=\frac{a}{2\bar{\alpha}}+\mathcal{O}(1)\,. \tag{2.14}$$

Again, the previous expression is valid as long as a is not too small.

[5]Notice that this equation can be also obtained by applying the $a\mapsto 1/a$ transformation in eq. (2.13).

JHEP07(2021)020

Finally, in the case of $k_{t,g}$, we find that

$$\ln\left(\frac{1}{k_{t,\max}}\right) = \frac{a^{\mathrm{sgn}(a-1)}}{2\bar{\alpha}} + \mathcal{O}(1)\,. \tag{2.15}$$

This analytic estimate confirms the ordering observed numerically in figure 9 of ref. [25], i.e. the $a=0.1$ curve is peaked at a smaller k_t than the $a=2$ case is, being $a=1$ the curve peaking at the largest value.

Infra-red and collinear safety. The first step towards boosting the accuracy of our calculation is to analyse the IRC (un)safety of the observables that we are dealing with. As we have already mentioned, and was shown in ref. [25], dynamically groomed observables are collinear unsafe for $a \leq 0$. For $a > 0$, while $k_{t,g}$ is a *standard* IRC safe observable, both z_g and θ_g are Sudakov safe only [5]. This means that the all-order resummation encompassed in the Sudakov form factor regulates the singularities that appear at each order in α_s when $\theta_g \to 0$ or $z_g \to 0$. Notice that for θ_g, this behavior represents a stark difference with respect to Soft Drop grooming, where this observable is, in fact, IRC safe [4]. This can be understood as a result of the z_{cut} that appears in the Soft Drop condition and regulates the soft singularity. In turn, Dynamical Grooming does not introduce any sharp cut-off on the radiation phase-space and thus nothing forbids the hardest splitting to be in the soft ($z \sim 0$) region.

A well-known consequence of Sudakov safety [5, 30] is that the z_g and θ_g-distributions have an ill-defined expansion in (integer) powers of α_s. To illustrate this fact, we introduce the cumulative distribution, that defines the probability to measure an observable below a certain value ν, i.e.

$$\Sigma(\nu) = \int_0^{\nu} \mathrm{d}\nu' \frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\nu'}\,. \tag{2.16}$$

The $k_{t,g}$ cumulative distribution at DLA reads

$$\begin{aligned}\Sigma(k_{t,g}) = \frac{1}{a-1}\Bigg[a\exp\left(-\frac{\bar{\alpha}}{a}\ln^2(k_{t,g})\right) - \exp\left(-\bar{\alpha}a\ln^2(k_{t,g})\right) \\ + \sqrt{\pi a\bar{\alpha}}\ln(k_{t,g})\left[\mathrm{erf}\left(\sqrt{\frac{\bar{\alpha}}{a}}\ln k_{t,g}\right) - \mathrm{erf}\left(\sqrt{\bar{\alpha}a}\ln k_{t,g}\right)\right]\Bigg]\,,\end{aligned} \tag{2.17}$$

and its expansion in α_s (or equivalently in $\bar{\alpha}$)

$$\Sigma(k_{t,g}) = 1 - \bar{\alpha}\ln^2\left(\frac{1}{k_{t,g}}\right) + \frac{1+a+a^2}{6a}\bar{\alpha}^2\ln^4\left(\frac{1}{k_{t,g}}\right) + \mathcal{O}(\bar{\alpha}^3)\,. \tag{2.18}$$

From the previous expression, it is clear that $\Sigma(k_{t,g})$ admits an analytic expansion in $\bar{\alpha}$, as it is expected for an IRC safe observable.

In contrast, the θ_g-cumulative distribution is

$$\Sigma(\theta_g) = \exp\left(-\bar{\alpha}a\ln^2\left(\frac{1}{\theta_g}\right)\right) - \sqrt{\bar{\alpha}\pi a}\ln\left(\frac{1}{\theta_g}\right)\left[\mathrm{erf}\left(-\sqrt{\bar{\alpha}a}\ln\left(\frac{1}{\theta_g}\right)\right) + 1\right]\,, \tag{2.19}$$

JHEP07(2021)020

and its expansion in powers of $\bar{\alpha}$ is given by

$$\Sigma(\theta_g) = 1 - \sqrt{\bar{\alpha}\pi a}\ln\left(\frac{1}{\theta_g}\right) + \bar{\alpha}a\ln^2\left(\frac{1}{\theta_g}\right) + \mathcal{O}(\bar{\alpha}^2)\,. \tag{2.20}$$

In this case, the resumming function is not analytic as the second term in eq. (2.20) is of order $\sqrt{\bar{\alpha}}$. The non-analyticity on the dynamically groomed θ_g is caused uniquely by the $\sqrt{\bar{\alpha}}$ term. That is, all other powers of $\bar{\alpha}$ appearing in eq. (2.20) are integer and the function

$$\Sigma(\theta_g) + \sqrt{\bar{\alpha}\pi a}\ln\left(\frac{1}{\theta_g}\right) \tag{2.21}$$

is, in fact, analytic. The same arguments apply to z_g where again one can utilise the $a \mapsto 1/a$ transformation, to confirm that

$$\Sigma(z_g) + \sqrt{\frac{\bar{\alpha}\pi}{a}}\ln\left(\frac{1}{z_g}\right) \tag{2.22}$$

has an analytic dependence on $\bar{\alpha}$ at any perturbative order. Interestingly, this α_s-expansion of z_g is remarkably different from its Soft Drop counterpart when $\beta > 0$. For Soft Drop, the expansion is driven by $\alpha_s^{n/2}$ terms where the integer $n \geq 1$ [5]. Whether this is a purely mathematical statement, or an explanation in physical terms exists, is beyond our degree of understanding and further work is required to clarify it.

To sum up, in this section we have shown that the opening angle and momentum sharing fraction of the splitting tagged by Dynamical Grooming are unconventional observables from a pQCD point of view. The Sudakov safety of the z_g and θ_g distributions leads to an ambiguous definition of the logarithmic accuracy in their resummation, as was noted in ref. [30]. Furthermore, the standard matching to fixed-order calculations is not trivial due to the non-analyticity of the resummed result. In this context, a careful definition of logarithmic accuracy in the resummation is required and will be provided next.

2.2 Revisiting the meaning of accuracy: from IRC to Sudakov safe observables

We start by considering a general resummed formula for an IRC safe distribution obtained with Dynamical Grooming. Following our previous notation, we denote $\kappa^{(b,c)}$ the observable that we measure on the splitting whose hardness, $\kappa^{(a)} = z\theta^a$, is the largest in the shower. The cumulative distribution to measure $\kappa^{(b,c)} \ll 1$ reads

$$\Sigma(\kappa^{(b,c)}) = \int_0^1 \mathrm{d}z \int_0^1 \mathrm{d}\theta\, \widetilde{P}(z,\theta)\Delta(\kappa|a)\Theta(\kappa^{(b,c)} - z^b\theta^c)\,, \tag{2.23}$$

where we have omitted the flavour index for simplicity. An important comment regarding the values of (a, b, c) is in order. Only when $b \leq 1$ and $c \leq a$, the hierarchy $\kappa^{(a)} \leq \kappa^{(b,c)} \leq 1$ is satisfied and thus $\Delta(\kappa|a)$ can be accurately computed through resummation techniques. Any other combination of (a, b, c) leads to a situation in which $\kappa^{(b,c)} \ll 1$ does not necessarily imply $\kappa^a \ll 1$ such that fixed order contributions to $\Delta(\kappa|a)$ become relevant.

JHEP07(2021)020

Having these constraints in mind, we derive the Sudakov safe distributions of $z_g \equiv \kappa^{(1,0)}$ and $\theta_g \equiv \kappa^{(0,1)}$ as two limits of eq. (2.23), i.e.

$$\Sigma(z_g) = \lim_{c\to 0} \Sigma(\kappa^{(1,c)}) \,, \tag{2.24}$$

and

$$\Sigma(\theta_g) = \lim_{b\to 0} \Sigma(\kappa^{(b,1)}) \,. \tag{2.25}$$

The key point is that we define the accuracy of z_g and θ_g through the accuracy of the IRC safe distribution $\Sigma(\kappa^{(b,c)})$. For instance, we shall state that $\Sigma(z_g)$ is known at DLA, if $\Sigma(\kappa^{(b,c)})$ is known at the same degree of accuracy for all $c > 0$, or at least in the neighbourhood of $c = 0$. Our prescription to define the accuracy of Sudakov safe observables follows the spirit of ref. [30]. However, instead of defining the accuracy of the Sudakov-safe observable by marginalization of an IRC safe double differential distribution, we exploit the IRC safety of the $\kappa^{(b,c)}$ observable itself. It's important to realise that the perturbative expansion of the Sudakov safe observables is only defined after taking first the appropriate limit on $\Sigma(\kappa^{(b,c)})$ as given by eqs. (2.24), (2.25). If these steps are taken in reverse order, i.e. expanding $\Sigma(\kappa^{(b,c)})$ in powers of α_s first and subsequently taking the limit of $b(c) \to 0$, one can show that the correct α_s-expansion, given by eqs. (2.20)–(2.22) at DLA, is not recovered. In short, these two operations do not commute.

The perturbative expansion of $\Sigma(\kappa^{(b,c)})$ can be written as

$$\Sigma(\kappa^{(b,c)}) = \sum_{n=0}^{\infty} \alpha_s^n \sum_{m=0}^{2n} c_{nm} \, \ln^m(\kappa^{(b,c)}) \,, \tag{2.26}$$

where the c_{nm} coefficients have to be determined. Then, we adopt the following convention [35, 36]: the logarithmic accuracy of $\Sigma(\kappa^{(b,c)})$ is said to be N^{p}DL if the c_{nm} coefficients are known for all n and $2n - p \leq m \leq 2n$. Notice that in many other jet substructure calculations it is customary to define the logarithmic accuracy at the level of $\ln \Sigma$ instead of on the cumulative distribution itself. The reason why we use $\Sigma(\kappa^{(b,c)})$ is because, in general, due to the marginalisation procedure stated in eq. (2.23) the resummation of DyG observables does not exponentiate [37, 38] as it clear from eq. (2.17). This no exponentiation property is part of other jet substructure observables such as subjet multiplicities. Yet, there is a specific case for which it does: when $b = 1$ and $c = a$. That is, when the kinematic variable used for tagging coincides with the measured observable. For instance, select the splitting with the largest k_t in the shower, and compute its k_t-distribution. In this case, the cumulative distribution is simply the Sudakov form factor, i.e.

$$\Sigma(\kappa^{(1,a)}) = \Delta(\kappa|a) \tag{2.27}$$

that is equivalent to the plain distribution.

A natural question at this point is how does one relate the c_{nm} coefficients with the accuracy of $\widetilde{P}(z,\theta)$ and $\Delta(\kappa|a)$. In other words, which are the relevant terms that one needs to include in the branching kernel and in the Sudakov form factor in order to reach

JHEP07(2021)020

a given accuracy? To answer this question we rely on the exponentiate property of $\Delta(\kappa|a)$, to write its logarithmic structure in the traditional form[6] [38]

$$\Delta(\kappa|a) = \left(1 + \sum_{n\geq 1} \alpha_s^n C_n\right) e^{\ln(\kappa) g_1(x) + g_2(x) + \alpha_s g_3(x) + \mathcal{O}(\alpha_s^{n+2} \ln^n \kappa)}\,, \tag{2.28}$$

with C_n being constant coefficients, g_i analytic functions

$$g_i(x) = \sum_{i=1}^{\infty} g_{ij} x^j\,, \tag{2.29}$$

and $x \equiv \alpha_s \ln \kappa$. In the N^{p}LL type of counting, the resuming function g_1 would be referred as LL, g_2 as NLL and so on. Our targeted accuracy is N^{2}DL in the rest of the paper, with the possibility of keeping sub-leading terms. After expanding eq. (2.28) in powers of α_s we realize that one has to account for the following g_{nm} coefficients at the corresponding level of accuracy

$$\text{DL}(p=0): \; g_{11}\,, \tag{2.30}$$

$$\text{NDL}(p=1): \; g_{11}, g_{12}, g_{21}\,, \tag{2.31}$$

$$\text{N}^2\text{DL}(p=2): \; g_{11}, g_{12}, g_{13}, g_{21}, g_{22}, C_1\,. \tag{2.32}$$

The g_{11} was already computed in section 2.1 where we accounted for soft and collinear emissions only

$$g_{11} = -\frac{C_i}{a\pi}\,. \tag{2.33}$$

The other coefficients and their physical interpretation are provided in the following section up to N^{2}DL. Given that the constant C_1 term is related to the interplay between the resummation and fixed-order calculations, we postpone its discussion to section 2.3.2 and neglect it in the resummation-related part.

Turning to the terms that are needed in $\widetilde{P}(z,\theta)$, we start by working out the plain case ($b{=}1$ and $c{=}a$). The exponentiation property of the resummation, in this particular case, leads to a one-to-one mapping between the terms in the Sudakov and in the branching kernel. More concretely, following eq. (2.23) one gets

$$-\int_0^1 \mathrm{d}z \int_0^1 \mathrm{d}\theta \widetilde{P}(z,\theta)\Theta(z\theta^a - \kappa) = \ln(\kappa) g_1(\alpha_s \ln\kappa) + g_2(\alpha_s \ln\kappa) + \cdots\,, \tag{2.34}$$

that reduces in the N^{2}DL case to

$$-\int_0^1 \mathrm{d}z \int_0^1 \mathrm{d}\theta \widetilde{P}(z,\theta)\Theta(z\theta^a - \kappa) = \ln(\kappa)(g_{11}x + g_{12}x^2 + g_{13}x^3) + g_{21}x + g_{22}x^2\,, \tag{2.35}$$

where, again, $x \equiv \alpha_s \ln\kappa$. The previous equation, derived exploiting the exponentiation property of the plain case, is sufficient to reach N^{2}DL for all values of (b,c). Using eq. (2.35)

[6] Notice that, in contrast to some cases in the literature, the g-functions contain both collinear and soft, non-global terms, i.e. we do not write a separate $\mathcal{S}$ factor as in [38].

JHEP07(2021)020

with any b and c may produce power suppressed or sub-leading ($p \geq 3$) logarithmic corrections in front of $\alpha_s^n \ln^{2n-2}(\kappa)$ terms, that are nevertheless negligible in the resummation region. The physical insight behind the 'universality' of eq. (2.35) relates to the fact that $\tilde{P}(z,\theta)$ is just a probability to have a splitting with a given z and θ. Thus, the branching kernel should be a priori independent of both a and the observable we measure on this branching.

2.3 $k_{t,g}$ at LO+N^{2}DL accuracy

After this rather formal discussion, we would like to shed light on our statements through an explicit calculation. Namely, we compute the IRC safe $k_{t,g}$ distribution in the small jet radius limit at N^{2}DL accuracy on the resummation side and include its matching to a fixed-order calculation at leading order, thus achieving a solid analytic description for all values of $k_{t,g}$.

2.3.1 Resummation

From the general formula given by eq. (2.23) it is straightforward to calculate the cumulative $k_{t,g}$ distribution by setting $b=c=1$. It reads,

$$\Sigma(k_{t,g}) = \int_0^1 \mathrm{d}z \int_0^1 \mathrm{d}\theta \widetilde{P}(z,\theta) \Delta(\kappa|a) \Theta(k_{t,g} - z\theta) \tag{2.36}$$

such that the differential cross section is

$$\frac{1}{\sigma_0} \frac{\mathrm{d}\sigma}{\mathrm{d}k_{t,g}} = \frac{\mathrm{d}\Sigma(k_{t,g})}{\mathrm{d}k_{t,g}} \tag{2.37}$$

where σ_0 represents the Born level total cross-section. In what follows, we calculate the necessary g_{nm} coefficients that enter in the Sudakov form factor and the branching kernel, see eqs. (2.28), (2.35), and organise them according to the underlying physical effect.

Hard-collinear emissions. Due to its simplicity, the first term that we add to our calculation is the one arising from including hard-collinear corrections ($z \sim 1$, $\theta \ll 1$) in the splitting function. This amounts to take into account the finite part of the splitting functions as follows:

$$P_i^{(\mathrm{h-c})}(z) = \frac{2C_i}{z} \Theta\left(e^{-B_i} - z\right) , \tag{2.38}$$

where $B_q = 2/C_F$, $B_g = 11/12 - n_f T_r/(3C_A)$, $T_r = 1/2$, and we fix the number of flavors to $n_f = 5$. The analytic integration of the new finite piece that appears both in the Sudakov and the branching kernel is useful to illustrate the point about sub-leading terms that appear naturally in the calculation. In fact,

$$-\int_0^1 \mathrm{d}z \int_0^1 \frac{\mathrm{d}\theta}{\theta} \frac{\alpha_s}{\pi} P_i^{(\mathrm{h-c})}(z) \Theta(z\theta^a - \kappa) = -\frac{\alpha_s C_i}{\pi a} \left(B_i + \ln(\kappa)\right)^2 . \tag{2.39}$$

From the previous equation one can easily read off the g_{21} coefficient

$$g_{21} = -\frac{2C_i B_i}{a\pi} , \tag{2.40}$$

JHEP07(2021)020

while the term proportional to B_i and no $\ln(\kappa)$ dependence is sub-leading, although might be large when $a \ll 1$. Strictly speaking, this latter contribution is not required to reach N^{2}DL in our calculation, but we will check its numerical impact by the end of this section. Notice that since there is no soft singularity for flavor switching splittings, they contribute as a power correction to $\kappa^{(a)}$ in our Sudakov form factor and we do not include them here. This argument is valid as long as $\kappa \ll 1$ along the lines of the role played by $y_{\rm cut}$ in appendix B of ref. [12].

Running coupling. Up to now, we have fixed the coupling in order to achieve compact, fully analytic expression. However beyond DLA, the running of the coupling has to be taken into account. At 1-loop in perturbation theory it is given by:

$$\alpha_s^{1\ell}(k_t) = \frac{\alpha_s}{1 + 2\beta_0 \alpha_s \ln\left(\frac{k_t}{Q}\right)} \tag{2.41}$$

$$= \alpha_s \left[1 - 2\beta_0 \alpha_s \ln\left(\frac{k_t}{Q}\right) + 4\beta_0^2 \alpha_s^2 \ln^2\left(\frac{k_t}{Q}\right)\right] + \mathcal{O}(\alpha_s^4)\,, \tag{2.42}$$

with the reference value $\alpha_s \equiv \alpha_s(Q)$ is set at the jet scale $Q \equiv p_{t,\rm jet} R$, $\beta_0 = (11C_A - 4n_f T_r)/(12\pi)$ and $k_t = z\theta p_{t,\rm jet}$.

Next, we integrate analytically the branching kernel with the 1-loop running coupling

$$-\int_0^1 \mathrm{d}z \int_0^1 \mathrm{d}\theta \frac{\alpha_s^{1\ell}(k_t)}{\pi\theta} P_i(z)\Theta(z\theta^a - \kappa) =$$

$$\ln\kappa \left(g_{11}\alpha_s \ln\kappa + \frac{2C_i\beta_0(1+a)}{3a^2\pi}\alpha_s^2 \ln^2\kappa - \frac{2C_i\beta_0^2(1+a+a^2)}{3a^3\pi}\alpha_s^3 \ln^3\kappa\right)$$

$$+ g_{21}\alpha_s \ln(\kappa) + \frac{2B_iC_i\beta_0}{a^2\pi}\alpha_s^2 \ln^2\kappa + \mathcal{O}(\mathrm{N}^3\mathrm{DL})\,. \tag{2.43}$$

Note that in the previous expression we have only kept the relevant terms up to N^{2}DL, as indicated by the $\mathcal{O}$(N^{3}DL) notation. Now, we can identify the terms corresponding to the soft and collinear piece of the splitting function to be

$$g_{12} = \frac{2\beta_0 C_i(1+a)}{3a^2\pi}\,, \tag{2.44}$$

$$g_{13} = -\frac{2\beta_0^2 C_i(1+a+a^2)}{3a^3\pi}, \tag{2.45}$$

while the hard-collinear correction results into

$$g_{22}^1 = \frac{2B_iC_i\beta_0}{a^2\pi}. \tag{2.46}$$

In the last equation, the upper subscript in the coefficient indicates that this is not the only term that contributes to the g_{22} coefficient, i.e. $g_{22} = \sum_i g_{22}^i$.

To achieve N^{2}DL accuracy, we need to go to the next order in the running coupling. We work in the CMW scheme [39] which enables to include also the 2-loop contribution of the splitting functions in the soft limit. Then, the running of the coupling at two loops is

JHEP07(2021)020

given by

$$\alpha_s^{2\ell}(k_t) = \frac{\alpha_s}{1+2\alpha_s\beta_0\ln(\frac{k_t}{Q})} - \frac{\beta_1\alpha_s^2}{\beta_0}\frac{\ln(1+2\alpha_s\beta_0\ln(\frac{k_t}{Q}))}{[1+2\alpha_s\beta_0\ln(\frac{k_t}{Q})]^2} + \frac{K}{2\pi}\frac{\alpha_s^2}{[1+2\alpha_s^2\beta_0\ln(\frac{k_t}{Q})]^2}\,, \quad (2.47)$$

with $\beta_1 = (17C_A^2 - 5C_An_f - 3C_f)/(24\pi^2)$ and $K = (67/18 - \pi^2/6)C_A - 5n_f/9$. Again, we can integrate the branching kernel with the 2-loop running coupling to identify another contribution to the g_{22} coefficient:

$$g_{22}^2 = -\frac{KC_i}{2a\pi^2}\,. \quad (2.48)$$

Soft emissions at large angles. The dynamically groomed $k_{t,g}$ pertains to the category of so-called non-global observables,[7] i.e. it is sensitive to a certain region of the radiation phase space. As such it is affected by a particularly complex class of logarithms known as non-global logs [40–42]. A typical configuration that can give rise to these contributions is a collection of large angle gluons outside the jet which subsequently radiate softer gluons inside. In order to understand how this topology contributes to the $k_{t,g}$ distribution, we first calculate the lowest $\mathcal{O}(\alpha_s^2)$ term coming from such configurations. For illustrative purposes, we start with the calculation of the leading non-global logarithm in e^+e^- annihilation, in which the color structure of the event is simpler. We discuss the straightforward generalization to $p+p$ collisions in the following paragraph. Once again, we rely on the small-R limit and sketch how to lift this approximation in the next section.

The calculation of the non-global contribution at $\mathcal{O}(\alpha_s^2)$ is standard: one calculates the cross-section for two correlated gluon emissions strongly ordered in energy, with the first emission outside the jet and the second inside. For Dynamical Grooming, one can rely on the fact that the gluon inside the jet is necessarily the hardest, since it is the only one at this order. Thus, the double differential distribution for having a dynamically groomed z_g and θ_g value from a non-global configuration initiated by a $q\bar{q}$ dipole is:

$$\begin{aligned}
&\frac{1}{\sigma_0}\frac{\mathrm{d}^2\sigma^{\mathrm{NG}}}{\mathrm{d}z_g\,\mathrm{d}\cos(R_g)} \\
&\quad = 4C_FC_A\left(\frac{\alpha_s}{2\pi}\right)^2\int_0^{p_T}\frac{\mathrm{d}\omega_1}{\omega_1}\int_0^{\omega_1}\frac{\mathrm{d}\omega_2}{\omega_2}\int_{-1}^{1}\mathrm{d}\cos R_1\int_{-1}^{1}\mathrm{d}\cos R_2\,\Omega(\cos R_1,\cos R_2) \\
&\qquad \Theta(\cos(R)-\cos(R_1))\Theta(\cos(R_2)-\cos(R))\delta\left(z_g-\frac{\omega_2}{p_T}\right)\delta(\cos(R_2)-\cos(R_g)) \\
&\quad = 4C_FC_A\left(\frac{\alpha_s}{2\pi}\right)^2\frac{1}{z_g}\ln\left(\frac{1}{z_g}\right)\Theta(\cos(R_g)-\cos(R))\int_{-1}^{\cos(R)}\mathrm{d}\cos R_1\,\Omega(\cos R_1,\cos R_g)\,,
\end{aligned} \quad (2.49)$$

where the first Θ-function in the second line enforces the first gluon to be outside the jet, while the second Θ-function constrains the tagged emission to be inside. Note that this is the real term only.[8] The function Ω is the azimuthal average of the real cross-section for

[7] This also applies to z_g and θ_g.

[8] In principle the lower bound in the integration range of $\cos R_1$ depends on the jet selection. However, in the small R approximation (see the discussion thereafter), those corrections are power of R suppressed.

JHEP07(2021)020

g_{nm}	Physical origin
$g_{11} = -\dfrac{C_i}{a\pi}$	Soft and collinear
$g_{12} = \dfrac{2\beta_0 C_i(1+a)}{3a^2\pi}$	Soft and collinear + $\alpha_s^{1\ell}(k_t)$
$g_{13} = -\dfrac{2\beta_0^2 C_i(1+a+a^2)}{3a^3\pi}$	Soft and collinear + $\alpha_s^{1\ell}(k_t)$
$g_{21} = -\dfrac{2C_i B_i}{a\pi}$	Hard and collinear
$g_{22} = \dfrac{2C_i B_i \beta_0}{a^2\pi} - \dfrac{KC_i}{2a\pi^2} - \dfrac{\pi^2}{3}\dfrac{C_i C_A}{(2\pi)^2}$	Hard and collinear $+\alpha_s^{1\ell}(k_t)$, $\alpha_s^{2\ell}(k_t)$, non-global soft

Table 1. Relevant coefficients for the Sudakov form factor at N^{2}DL accuracy.

correlated double gluon emission from a quark [3]:

$$\Omega(\cos R_1, \cos R_2) = \frac{2}{(\cos(R_2) - \cos(R_1))(1-\cos(R_1))(1+\cos(R_2))}\,. \tag{2.50}$$

The $R_g = \theta_g R$ integral in eq. (2.49) is non-singular in the collinear limit, so that one can perform the two angular integrals exactly to get the leading term in the soft and $R \to 0$ limit:

$$\frac{1}{\sigma_0}\frac{\mathrm{d}\sigma^{\mathrm{NG}}}{\mathrm{d}z_g} = 2C_F C_A \left(\frac{\alpha_s}{2\pi}\right)^2 \frac{1}{z_g} \ln\left(\frac{1}{z_g}\right) \frac{\pi^2}{3}\,. \tag{2.51}$$

In the previous equation, the soft singularity when $z_g \to 0$ induces a single log contribution which has to be taken into account at N^{2}DL as part of the g_2 function in eq. (2.28).

In $p+p$ collisions, the situation is *a priori* more involved. Each Born level partonic configuration needs to be broken into distinct hard dipoles. However, as shown in ref. [17], only the dipoles involving the measured jet matter in the small R limit (i.e. neglecting terms proportional to θ^n), and all such contributions are enhanced by the same $\pi^2/3$ factor as in the e^+e^- result in eq. (2.51). Consequently, the non-global contribution to the resummed distributions factorize according to the flavour of the jet, in the same way as the collinear piece calculated above. In other words, by imposing the small R limit we can use the e^+e^- result from eq. (2.51) for the $p+p$ case. By doing so, we can extract the last piece of the g_{22} coefficient, namely

$$g_{22}^3 = -\frac{\pi^2}{3}\frac{C_i C_A}{(2\pi)^2}\,. \tag{2.52}$$

At this point let us summarize the main ingredients obtained so far in a compact way that shall facilitate the reproducibility of our results. At N^{2}DL accuracy, and in the small-R limit, the Sudakov form factor given by eq. (2.28) involves the coefficients provided in table 1.

JHEP07(2021)020

Equivalently, using eq. (2.35) we arrive to the following, non unique expression for the branching kernel

$$\widetilde{P}(z,\theta) = \frac{2\alpha_s C_i}{\pi z\theta}(1 - 2\alpha_s\beta_0 \ln(\mu_K z\theta) + 4\alpha_s^2\beta_0^2 \ln^2(\mu_K z\theta)) + \frac{2\alpha_s C_i B_i}{\pi\theta}(1 - 2\alpha_s\beta_0 \ln(\mu_K\theta)) + \frac{K}{2\pi}\frac{2C_i\alpha_s^2}{\pi z\theta} - 2C_iC_A\left(\frac{\alpha_s}{2\pi}\right)^2\frac{\pi^2}{3}\frac{\ln(\mu_K z)}{z}\,. \tag{2.53}$$

In order to estimate the uncertainty of the resummation, we have introduced the dimensionless multiplicative factor μ_K that will be varied between 0.5 and 2. A subtle issue[9] concerning this μ_K variation is that the first non-trivial correction that arises after integrating over (z,θ) eq. (2.53) is given by $\propto \alpha_s^2 \ln^2(\kappa)\ln(\mu_K)$. This is of the same order as the corresponding K-term, i.e. $\propto \alpha_s^2 \ln^2(\kappa)K$. To overcome the non-desirable variation of a g_{nm} coefficient, we vary μ_K under the condition that the K term is constant, i.e. K is shift to $K + 4\pi\beta_0 \ln(\mu_K)$ in the calculation (and similarly when considering varition of the renormalization scale Q).

Note that the approximations made to derive the coefficients inside the Sudakov factor $\Delta(\kappa|a)$ and the branching kernel $\widetilde{P}$ lead to a differential cross-section which is not necessarily normalized to the Born jet cross-section. To restore the correct normalization, one can simply divide the cumulative distribution eq. (2.36) by $\Sigma(1)$. This overall normalization factor is a non-logarithmic correction which does not spoil our targeted accuracy.

Lastly, we would like to draw the reader's attention to the fact that multiple gluon emissions were not considered in this calculation. Generically speaking, at leading logarithmic accuracy, a single emission dominates jet substructure observables. This strong ordering might be broken beyond leading-log, like in the jet mass case, such that an arbitrary number of emissions give comparable contributions to the final measured value. The region of phase space for which this happens has to be determined on an observable basis thus increasing the complexity of analytic calculations. In the Dynamical Grooming case, multiple emissions do not have to be considered for the observables computed in this paper. This property is a direct consequence of how the method is built. That is, dynamically groomed observables in tagging mode are not additive but defined on the hardest emission and thus it is the only one that contributes to all orders in the resummation.

N^{2}DL and N^{2}DL'. Insofar, we have provided the minimal set of g_{nm} coefficients that lead us to N^{2}DL accuracy. For that purpose we have neglected all terms that are not logarithmically enhanced. In order to gauge the impact of these sub-leading contributions, we will also provide results with the 'complete' branching kernel, i.e.

$$\widetilde{P}(z,\theta) = \left[\frac{2\alpha_s^{2\ell}(\mu_K z\theta Q)C_i}{\pi z\theta} - 2C_iC_A\frac{\pi^2}{3}\left(\frac{\alpha_s}{2\pi}\right)^2\frac{\ln(\mu_K z)}{z}\right]\Theta\left(e^{-B_i} - z\right)\,, \tag{2.54}$$

and the Sudakov $\Delta(\kappa|a)$ calculated *exactly* from this complete branching kernel whose explicit expression can be found in appendix A. Note that the resulting differential cross-section is then normalized by construction. The running of the coupling is neglected in

[9] We are grateful to Gregory Soyez for pointing this out.

JHEP07(2021)020

the non-global term for simplicity, adding it would enable to account for part of the full resummation of the non-global soft function [40]. We will refer to this resummation as N^{2}DL', where the prime indicates that the resummation actually includes some of the sub-leading logarithmic corrections with $p \geq 3$. That said, we emphasize that in all rigour, both ways of doing the resummation — either 'minimally' using eq. (2.53) and the coefficients in table 1 in the Sudakov or with the complete branching kernel given by eq. (2.54) — reach the same N^{2}DL logarithmic level accuracy, and not more.[10] Therefore, we will use this resummation scheme freedom to leverage our uncertainty.

Beyond N^{2}DL and the small-R limit Before we move on to the fixed-order section, we would like to sketch which steps have to be taken in order to extend the calculation that we have just presented.

In the first place, if the small-R constraint is lifted, one has to account for process dependent terms that enter the calculation as a power series in the jet radius. Physical scenarios that lead to such contributions involve soft and large angle emissions that end up being clustered in the reconstructed jet. For example, a splitting originated from the initial state partons can be tagged by Dynamical Grooming and induce single logarithmic terms suppressed by powers of the jet radius R in the resummation. The difficulty with soft emissions at large angles comes from the fact that such emissions have a complicated color structure which depend on the full Born level event and not only on the Casimir factor of the measured jets. In order to handle such corrections, which are expected to be important for $R \sim 1$, one could decide to rely on the large N_c limit and decompose each Born processes into different colour flows, as done in refs. [8, 40] in the context of the Lund plane density. Then, each color flow corresponds to a superposition of hard dipoles, which can radiate a soft large angle gluon into the measured jet. In practical terms, adding these contributions would promote the jet flavour dependence of $\widetilde{P}(z,\theta)$ and $\Delta_i(\kappa|a)$ to a color flow one. Once these new terms are taken into account N^{2}DL accuracy is reached beyond the small jet radius limit, but in the large N_c approximation. If one does not resort to the large N_c limit, one has to deal with matrix formulae in color space, as in ref. [17]. It is however unclear if the simple structure of eq. (2.36) remains when the exponentiation has a matrix form and it deserves a dedicated study.

From the non-global logarithms side, their full resummation is required if a higher accuracy in the resummation is intended. This is a complicated task for the $\kappa^{(b,c)}$ observable in $p+p$ collisions, even in the large N_c limit. With the latter approximation, one could resort to the same numerical method as in ref. [17] (see also ref. [18] for the θ_g distribution defined with Soft Drop).

2.3.2 Matching to fixed-order

In order to produce reliable predictions when $k_{t,g} = \mathcal{O}(1)$ and to achieve N^{2}DL via the C_1 term, the resummed distribution obtained in the previous section needs to be matched with a fixed-order calculation. Several matching schemes are available in the literature. For

[10] Indeed, a complete N^{3}DL resummation would require at least the first term in the analytic expansion of the g_3 (NNLL) function inside $\Delta(\kappa|a)$.

JHEP07(2021)020

our purposes, it is clearly desirable to have a matching scheme satisfying the two following conditions: (i) the matching scheme should produce 'for free' the C_1 term, (ii) the matching scheme should preserve the fixed order endpoint of the distribution at $k_{tg,\max} = 0.5$. Two possible matching schemes that satisfy these requirements are the multiplicative and the $\log(R)$ matching [37, 43]. In what follows, we shall use multiplicative matching at leading order ($\mathcal{O}(\alpha_s)$) and discuss how to extend it to next-to-leading order ($\mathcal{O}(\alpha_s^2)$).

As the colour structure of the resummation is tremendously simplified within our targeted accuracy, i.e. it only depends on the jet flavor. The matching formula can be decomposed accordingly as follows [43]:

$$\Sigma^{\mathrm{LO+N^2DL}}(k_{t,g}) = \frac{1}{\sigma_0+\sigma_1}\left\{\sum_{i=q,g}\tilde{\Sigma}_i^{\mathrm{N^2DL}}(k_{t,g})\left(1+\frac{\Sigma_i^{\mathrm{LO}}(k_{t,g})-\tilde{\Sigma}_{i,1}^{\mathrm{N^2DL}}(k_{t,g})}{\sigma_{0,i}}\right)+\Sigma_{\mathrm{else}}^{\mathrm{LO}}(k_{t,g})\right\}. \tag{2.55}$$

We proceed to describe the ingredients entering the previous equation, except the meaning of the last term that will become clear later on. First, σ_0 and σ_1 are the inclusive dijet cross-section at leading order and next-to-leading order respectively. The $\tilde{\Sigma}_i$ is the resummed cumulative $k_{t,g}$ cross-section for i-jets, that shares the same endpoint, $k_{tg,\max}$, as the fixed order distribution. At a given accuracy, this is achieved through the following transformation:[11]

$$\tilde{\Sigma}_i(k_{t,g}) = \sigma_{0,i}\Sigma_i\left[\exp\left(-\log\left(\frac{1}{k_{t,g}}-\frac{1}{k_{tg,\max}}+1\right)\right)\right]. \tag{2.56}$$

Notice that, besides the shift in the endpoint, we have multiplied the resummed cumulative distribution by $\sigma_{0,i}$ in order to ensure that $\tilde{\Sigma}^{\mathrm{N^2DL}}$ and Σ^{LO} have the same units. Following up with the pieces entering eq. (2.55), $\tilde{\Sigma}_{i,1}^{\mathrm{N^2DL}}$ is the $\mathcal{O}(\alpha_s)$ term in the expansion of $\tilde{\Sigma}_i^{\mathrm{N^2DL}}$, while Σ_i^{LO} is the leading order distribution defined as

$$\Sigma_i^{\mathrm{LO}}(k_{t,g}) = \sigma_{1,i} - \int_{k_{t,g}}^1 \mathrm{d}k'_{t,g}\frac{\mathrm{d}\sigma_i^{\mathrm{LO}}}{\mathrm{d}k'_{t,g}}. \tag{2.57}$$

Regarding the normalization, we find that, by construction, $\Sigma^{\mathrm{LO+N^2DL}}(k_{tg,\max}) = 1$.

One can check that the limiting behavior of the matched distribution is correct. Indeed, eq. (2.55) gives back the LO distribution for $k_{t,g} \sim k_{tg,\max}$. In turn, when $k_{t,g} \ll 1$ the distribution behaves like

$$\Sigma^{\mathrm{LO+N^2DL}}(k_{t,g}) \simeq \frac{1}{\sigma_0+\sigma_1}\sum_{i=q,g}\tilde{\Sigma}_i^{\mathrm{N^2DL}}(k_{t,g})(1+\alpha_s C_{1,i}), \tag{2.58}$$

where the C_1 term is given by its standard definition:

$$\alpha_s C_{1,i} = \lim_{k_{t,g}\to 0}\frac{\Sigma_i^{\mathrm{LO}}(k_{t,g})-\tilde{\Sigma}_{i,1}^{\mathrm{N^2DL}}(k_{t,g})}{\sigma_{0,i}}. \tag{2.59}$$

[11] When the resummed distribution has an endpoint different from 1, eq. (2.56) needs to be modified accordingly. In particular, when using the calculation of Σ at $\mathrm{N^2DL'}$, the $+1$ inside the logarithm is replaced by $\exp(B_q)$.

JHEP07(2021)020

Thus, eq. (2.58) shows that the matched distribution reproduces the resummed result in its regime of validity. In our calculation, C_1 is a constant up to R^2-suppressed single logarithmic contributions.

In practice, the LO $k_{t,g}$ differential cross-section and the LO and NLO jet cross-sections σ_0 and σ_1 are obtained using MadGraph5 [44] (in fixed order mode) with CT10nlo PDF set [45]. The factorisation scale for the PDF convolution is set to $\mu_F Q$, with μ_F a dimensionless factor introduced to estimate the uncertainty relative to this prescription. For a given jet selection $[p_{t,\min}, p_{t,\max}]$, a unique generation cut is imposed in the fixed order calculation. Namely, the sum of the transverse momenta of the partons is required to be larger than $p_{t,\min}$ and one asks for at least one jet with $p_t > p_{t,\min}/4$. We have checked that the resulting cross-sections are insensitive to the precise value of these cuts. The reference value of the strong coupling at the jet scale, $\alpha_s(\mu_R Q)$, is evaluated in the $\overline{\text{MS}}$ scheme. The μ_R is a dimensionless factor used to gauge the uncertainty with respect to the renormalization scale. When the p_t selection is broad, such as in the ATLAS set-up detailed in the following section, the p_t range is divided into smaller bins in which the inclusive jet and k_{tg} cross-sections are calculated. The extension of eq. (2.55) in this case is straightforward.

The last ingredient in eq. (2.55) involves the decomposition according to the flavour of the jet. This is done in an IRC safe way for both the $k_{t,g}$ differential and inclusive jet cross-section. More concretely, at LO the jets have at most two constituents. Then, when the jet has zero or one net flavour, the jet is tagged as a gluon or quark jet, respectively. Otherwise, whenever the jet is multi-flavored, i.e. it contains two (anti)-quarks of different flavor, it pertains to what we call the 'else' category. The LO $k_{t,g}$ differential cross-section for these multi-flavored jets goes to zero at small $k_{t,g}$ and contributes to the full match result via the $\Sigma^{\text{LO}}_{\text{else}}(k_{t,g})$ term in eq. (2.55).

Finally, as for the resummation part, we would like to comment on how to further extend the matching procedure to higher accuracy. In this case, the equivalent of the multiplicative matching formula eq. (2.55) at NLO can be found in ref. [43] and it involves the NLO $k_{t,g}$ differential cross-section. The latter can be obtained by generating 3-jet events at NLO with MadGraph, or any other code dedicated to matrix element calculations. Even if there is no conceptual difficulty in promoting our matching to NLO, we postpone it for further studies given that its quantitative impact on the resulting distributions could be as sizeable as the missing power of R suppressed terms on the resummation. Therefore, we believe that these two endeavors should be pursued simultaneously and must be included for refining our phenomenological studies presented in section 3.

2.3.3 Results

Once the analytic framework has been presented, we proceed to show some numeric results for high-p_t jets ($800 < p_t < 1000\,\text{GeV}$) at top LHC energy $\sqrt{s} = 13\,\text{TeV}$ with cone size $R = 0.4$. The central value of the following curves is obtained with $\mu_F = \mu_R = 1$. Further, the error bars are obtained by varying a factor of two the following parameters in the calculation: factorization and renormalization scales through the 7-point rule [46], the parameter μ_K that controls the scale at which the strong coupling runs (see eq. (2.53)) and, in the case of

JHEP07(2021)020

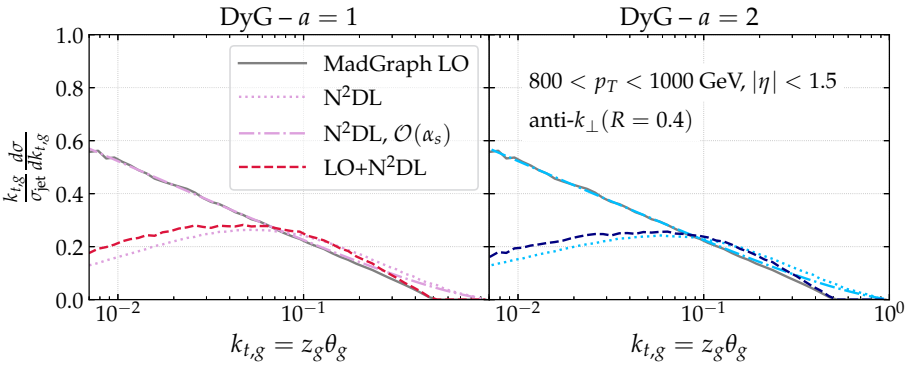


Figure 1. The $k_{t,g}$-distribution computed in different ways: at leading-order with MadGraph, resumed at N^{2}DL as given by eq. (2.35) and table 1, first order expansion of the resumed result, and the matched distribution (see eq. (2.55)) for $a = 1$ (left) and $a = 2$ (right). The normalization factor $\sigma_{\rm jet}$ reduces to $\sigma_0 + \sigma_1$ for the resummed and matched distributions and to σ_0 in the other two cases.

N^{2}DL', the freezing scale $\mu_{\rm fr}$ used to avoid the Landau pole in eq. (A.1). Then, we combine the various uncertainties by taking the envelope of all distributions.

In figure 1 we present four distributions of $k_{t,g}$: (i) the fixed-order result, (ii) the resummed result at N^{2}DL, (iii) the $\mathcal{O}(\alpha_s)$ expansion of the latter and (iv) the matched distribution at LO+N^{2}DL. It is clear from this figure that the matching procedure works as expected, i.e. the LO+N^{2}DL recovers the N^{2}DL result at small $k_{t,g}$, while it tends towards the leading-order curve in the opposite regime. In addition, the endpoint of the resummation is shifted by the matching procedure to the fixed-order one at $k_{t,g} = 0.5$. By comparing the fixed-order result and the first term in the α_s-expansion of the resummed result, we can get a hint on the size of the $\mathcal{O}(R^n)$ logarithmically enhanced terms that we have so far neglected. In fact, the difference between the $\mathcal{O}(\alpha_s)$ term of the N^{2}DL curve and the exact leading order result converges towards a constant at small $k_{t,g}$. This indicates that these power suppressed terms enter with a small coefficient in the cumulative distribution and can be safely neglected for the setup studied in this work. All the previous statements hold for both values of a. In particular, the fixed order result is independent of a because there is only one splitting tagged.

Next, we compare in figure 2 the two prescriptions to perform the resummation that we have discussed above, i.e. keeping uniquely the logarithmically enhanced terms at N^{2}DL or including sub-leading corrections (N^{2}DL'). In the large $k_{t,g}$ regime, we observe no difference between the LO+N^{2}DL and the LO+N^{2}DL' as it is expected since in this limit the fixed-order contribution dominates the matched result. This is no longer the case for $k_{t,g} \ll 1$, where details of the resummation structure do matter. An important remark is that the discrepancy between the two curves diminishes when increasing the parameter a that determines the hardness condition in the grooming algorithm. We attribute this to the a-

JHEP07(2021)020

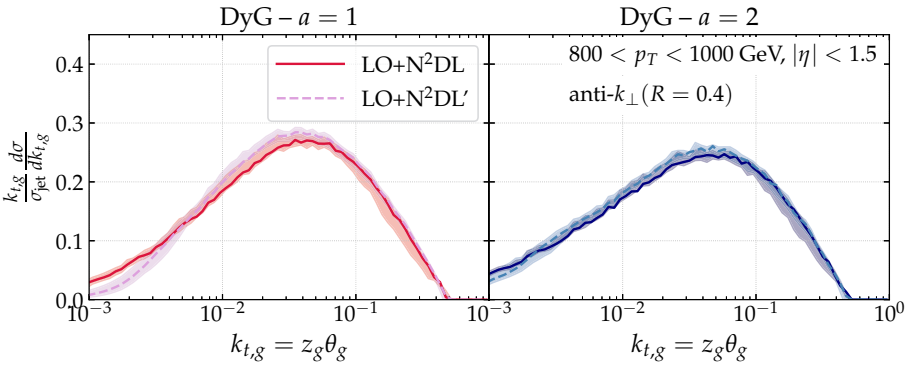


Figure 2. The $k_{t,g}$-distribution with a minimal N^{2}DL resummation (see eq. (2.53) and table 1), and including sub-leading contributions N^{2}DL' (see eq. (2.54) and appendix A). Both curves are normalized to $\sigma_0+\sigma_1$.

scaling of the g_{nm} parameters that, as one can see in table 1, satisfies $g_{nm} \sim 1/a$. Hence, the larger the value of a is, the smaller the coefficients in front of the higher order terms are and the narrower the difference between N^{2}DL and N^{2}DL' becomes. In the phenomenological section, we will include the differences between N^{2}DL and N^{2}DL' as part of our uncertainty band given that, from a logarithmic counting point of view, there is no preferred option.

2.4 z_g, θ_g at LO+N^{2}DL accuracy

As discussed at length in section 2.1, the momentum sharing fraction, z_g, and opening angle, θ_g, of the splitting tagged by Dynamical Grooming are Sudakov safe observables. In eqs. (2.24)–(2.25), we defined their distribution as the limit of the IRC safe $k_{t,g}$ distribution. This allows us to follow the same logic as in the previous section to obtain the resummation part of their distribution. In turn, the fixed-order result is not even well defined, as shown in eqs. (2.22) and (2.20), and thus the matching strategy differs to that presented in section 2.3.2. In what follows, we provide the necessary ingredients to reach LO+N^{2}DL accuracy in the small-R limit.

2.4.1 Boundary logarithms for the θ_g distribution

In the case of z_g, the resummation proceeds in exactly the same fashion as for $k_{t,g}$. In turn, for θ_g, another source of logarithmic enhancement appears, caused by the interplay between the anti-$k_\perp$ algorithm [47] used to cluster the jet, and the C/A algorithm to decluster it in the Dynamical Grooming procedure. These so-called clustering boundary logarithms [8] are of the form:

$$\alpha_s^2 \frac{1}{z} \ln\left(\frac{1}{z}\right) \ln\left(\frac{R}{R-R_g}\right) , \tag{2.60}$$

where the double logarithmic enhancement becomes important when $R_g \to R$ ($\theta_g \equiv R_g/R$ close to 1).

JHEP07(2021)020

Boundary logarithms arise from a non-global configuration where the first emission is outside the anti-$k_\perp$ jet, while the second is inside and almost collinear to the first one. In this situation, C/A algorithm would cluster these two emissions together, as part of the jet. As such, their leading contribution to the branching kernel can be obtained from the same calculation as the one done in eq. (2.49), but focusing in the regime where $\theta_1 \simeq \theta_2 \simeq R \ll 1$ [8, 48], such that:

$$\Omega(\theta_1, \theta_2) \simeq \frac{4}{\theta_1^2(\theta_1^2 - \theta_2^2)}\,, \tag{2.61}$$

and

$$\begin{aligned}
\frac{1}{\sigma_0}\frac{\mathrm{d}^2\sigma^{\mathrm{NG,cl}}}{\mathrm{d}z_g \mathrm{d}\theta_g} &\simeq 4C_F C_A \left(\frac{\alpha_s}{2\pi}\right)^2 \frac{1}{z_g} \ln\left(\frac{1}{z_g}\right) R^2 \int_R^\infty \mathrm{d}\theta_1 \theta_1 \theta_g\, \Omega(\theta_1, \theta_g R) \\
&= 8C_R C_A \left(\frac{\alpha_s}{2\pi}\right)^2 \frac{1}{z_g} \ln\left(\frac{1}{z_g}\right) \frac{1}{\theta_g} \ln\left(\frac{1}{1-\theta_g^2}\right) \qquad (2.62) \\
&\simeq 8C_R C_A \left(\frac{\alpha_s}{2\pi}\right)^2 \frac{1}{z_g} \ln\left(\frac{1}{z_g}\right) \frac{1}{\theta_g} \ln\left(\frac{1}{1-\theta_g}\right). \qquad (2.63)
\end{aligned}$$

The upper boundary in the θ_1 integral can be safely sent to ∞ since the integral is dominated by the region $\theta_1 \simeq \theta_2 \simeq R$. To get the last line, we have kept the dominant contribution at $\theta_g \sim 1$.

As stated above, this $\mathcal{O}(\alpha_s^2)$-contribution is enhanced by two soft logarithms of the type $\ln(z_g)/z_g$ and one boundary logarithm $\ln(1-\theta_g)$. However, the logarithmic divergence associated with $\theta_g \sim 1$ is integrable in a neighbourhood of 1. Consequently, as long as one deals with an observable in which the angle $R_g \sim R$ is integrated out between some lower bound and 1 (such as z_g or $k_{t,g}$ distributions), the boundary divergence is harmless. More precisely, its integral over θ_g should give back the soft single logarithmic divergence that was part of our treatment of non-global configurations. This argument also applies to the Sudakov factor, since vetoing all emissions with hardness larger than κ translates into the following integral

$$\int_0^1 \mathrm{d}z' \int_0^1 \mathrm{d}\theta' \, \frac{1}{\sigma_0}\frac{\mathrm{d}^2\sigma^{\mathrm{NG,cl}}}{\mathrm{d}z'\mathrm{d}\theta'} \Theta(z'\theta'^a - \kappa)\,, \tag{2.64}$$

where it is clear that θ' is always marginalized in the neighbourhood of 1. In other words, the Dynamical Grooming procedure lowers the singularity associated with boundary logarithms from double- to single-log, and this single-log term is already taken into account by the coefficient g_{22}^3 in the Sudakov.

From that perspective, the θ_g distribution is peculiar since the $R_g \to R$ logarithmic divergence from eq. (2.63) is not integrated out. To effectively include boundary logarithms for the θ_g distribution, we replace the last term in the branching kernel $\widetilde{P}(z, \theta)$ given in eq. (2.53) by

$$-2C_i C_A \left(\frac{\alpha_s}{2\pi}\right)^2 \frac{\pi^2}{3}\frac{\ln z}{z} \to -8C_i C_A \left(\frac{\alpha_s}{2\pi}\right)^2 \frac{\ln z}{z}\frac{1}{\theta_g} \ln\left(\frac{1}{1-\theta_g}\right) \Theta(\theta_g - \bar{\theta})\,, \tag{2.65}$$

JHEP07(2021)020

with $\bar{\theta}$ defined such that

$$\int_{\bar{\theta}}^{1} \frac{\mathrm{d}\theta_g}{\theta_g} \ln\left(\frac{1}{1-\theta_g}\right) = \frac{\pi^2}{12}\,. \tag{2.66}$$

Numerically, one finds $\bar{\theta} \simeq 0.66$. The step function guarantees that the single logarithmic term from soft non-global configurations is correctly accounted for within our targeted accuracy and without double counting.[12] Such a constraint is also physically expected since boundary logarithms come from the region where $\theta_g \sim 1$ by definition.

Finally, we would like to discuss how this new logarithmic divergence affects the logarithmic counting provided in section 2.2. For the θ_g distribution, we have found two sources of logarithmic enhancement that are either of the form $\ln(\theta_g)$ or $\ln(1-\theta_g)$. Since the veto factor in $\Delta(\kappa|a)$ suppresses boundary logarithms, there is only one power of $\ln(1-\theta_g)$ that appears in the α_s expansion of the θ_g distribution and it comes from the α_s^2 result given by eq. (2.65). In order to have the correct logarithms at $\mathrm{N^2DL}$ in front of this single power of $\ln(1-\theta_g)$, it is enough to solely include the first one-loop correction in the running coupling $\alpha_s^2 \to \alpha_s^2(1-4\alpha_s\beta_0\log(z))$ and the hard-collinear correction at fixed coupling $1/z \to \Theta(e^{-B_i}-z)/z$ in eq. (2.65).

2.4.2 Comparison between the resummation structure of Soft Drop and Dynamical Grooming

The idea of studying the momentum sharing fraction and opening angle of a given splitting in the shower was originally proposed in ref. [4]. In this work, the splitting at issue was selected through the Soft Drop procedure, that is, the first branching in the de-clustering sequence that satisfies $z > z_{\mathrm{cut}}\theta^{\beta}$. These observables, (z_g, θ_g), have been measured experimentally [31, 49] and resummed to modified-leading log [50] and next-to-leading log accuracy [18], respectively. An important comment at this point is that Soft Drop observables do exponentiate and, therefore, a $\mathrm{N^pLL}$ counting applies. Hence, strictly speaking, an apples-to-apples comparison on the resummation structure for Soft Drop and Dynamical Grooming does not exist.

We have identified one major simplification in the resummation structure of θ_g when it is defined through the Dynamical Grooming procedure instead of with Soft Drop: Dynamical Grooming is free of clustering logarithms. Let us briefly recap how these contributions arise for two correlated emissions [42]. Consider the emission of a gluon, p_1, off a hard quark p_0 together with a secondary emission, p_2, off p_1. These two emissions have commensurate angles $\theta_{01} \sim \theta_{02}$, while their energies (and thus transverse momentum) are strongly ordered $z_2 \ll z_1$ ($k_{t,2} \ll k_{t,1}$). The C/A algorithm will miss-cluster the secondary gluon as a primary if $\theta_{02} < \theta_{12}$, with θ_{12} being the relative distance between the two emissions. Then, if p_2 is a real emission it will trigger the Soft Drop condition, even though $z_1 \gg z_2$, and consequently $\theta_{02} = \theta_g$. In turn, if p_2 is virtual, the tagged splitting would be p_1 and $\theta_1 = \theta_g$. This mismatch between the real and virtual contributions lead to a tower of logarithms at NLL

[12] The spirit of the step function is essentially the same as in our treatment of hard collinear emissions via the effective splitting function given by eq. (2.38) Another way of including boundary logarithms without double counting is to use directly eq. (2.62) (without step function) since $-\int_0^1 \mathrm{d}\theta \ln(1-\theta^2)/\theta = \pi^2/12$.

JHEP07(2021)020

that were numerically computed, in the large N_c-limit, in ref. [18] for θ_g, as defined by Soft Drop, and also discussed in the context of the Lund plane in ref. [8]. In the Dynamical Grooming case, even if some secondary emissions can be 'wrongly' pushed by the C/A algorithm into the primary Lund plane, it will never be the hardest given that $z_1 \gg z_2$ and, therefore, $\kappa_1 = z_1\theta_1^a$ will be larger than $\kappa_2 = z_2\theta_2^a$ even if the angles are commensurate. The only effect of these emissions on DyG, beyond $\mathrm{N^2DL}$, would be a small contribution to the p_t degradation of the primary branch.

From the non-global logarithms side, that affect not only z_g but also $k_{t,g}$, we have shown in section 2.3, that they are proportional to $\ln(z_g)$ for Dynamical Grooming. In the Soft Drop case, the soft singularity is cured by the definition of the grooming condition. That is, non-global logs enter in the Soft Drop calculation as $\propto \ln(z_{\mathrm{cut}}\theta^\beta)$ and thus have a smaller impact that in the DyG option.

Therefore, the cleanest grooming procedure from a theoretical point of view in order to avoid the resummation of both non-global and clustering logarithms at NLL would be to combine the two methods. Then, the grooming procedure would be a two-step process: first, one removes all emissions with $z < z_{\mathrm{cut}}$ and then one looks for the hardest one in the Dynamical Grooming sense. This possibility will be further studied in an upcoming publication [51].

2.4.3 Matching to fixed-order

The first leading order matching scheme for Sudakov safe observables was proposed in ref. [30]. It is based on constructing a n-dimensional IRC safe distribution, that we dub 'IRC safe companion', and re-defining the Sudakov safe observable by an appropriate marginalization. In our case, the 2-dimensional IRC safe distribution would be $\mathrm{d}^2\sigma/\mathrm{d}z_g\mathrm{d}\theta_g$, that can be interpreted as the *joint* probability distribution for having a tagged branching with momentum sharing fraction z_g and (normalised) opening angle θ_g. Although z_g and θ_g are Sudakov safe observables by themselves, measuring them simultaneously, i.e. z_g in a given bin of θ_g or vice versa, restores IRC safety.

To define a matching scheme for a Sudakov safe observable, one then rely on the matching of the IRC safe companion. Such matching can be done for instance in a multiplicative way,

$$\mathrm{d}^2\sigma_i^{\mathrm{LO+N^2DL}} = \frac{\mathrm{d}^2\sigma_i^{\mathrm{LO}} \times \mathrm{d}^2\sigma_i^{\mathrm{N^2DL}}}{\mathrm{d}^2\sigma_{i,1}^{\mathrm{N^2DL}}}\,. \tag{2.67}$$

This formula guarantees that $\mathrm{d}^2\sigma_i^{\mathrm{LO+N^2DL}}$ has exactly the same $\mathcal{O}(\alpha_s)$ coefficient as the LO result and reproduces the resumed calculation in the kinematic region enhanced by large logarithms. Notice also that at LO, $\mathrm{d}^2\sigma_i^{\mathrm{LO}}$ coincides with the primary Lund plane density. We would like to point out that this matching scheme applies to jets with a given flavour i. As we shall see, this simplifies the resulting formula as our resummed result depends on the jet's flavour via the Casimir factor of the jet. At LO, such a decomposition is trivial, as explained in section 2.3.2. However, beyond LO, this requires to determine in an IRC safe manner the jet's flavour. This can be done using, for instance, the flavour-k_t clustering algorithm [52].

JHEP07(2021)020

Once $\mathrm{d}^2\sigma_i^{\mathrm{LO+N^2DL}}$ is known, the z_g or θ_g distributions are computed by marginalization. In the case of z_g, for example, this amounts to

$$\frac{1}{\sigma}\frac{\mathrm{d}\sigma^{\mathrm{LO+N^2DL}}}{\mathrm{d}z_g} = \frac{1}{\sigma}\sum_{i=q,g}\int_0^1 \mathrm{d}z \int_0^1 \mathrm{d}\theta \, \frac{\mathrm{d}^2\sigma_i^{\mathrm{LO+N^2DL}}}{\mathrm{d}z\mathrm{d}\theta}\delta(z-z_g)\,, \tag{2.68}$$

where σ is the inclusive jet cross-section. An important feature of Sudakov safe observables is hidden behind the apparent simplicity of eq. (2.68). In fact, not all matching schemes for the IRC safe companion lead to a well-defined integral after marginalization. For instance, choosing an additive matching,

$$\mathrm{d}^2\sigma_i^{\mathrm{LO+N^2DL}} = \left[\mathrm{d}^2\sigma_i^{\mathrm{LO}} - \mathrm{d}^2\sigma_i^{\mathrm{N^2DL},(1)}\right] + \mathrm{d}^2\sigma_i^{\mathrm{N^2DL}}\,, \tag{2.69}$$

induces a collinear divergent term (the one inside the bracket) which is not cured by the Sudakov. On the contrary, the integral in eq. (2.68) is well defined because the Sudakov factor in $\mathrm{d}^2\sigma_i^{\mathrm{N^2DL}}$ shields the $\theta=0$ logarithmic divergence.

We now turn the concrete implementation of eq. (2.68) used in this paper. In principle we could compute $\mathrm{d}^2\sigma_i^{\mathrm{LO}}$ using MadGraph as in our matched calculation of $k_{t,g}$. However, we decide here to take another path that we find more enlightening from the physics point of view and, at the same time, easier to implement numerically. Namely, in the small jet radius limit that we consider throughout this paper, it is possible to provide an explicit analytic expression for $\mathrm{d}^2\sigma_i^{\mathrm{LO}}$. Up to powers of θ_g corrections, it reads

$$\frac{\mathrm{d}^2\sigma_i^{\mathrm{LO}}}{\mathrm{d}z_g\mathrm{d}\theta_g} \simeq \sigma_{0,i}\frac{2\alpha_s C_i}{\pi}\frac{1}{\theta_g}\overline{P}_i(z_g) + \mathcal{O}(\theta_g^n)\,, \tag{2.70}$$

where $\overline{P}_i$ is the symmetrized splitting function of a parton i: $\overline{P}_i(z) = P_i(z) + P_i(1-z)$, summed over all decay channels. Matching our resummed distribution to this form of the LO result is then straightforward as it amounts to replace $2C_i\Theta(e^{-B_i}-z)/z$ in $\tilde{P}(z,\theta)$ (eq. (2.54)) by $2C_i\overline{P}_i(z)$.[13]

In figure 3, we show a comparison between the exact result of $\mathrm{d}^2\sigma_i^{\mathrm{LO}}/\mathrm{d}z_g\mathrm{d}\theta_g$ obtained through MadGraph and the approximation given by eq. (2.70). In addition, we compare these two options with the one that we get after replacing the full symmetrized splitting function by its soft limit in eq. (2.70). More concretely, we have computed $\mathrm{d}^2\sigma_i^{\mathrm{LO}}$ for the gluon channel in the three ways that we have just mentioned and show the z_g and θ_g-projections for two bins of θ_g and z_g, respectively. We see that eq. (2.70) matches the exact leading order result while the soft limit of the splitting function is not enough to accurately reproduce the MadGraph output throughout the whole range of z_g. Similar conclusions can be drawn by analysing the θ_g-projection of $\mathrm{d}^2\sigma_i^{\mathrm{LO}}$. Both for z_g and θ_g, the deviation of eq. (2.70) to the exact result remains below 5%. We have deliberately chosen a low p_t bin to ensure that using $\overline{P}_i$ as a proxy for $\mathrm{d}^2\sigma_i^{\mathrm{LO}}$ is valid in the regime in which the ALICE measurement has been recorded.

We decide to normalize the z_g and θ_g distributions to the Born level cross-section, σ_0, in contrast to the $k_{t,g}$ case where the NLO correction to the inclusive jet cross section, σ_1,

[13] See also refs. [5, 50] for a similar trick in the calculation of the Soft Drop z_g distribution.

JHEP07(2021)020

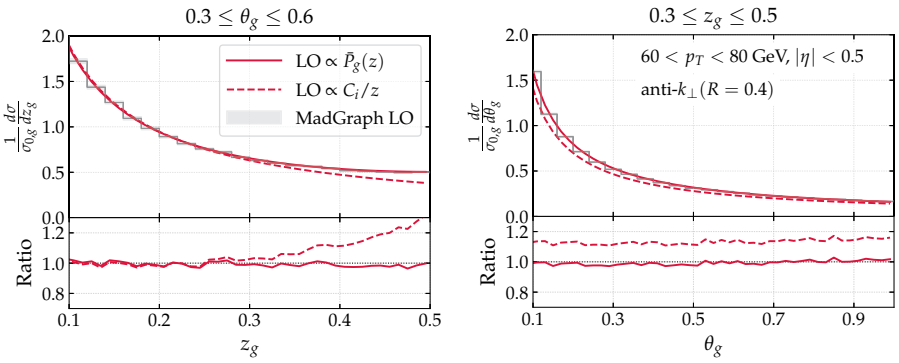


Figure 3. Left: z_g differential distribution at leading order in α_s for the gluon channel in a bin of θ_g computed in three different ways: with MadGraph (gray), using the splitting function as a proxy for the leading order result, see eq. (2.70), either with the leading order expression of $P_i(z)$ (solid, red) or taking the soft limit (dashed, red). Right: same as left panel but for θ_g.

was taken into account. For k_{tg}, the reason we included this correction is to account for the C_1 term, but in principle, a LO matched cross-section can safely be normalized by the Born cross-section without spoiling the targeted accuracy. For Sudakov safe cross-sections such as z_g and θ_g, the question of finding a matching scheme which makes possible the inclusion of the NLO correction to σ in a consistent way with respect to the resummation counterpart is in fact closely related to the C_1 problem that we proceed to tackle.

The C_1 term in Sudakov safe observables. The impossibility to perturbatively expand Sudakov safe observables in powers of α_s invalidates the definition of the C_1 term given by eq. (2.59). Up to now, the question on whether a 'C_1-like' contribution to the resummation exists for this type of observables has not even been addressed in the literature. In this paper, we would like to outline a new, dedicated matching scheme for Sudakov safe observables. The main difference with respect to the original proposal by the authors in ref. [30] is to rely on an IRC safe cross-section which is one-dimensional. This IRC safe companion is built from the Sudakov safe distribution with an additional cut on the kinematic variable that is integrated out. More explicitly, for the dynamically groomed z_g distribution, one defines the IRC safe cumulative distribution $\Sigma(z_g|\theta_{\rm cut})$ using the same grooming procedure, but with an additional cut-off on the angle of the splitting, θ_g, that is denoted $\theta_{\rm cut}$. Then, our matching formula is:

$$\Sigma^{\rm LO+N^2DL}(z_g) = \Sigma^{\rm LO+N^2DL}(z_g|\theta_{\rm cut}) + \Sigma^{\rm N^2DL}(z_g) - \Sigma^{\rm N^2DL}(z_g|\theta_{\rm cut})\,, \tag{2.71}$$

with $\Sigma^{\rm LO+N^2DL}(z_g|\theta_{\rm cut})$ defined using multiplicative matching as in eq. (2.55).[14] Notice that the normalization of $\Sigma^{\rm LO+N^2DL}(z_g)$ to $\sigma_0+\sigma_1$ is ensured. Clearly, this formula has an explicit dependance on $\theta_{\rm cut}$, but it is the price to pay in order to define a C_1-like

[14]The generalization of eq. (2.71) to θ_g and beyond leading order is straightforward.

JHEP07(2021)020

contribution to the logarithmic resummation of Sudakov safe cross-sections. That said, if $\theta_{\rm cut}$ is low enough the distribution is not sensitive to its value over a large z_g domain. To understand this, recall that the Sudakov form factor in $\Sigma^{\rm N^2DL}(z_g)$ provides a natural cut-off, θ_c, on the angular integration of the branching kernel that scales at DLA with $\theta_c \simeq \exp(-1/\sqrt{\bar{\alpha} a})$ see eq. (2.13) (also in ref. [25]). Thus, if $\theta_{\rm cut}$ is chosen smaller than θ_c, we expect that

$$\Sigma^{\rm N^2DL}(z_g) = \Sigma^{\rm N^2DL}(z_g|\theta_{\rm cut}) \tag{2.72}$$

for $z_g \gtrsim \theta_{\rm cut}^a$. Consequently, far from the resummation region, the dominant term in eq. (2.71) is $\Sigma^{\rm LO+N^2DL}(z_g|\theta_{\rm cut})$ which correctly captures the large $z_g \sim 0.5$ domain. On the contrary, in the small z_g limit (but not smaller than $\theta_{\rm cut}^a$), we obtain:

$$\Sigma^{\rm LO+N^2DL}(z_g) \simeq \Sigma^{\rm N^2DL}(z_g) + \alpha_s C_1(\theta_{\rm cut}) \Sigma^{\rm N^2DL}(z_g)\,. \tag{2.73}$$

The second term in the previous equation is a correction to our resummed formula, which looks like a C_1 term. It depends on the value of $\theta_{\rm cut}$ as a reminiscence of the non IRC safety of the z_g distribution. To see that, one notices that up to a constant factor, $C_1(\theta_{\rm cut}) \propto \ln(\theta_{\rm cut})$. Since $\theta_{\rm cut}$ cannot be larger than θ_c, the C_1 correction is actually of order $\mathcal{O}(\alpha_s/\sqrt{\alpha_s a}) = \mathcal{O}(\sqrt{\alpha_s/a})$, at least. There are two interesting features in this scaling behavior. First, the appearance of the square root of α_s is characteristic of Sudakov safe quantities. Second, we observe how C_1 can become sizeable for $a \ll 1$. The latter point reminds us that introducing and ad-hoc parameter, $\theta_{\rm cut}$, in the matching scheme comes with some associated difficulties. In short, from the resummation point of view, one would like $\theta_{\rm cut}$ to be as small as possible such that eq. (2.72) holds. However, the smallness of $\theta_{\rm cut}$ can lead to a sizeable C_1 correction in eq. (2.73), thus spoiling the correct asymptotic limit. A clear trade-off exists and the concrete value of $\theta_{\rm cut}$ in the proposed matching scheme and its applicability to phenomenological applications deserve further investigation. We emphasize that the proposed scheme has not been used in the results of this paper.

2.4.4 Results

Following the reasoning of the $k_{t,g}$ section, we would like to highlight some features of the z_g and θ_g analytic distributions before moving on to the comparison against Monte-Carlo simulations and ALICE preliminary data. In the left panel of figure 4, we quantify the difference between the double-log calculation of the z_g distribution and the LO+N^{2}DL'. The purpose of this figure is to highlight the deviation of the z_g-distribution from the $1/z$ behavior when higher orders in the resummation are included. Indeed, we have shown in eq. (2.14), the z_g-distribution has a dynamically generated cut-off at $z_{\rm cut} \sim e^{-a/\bar{\alpha}}$. For $z > z_{\rm cut}$, it was shown in ref. [25] that the distribution falls off like the soft limit of the Altarelli-Parisi splitting function. We observe that NDL and N^{2}DL contributions such as the running of the strong coupling or the presence of non-global logarithms induce an almost 50% difference with respect to the DLA result. This should be taken into account when interpreting the experimental data specially when searching for modifications in heavy-ion measurements [53, 54].

In the right panel of figure 4, we asses the impact of the boundary logarithms in the θ_g distribution that were discussed in section 2.4.1. As expected, they only matter at large

JHEP07(2021)020

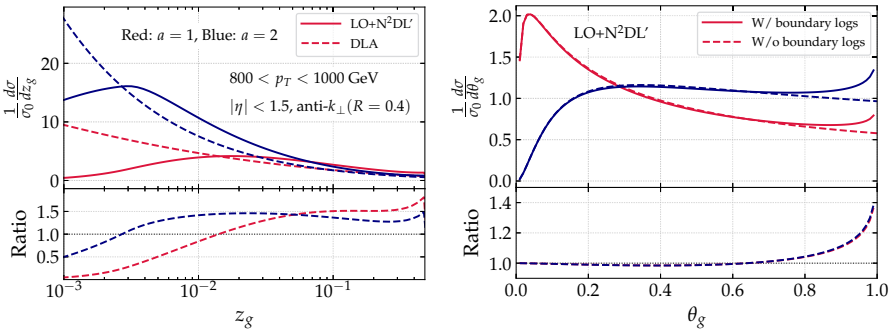


Figure 4. Left: z_g-distribution for $a = 1$ (red) and $a = 2$ (blue) at two different accuracies, DLA and LO+N^{2}DL' and their ratio. Right: θ_g-distribution for $a = 1$ (red) and $a = 2$ (blue) with and without boundary logarithms in the LO+N^{2}DL' result and their ratio.

angles and diverge when $\theta_g \to 1$. Their contribution amounts to a 10–20% and is therefore mandatory to include them if a theory-to-data comparison is aimed.

3 Phenomenology at LHC energies

The analytic calculations that we have presented above rely mainly on two approximations: the narrow jet limit and the use of the Altarelli-Parisi splitting function in the matching as a proxy for the leading order result in the case of z_g and θ_g. In order to evaluate the goodness of such approximations, we test our analytic results against parton level simulations in realistic experimental conditions together with the available experimental data. The results are presented for two values of a in the Dynamical Grooming condition: $a = 1$ and $a = 2$. The reason why we do not consider smaller values of a and, in particular, $a = 0.1$ as done in the ALICE measurement, is because non-perturbative phenomena, beyond the reach of our analytic pQCD calculation, notably affect dynamically groomed observables when $a < 1$. In addition, we utilise the N^{2}DL' prescription on the resummation side.

3.1 Analytics vs. Monte-Carlo parton level

In this section, we compare our analytic calculation for $(k_{t,g}, z_g, \theta_g)$ to parton level simulations of dijet events with Pythia8.235 [34] and Herwig7.1.2 [55]. For the latter we use both the default angular-ordered shower that we denote 'Herwig7-AO' [56] and the dipole-type shower, 'Herwig7-Dip', based on ref. [57]. Given that non-perturbative effects are reduced when going to larger p_t, we study an experimental setup, that lies within the ATLAS capabilities [49], where the comparison to pQCD calculations are deemed to be cleaner. The centre of mass energy is set to $\sqrt{s} = 13$ TeV, jets are clustered with the anti-k$_\perp$ algorithm with $R = 0.4$ and re-clustered with Cambridge/Aachen using FastJet3.3.1 [58]. The analysis is performed on those jets that satisfy: $800 < p_t < 1000$ GeV and $|\eta| < 1.5$. For the Monte-Carlo studies, we used the DyG condition given by eq. (1.1).

JHEP07(2021)020

We show the comparison between our analytic result and parton level Monte-Carlo simulations with Pythia8 and Herwig7 for the relative transverse momentum of the dynamically groomed splitting in figure 5. A crucial point to understand the fixed-order dominated regime, i.e. $k_{t,g} \gtrsim 10^{-1}$, is that in the default setting of both Monte-Carlos, the parton shower starts off a leading-order $2\to2$ matrix element.[15] Therefore, the fixed-order contribution to the $k_{t,g}$ distribution is exactly zero for these event generators. Hence, at large $k_{t,g}$, an exact agreement between our analytic result, dominated by the exact NLO matrix element, and the Monte-Carlos, where $k_{t,g}$ is exclusively generated by the parton shower, is not expected. Nevertheless, both event generators use at the very least the leading order Altarelli-Parisi splitting functions. As we have discussed in section 2.4.3, the use of the full splitting function in the resummation (or, similarly, in the parton shower) effectively reproduces the fixed-order result in the narrow jet limit. Then, part of the higher-order corrections to the Born-level process are incorporated through the splitting function in the parton shower. This can explain the nice agreement between the analytic result and the Monte-Carlo curves for $k_{t,g} \gtrsim 10^{-1}$.

On the resummation side, both showers in Herwig are in relatively good agreement and notably differ from Pythia. This is, a priori, rather counterintuitive based on the nature of the three parton showers that we are evaluating. The dipole-style Herwig shower and the Pythia one use a Catani-Seymour like [59] dipole map, transverse momentum ordering and implement a local recoil scheme. In turn, Herwig7-AO evolves through $1\to2$ splittings by means of a generalised angular variable and employs a global recoil scheme. Based on these general arguments, one would expect Herwig7-Dip and Pythia showers to deliver somewhat similar results. The opposite behavior observed in figure 5 points out to a more general Pythia-to-Herwig difference rather than to the showers themselves. We identify the scale at which the QCD shower is stopped to be the source of this discrepancy. In fact, in the default setting, Pythia imposes a relatively low infra-red cut-off of 0.5 GeV, while Herwig uses a more conservative scale of $\sim$ 1 GeV that is common to both showers [60]. Then, more phase-space is available for radiation in the Pythia case and this leads to the differences observed on the low $k_{t,g}$ side in figure 5. Thus, we conclude that the small-$k_{t,g}$ part of the differential distribution is sensitive to the way the infra-red is handled and thus to hadronisation. This point will be further emphasized in the following section. Moreover, any higher order term contained in the Monte-Carlo and not present at $\mathrm{N^2DL}$ in the resummation, e.g. energy-momentum conservation, would affect the low $k_{t,g}$ regime.

In what concerns the comparison between MCs and the analytic result, an enhancement at low $k_{t,g}$ values appears. A very similar trend is observed in figure 11 of appendix B where we evaluate the impact of removing the $1-z$ in the hardness variable κ (see eq. (1.1)) for the Monte-Carlo results. We remind the reader that this factor is a sub-leading, non-logarithmic correction in our analytic calculation at $\mathrm{N^2DL}$ accuracy. However, this mismatch in the κ definition on the analytics and the Monte-Carlos amounts to a $\sim 10\%$ difference on the low $k_{t,g}$ regime and is, therefore, partly responsible for the bump at $k_{t,g} \sim 10^{-3}$.

[15] The α_s counting might be misleading at this point. Notice that what we refer to as LO in the analytic result is actually a NLO contribution in the sense that it enters at order α_s, i.e. we consider $p + p \to jj$ at NLO.

JHEP07(2021)020

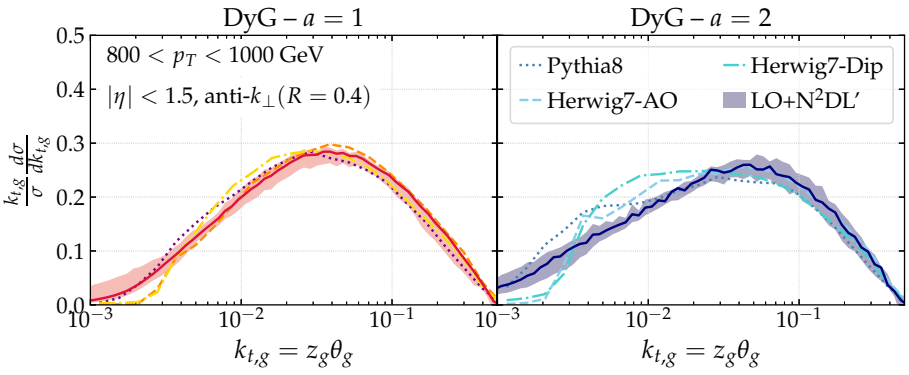


Figure 5. Theory to parton level comparison of $k_{t,g}$ in the ATLAS-like scenario for $a=1$ (left) and $a=2$ (right) in the Dynamical Grooming condition, see eq. (1.1).

The distribution of the opening angle θ_g is displayed in figure 6. On the Monte-Carlo side, we again observe a strong sensitivity to the momentum scale at which the shower is cut-off. In fact, no significant differences are observed between the angle tagged by Herwig7-AO and Herwig7-Dip, thus indicating a strong dominance of the choice of IR-scale. In particular, we have checked that the small angle bump for $a=2$ disappears if the infra-red scale is lowered down in Herwig. We have pinned down two other sources for the analytic-to-Monte-Carlo discrepancy. The first one concerns again the $1-z$ factor in the definition of κ. In figure 11 in appendix B, we quantify this effect and observe that the small θ_g region can be distorted by $\sim 20-40\%$, depending on the value of a. The enhancement of θ_g at large angles in the MCs with respect to the analytic curve could be explained by the $\mathcal{O}(\theta_g^n)$ terms that we have neglected all along our calculation, both on the resummation side and also on the matching procedure where power suppressed terms in the fixed order result were ignored. On the other hand, it is not guaranteed that the branching kernels implemented in the Monte-Carlos recover the exact soft and large angle limit. Then, we conclude that the disagreement between the analytic calculation of the θ_g-distribution and the parton shower results can be understood as a result of the choice of the infra-red scale, the finite z corrections in the κ definition and the jet clustering procedure, being the first one the strongest effect.

Finally, the momentum sharing splitting fraction z_g is presented in figure 7. We clearly observe the presence of the dynamically generated cut-off that separates the fall-off of the distributions from the flattening. The latter starts earlier for $a=2$ given that $z_{\rm cut}$ is smaller in this case, see eq. (2.14). The agreement between the theory calculation and the Monte-Carlos is reasonable in the intermediate regime of $10^{-2}<z_g<10^{-1}$. Outside this interval, the recoil factor in the hardness definition is responsible for both the depletion at large z_g in the MC's with respect to the analytic as well as for the excess at small-z, as can be seen in appendix B. In addition, the reduced phase space for emissions at infra-red scales in Herwig as compared to Pythia is manifest and further studied in appendix D.

JHEP07(2021)020

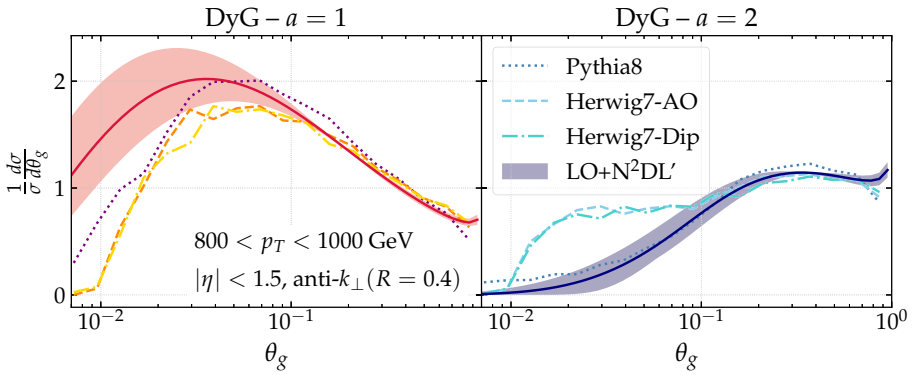


Figure 6. Same as figure 5 but for θ_g.

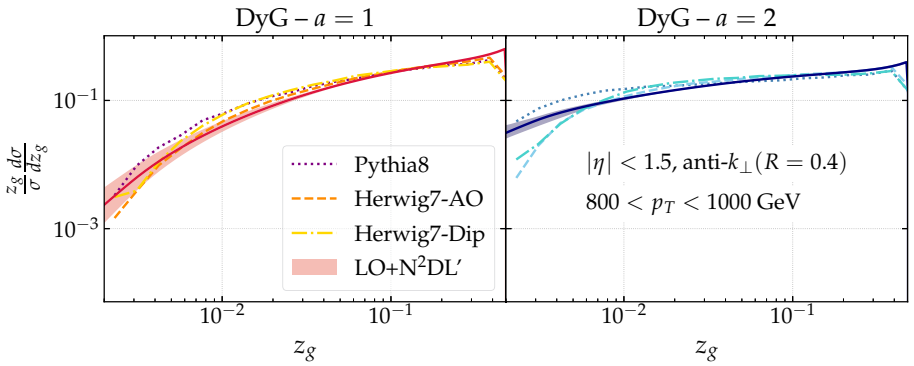


Figure 7. Same as figure 6 but for z_g.

3.2 Comparison to preliminary ALICE data

In this section, we scrutinise our analytic calculation against preliminary measurements of DyG observables [31, 32]. The experimental analysis is performed at $\sqrt{s} = 5.02$ TeV on jets clustered with the anti-k$_\perp$ algorithm with $R = 0.4$. An important aspect is that only charged jets that satisfy $60 < p_t^{ch} < 80$ GeV and $|\eta| < 0.5$ are considered. In the analytic calculation, no distinction is made between charged and neutral particles, thus a method to translate the charged p_t bin of the data into its full counterpart has to be designed. A few possibilities exist to tackle this problem. One could be to identify via Monte-Carlo simulations the transverse momentum bin that, after subtracting the neutral component, yields most of the jets in the $60 < p_t < 80$ GeV interval. We have carried out this exercise and found that 80% of the jets that fulfil $64.5 < p_t < 102.5$ GeV, fall in the $60 < p_t^{ch} < 80$ GeV category. Another option is to absorb this p_t-shift from charged to full jets into a non-perturbative factor that also accounts for the effect of hadronization, initial

JHEP07(2021)020

state radiation and multi-parton interactions. This latter is the approach followed in this paper (see also ref. [8]) and works as follows. We perform the analytic calculation in the same p_t-bin as where the experimental measurement is carried out, i.e. $60 < p_t < 80\,\text{GeV}$. Then, to construct the non-perturbative factor two samples have to be generated with Monte-Carlo. The first one includes all non-perturbative effects and only charged particles are clustered. Then, the Dynamical Grooming analysis is performed on the charged jets that satisfy $60 < p_t^{ch} < 80\,\text{GeV}$. The second sample is generated as parton level events. Again, we select jets in the same p_t-bin as the theoretical calculation, i.e. $60 < p_t < 80\,\text{GeV}$, without any charge selection. Then, our non-perturbative factor is defined as the ratio of the charged hadron level sample and the parton level one, both computed in the same transverse momentum bin. These results are plotted in appendix D, where further details on the role of the infra-red cutoff in the parton shower are provided. Finally, the theoretical results are multiplied by this phenomenological parameter. Then, the theoretical error band includes both the uncertainty of the non-perturbative factor and the analytic uncertainties characterised in the end of section 2.3.3. We label these results as 'LO+N^{2}DL'+NP'.

Like in the previous section, we start the discussion with the $k_{t,g}$ distribution shown in figure 8. To begin with, an important remark is that a mismatch exists between how the $k_{t,g}$ is defined in the analytic calculation, i.e. $k_{t,g} = z_g R_g$,[16] and in ALICE's measurement, where $k_{t,g} = z_g \sin(R_g) p_t$. As we have already mentioned, p_t degradation is ignored in our calculation because the hardest branching is located on the primary Lund plane at our degree of accuracy. Then, to accommodate the p_t dependence of the experimental definition we simply multiply our analytic result by the lower bound of the p_t bin, 60 GeV in this case. We have checked that changing this factor by any other value within the explored p_t-bin leads to variations that are well covered by our uncertainty bands. The functional form of the angular dependence of the two $k_{t,g}$ definitions is a bit more delicate. This is so because considering $\sin(R_g)$ instead of R_g brings additional power-corrections in the calculation that we have so far neglected based on our narrow jet approximation. Besides this fact, we observe in figure 8 that the data points are only described by the theoretical calculation if the non-perturbative factor, displayed in figure 13, is included. In particular, its role is most prominent for the first bin and generates a large uncertainty. This is yet another manifestation of the different methods that Pythia and Herwig employ to regulate the infra-red sector in the shower. We also provide the Monte-Carlo to data comparison in appendix E and find that all three explored setups result into $10-20\%$ deviations with respect to the data both for $a = 1$ and $a = 2$. Therefore, we conclude that the agreement between the analytic result presented in this paper and ALICE data is satisfactory in spite of the low p_t selection where hadronization effects are very large.

Turning to θ_g, represented in figure 9, we observe that the ad-hoc non-perturbative factor completely dominates the result in the first bin for $a = 1$. The a-dependence of these results is also interesting from the point of view of missing terms in the analytic calculation. Indeed, we see a deficit in the analytic result for splittings with angles $\theta_g > 0.6$

[16] Notice that in the previous section we have used $k_{t,g} = z_g \theta_g$ and now we replace θ_g by $R_g = \theta_g R$ to follow ALICE convention.

JHEP07(2021)020

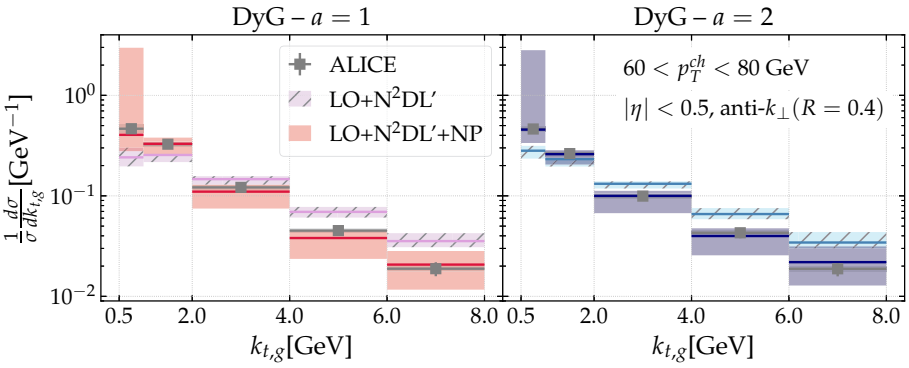


Figure 8. Comparison between the analytic result obtained in this paper and the preliminary ALICE data [32] of $k_{t,g}$.

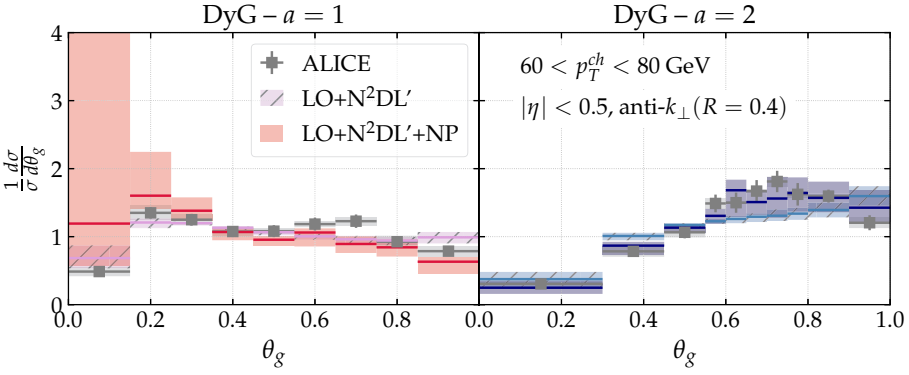


Figure 9. Same as figure 8 for θ_g. The experimental data was obtained from [31].

that disappears for the $a = 2$. We have already mentioned that the g_{nm} coefficients are inversely proportional to a and thus higher-order terms would impact less the $a = 2$ result than the $a = 1$ one. In addition, the discrepancy appears in the region where the power-suppressed terms that we have neglected all along the calculation based on the narrow jet approximation, i.e. contributions of $\mathcal{O}(\theta_g^n)$, may matter. An obvious way to confirm this hypothesis would be either to include them or, alternatively, to make a jet radius scan of this observable on the experimental side.

To end up this phenomenological section, we present the theory-to-data comparison for the momentum sharing fraction in figure 10. In this scenario, the non-perturbative factor is prominent at small z_g, but has a relatively mild effect for $z_g > 0.2$. In fact, in this interval both the LO+N^{2}DL' and LO+N^{2}DL'+NP results agree with the data within uncertainties. Clearly, this observable together with $k_{t,g}$ are the ones for which our theoretical calculation

JHEP07(2021)020

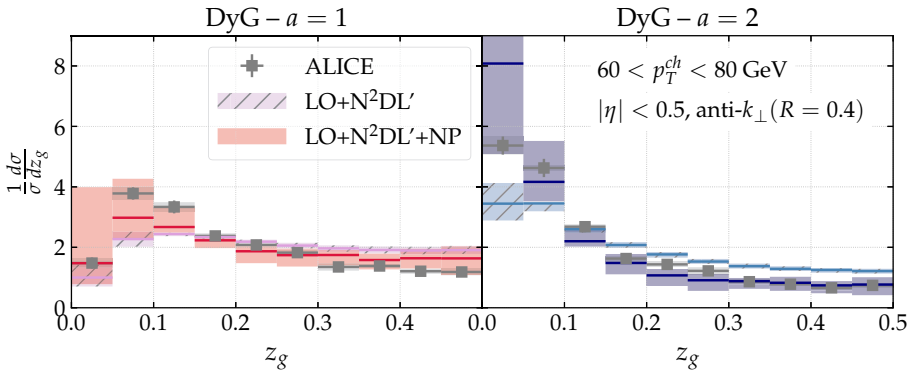


Figure 10. Same as figure 9 for z_g.

results provides the best description of the experimental measurements. This is a remarkable result given that the Sudakov nature of this observable complicates its theoretical analysis in many different ways, as we have discussed throughout the paper. The Monte-Carlo results are also consistent with the experimental measurement (see appendix E).

4 Conclusions and outlook

The work presented in this paper follows the current global effort towards a precise theoretical description of jet substructure observables that will help us to deepen our understanding of the space-time evolution of QCD jets, both in vacuum and in heavy-ion collisions. In particular, we have focused on three substructure observables defined on the splitting selected by the Dynamical Grooming method, that is, the hardest one in the jet tree. These three observables are the momentum sharing fraction z_g, the opening angle θ_g and the relative transverse momentum $k_{t,g}$. Out of the three, $k_{t,g}$ is the only infra-red and collinearly safe observable. Then, the definition of logarithmic accuracy for z_g and θ_g is far from trivial and we extensively discuss a possible approach to tackle the problem that consists in defining the accuracy of the Sudakov safe observable in terms of the cumulative distribution of an IRC safe companion. In this way, we demonstrate that the resummation of dynamically groomed observables does not exponentiate, in general, and that a logarithmic counting at the level of the cumulative distribution is therefore better suited. Further, we present all the necessary ingredients to reach next-to-next-to-double logarithmic accuracy in the narrow jet limit. This includes: (i) the resummation of collinear logarithms arising from the running of the coupling and the hard-collinear correction to the splitting function, (ii) the contribution of non-global logarithms at $\mathcal{O}(\alpha_s^2)$ and (iii) the $\mathcal{O}(\alpha_s^2)$ contribution of boundary logarithms in the case of θ_g. Remarkably, neither clustering logarithms nor multiple emissions affect these dynamically groomed distributions. We make use of a matching scheme that naturally includes the C_1 term and allows us to recover the exact leading-order result, computed with MadGraph, for large values of $k_{t,g}$. On the Sudakov

JHEP07(2021)020

safe cases, we employ the full splitting function as a proxy for the fixed-order result and propose a dedicated matching scheme that depends on an ad-hoc cut-off. All in all, we achieve LO+N^{2}DL accuracy in our analytic computation.

The analytic framework is tested against three different parton-level Monte-Carlo simulations at high-p_t: Pythia8, angular-ordered, and dipole-style showers in Herwig. In the resummation dominated regime, we find a strong dependence of the Monte-Carlo results on the transverse momentum cut-off at which the parton shower stops. This value is smaller for Pythia ($\sim 0.5\,\text{GeV}$) than for Herwig ($\sim 1\,\text{GeV}$) by default and, as such, more radiation is allowed in the former case. The analytic result regulates the infra-red singularity through a freezing of the running coupling below 1 GeV and, therefore, allows for emissions with all possible transverse momenta. Due to this fact, it is reasonable that the analytic result is closer to Pythia than to Herwig. Even in the region in which the analytic calculation reduces to the fixed-order contribution, the parton shower is fully responsible of the Monte-Carlo results, given that a leading-order matrix element is implemented by default. Despite this mismatch, an overall good agreement is found for the three jet substructure observables that we attribute to the use of the full splitting function in the parton showers that, as we have stated, generates the correct matrix element in the narrow jet limit.

Our last step is to compare the analytic predictions against the preliminary ALICE data. To do so, we supplement the perturbative results with a non-perturbative factor extracted from Monte-Carlo simulations with Pythia and Herwig that accounts for the use of charged tracks, hadronisation and underlying event. This ingredient is particularly crucial in this experimental setup given that only low-p_t jets, i.e. $p_t^{ch} \lesssim 200\,\text{GeV}$ are measurable by the ALICE detector. In fact, it dominates the theoretical prediction in the lower bins of the $k_{t,g}, z_g$ and θ_g distributions. A quantitative description of $k_{t,g}$ and z_g is achieved up to $5-10\%$ deviations in some bins. In the case of θ_g, we find deviations of up to 15% in the moderately large angle region. This is precisely the regime in which we have less confidence in our result considering that we have neglected all power suppressed terms of the type $\mathcal{O}(\theta_g^n)$.

The natural extension of this work is to go beyond the small-R limit and develop a numerical routine to account for the resummation of non-global and boundary logarithms. Notice that, as far as we are aware, the latter has yet never been achieved in the literature. On the collinear side of the resummation, we could include the recoil of the hard branch, make use of the NLO splitting function together with higher orders in the running coupling. Further, promoting our leading-order matching to NLO is straightforward from a conceptual point of view in the case of $k_{t,g}$ and would improve the agreement with data/parton level MC simulations both at low and high p_t and reduce the theoretical uncertainties. Regarding the Sudakov safe observables, we would like to understand the feasibility of the dedicated matching scheme proposed in this paper and, specifically, quantify its dependence on the ad-hoc cut off. Finally, it would be insightful to compare these analytic results for dynamically groomed observables to the newly developed parton showers that aim at achieving perturbative control beyond leading double logarithmic accuracy and leading color [35, 36, 61]. From an experimental perspective, these theoretical efforts would highly benefit from a high-p_t measurement where non-perturbative corrections are deemed to be milder.

JHEP07(2021)020

Beyond the possible improvements of the $p+p$ calculation, we would like to discuss two further extensions in terms of collision systems: $e+p$ (and eventually $e+A$), relevant for the future Electron Ion Collider and heavy-ion collisions. The former, despite the relatively low number of constituents per jet [62], provides a cleaner environment with respect to $p+p$ given that both multi-parton interactions and the underlying event will have a residual effect. On the heavy-ion side, dynamically groomed observables can be used to characterise the properties of an in-medium parton shower. In particular, the θ_g distribution can be used to experimentally measure the critical resolution angle of the Quark-Gluon Plasma [63], while deviations at large $k_{t,g}$ from respect to the vacuum baseline could suggest rare, hard scatterings between the propagating parton and the medium.

Acknowledgments

We thank Leticia Cunqueiro, Raymond Ehlers and James Mulligan for clarifications on the experimental aspects of the measurements reported in [31, 32]. Further, we acknowledge Keith Hamilton, Silvia Ferrario-Ravasio, Yacine Mehtar-Tani, Gavin Salam, Gregory Soyez and Konrad Tywoniuk for many insightful discussions on different aspects of this calculation and their feedback on the manuscript. A.S.O.'s work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 788223, PanScales). A.T. is supported by a Starting Grant from Trond Mohn Foundation (BFS2018REK01), the MCnetITN3 H2020 Marie Curie Initial Training Network, contract 722104, and wishes to thank the Institut de Physique Théorique (IPhT) and Gregory Soyez for the hospitality. PC is also grateful to the IPhT for its support during the early stages of this work.

A Details of analytic calculations at N^{2}DL'

The purpose of this section is to provide the analytic Sudakov form factor needed to achieve N^{2}DL' accuracy as explained in section 2.3. To that end, we need to perform the following integral:

$$-\ln(\Delta_i(\kappa|a)) = \frac{2C_i}{\pi}\int_0^{e^{-B_i}}\frac{\mathrm{d}z}{z}\int_0^1\frac{\mathrm{d}\theta}{\theta}\alpha_s(\mu_K z\theta Q)\Theta(z\theta^a-\kappa)\,, \tag{A.1}$$

with $Q = p_{t,\mathrm{jet}}R$ and the coupling is frozen in the infra-red as $\alpha_s(k_t) = \alpha_s^{2\ell}(k_t)\Theta(k_t - \mu_{\mathrm{fr}}) + \alpha_s(\mu_{\mathrm{fr}})\Theta(\mu_{\mathrm{fr}} - k_t)$. We would like to point out that this choice is completely ad-hoc and one could think of a more general functional form for the infra-red description of the coupling and systematically study its impact on jet substructure observables. We will investigate this possibility in future studies. In the perturbative domain $k_t > \mu_{\mathrm{fr}}$, the two-loop running coupling $\alpha_s^{2\ell}(k_t)$ is given by eq. (2.47) with the reference α_s value taken at the renormalization scale $\mu_R Q$.

We define the following dimensionless variable: $\lambda_\kappa = 2\alpha_s\beta_0\ln(\kappa)$, $\lambda_B = -2\alpha_s\beta_0 B_i$, $\lambda_K = 2\alpha_s\beta_0\ln(\mu_K/\mu_R)$ and $\lambda_{\mathrm{fr}} = 2\alpha_s\beta_0\ln(\mu_{\mathrm{fr}}/(\mu_K Q))$, and the following functions:

$$W(x) = -x + x\ln(x)\,, \tag{A.2}$$

$$V(x) = \ln(x)(2+\ln(x))\,. \tag{A.3}$$

JHEP07(2021)020

Due to the presence of the constant $k_t = \mu_{\rm fr}$ line in the (z, θ) phase space, the formulae depend on whether a is larger or smaller than 1.

Case $a > 1$. If $\lambda_\kappa \geq \lambda_{\rm fr}$:

$$\begin{aligned}
-\ln \Delta_i(\kappa|a) = &\frac{C_i}{2\pi\alpha_s\beta_0^2}\left[W(1+\lambda_K+\lambda_B) + \frac{1}{a-1}\Big(W(1+\lambda_K+\lambda_\kappa)\right.\\
&\left.-aW\left(1+\lambda_K+\frac{a-1}{a}\lambda_B+\frac{\lambda_\kappa}{a}\right)\Big)\right]\\
&+\frac{C_i\beta_1}{4\pi\beta_0^3}\left[V(1+\lambda_K+\lambda_B) + \frac{1}{a-1}\Big(V(1+\lambda_K+\lambda_\kappa)\right.\\
&\left.-aV\left(1+\lambda_K+\frac{a-1}{a}\lambda_B+\frac{\lambda_\kappa}{a}\right)\Big)\right]\\
&-\frac{C_iK}{4\pi^2\beta_0^2}\left[\ln(1+\lambda_K+\lambda_B) + \frac{1}{a-1}\Big(\ln(1+\lambda_K+\lambda_\kappa)\right.\\
&\left.-a\ln\left(1+\lambda_K+\frac{a-1}{a}\lambda_B+\frac{\lambda_\kappa}{a}\right)\Big)\right]. \qquad \text{(A.4)}
\end{aligned}$$

If $\lambda_\kappa \leq a\lambda_{\rm fr} + (1-a)\lambda_B$:

$$\begin{aligned}
-\ln \Delta_i(\kappa|a) = &\frac{C_i}{2\pi\alpha_s\beta_0^2}\left[-\lambda_B+\lambda_{\rm fr}+(1+\lambda_B+\lambda_K)\ln\left(\frac{1+\lambda_B+\lambda_K}{1+\lambda_{\rm fr}+\lambda_K}\right)\right]\\
&+\frac{C_i\beta_1}{4\pi\beta_0^3}\left[V(1+\lambda_B+\lambda_K) - \frac{2\lambda_B-2\lambda_{\rm fr}+2(1+\lambda_B+\lambda_K)\ln(1+\lambda_{\rm fr}+\lambda_K)}{1+\lambda_{\rm fr}+\lambda_K}\right.\\
&\left.+\ln^2(1+\lambda_K+\lambda_{\rm fr})\right]\\
&-\frac{C_iK}{4\pi^2\beta_0^2}\left[\frac{\lambda_{\rm fr}-\lambda_B}{1+\lambda_{\rm fr}+\lambda_K}+\ln\left(\frac{1+\lambda_B+\lambda_K}{1+\lambda_{\rm fr}+\lambda_K}\right)\right]\\
&+\frac{2C_R\alpha_s(\mu_{\rm fr})}{4\pi\alpha_s^2\beta_0^2}\left[\frac{(1-a){\lambda_B}^2+2a\lambda_B\lambda_{\rm fr}-2\lambda_B\lambda_\kappa-a{\lambda_{\rm fr}}^2+{\lambda_\kappa}^2}{2a}\right]. \qquad \text{(A.5)}
\end{aligned}$$

If $a\lambda_{\rm fr} + (1-a)\lambda_B < \lambda_\kappa < \lambda_{\rm fr}$:

$$\begin{aligned}
-\ln \Delta_i(\kappa|a) = &\frac{C_i}{2\pi\alpha_s\beta_0^2}\left[\frac{\lambda_{\rm fr}-\lambda_\kappa}{1-a}\ln(1+\lambda_K+\lambda_{\rm fr}) + W(1+\lambda_K+\lambda_B)\right.\\
&\left.\frac{1}{a-1}\left(W(1+\lambda_K+\lambda_{\rm fr}) - aW\left(1+\lambda_K+\frac{a-1}{a}\lambda_B+\frac{\lambda_\kappa}{a}\right)\right)\right]\\
&+\frac{C_i\beta_1}{4\pi\beta_0^3}\left[\frac{2(\lambda_{\rm fr}-\lambda_\kappa)}{1-a}\frac{(1+\ln(1+\lambda_K+\lambda_{\rm fr}))}{1+\lambda_K+\lambda_{\rm fr}} + V(1+\lambda_K+\lambda_B)\right.\\
&\left.\frac{1}{a-1}\left(V(1+\lambda_K+\lambda_{\rm fr}) - aV\left(1+\lambda_K+\frac{a-1}{a}\lambda_B+\frac{\lambda_\kappa}{a}\right)\right)\right]\\
&-\frac{C_iK}{4\pi^2\beta_0^2}\left[\frac{\lambda_{\rm fr}-\lambda_\kappa}{1-a}\frac{1}{1+\lambda_K+\lambda_{\rm fr}} + \ln(1+\lambda_K+\lambda_B)\right.\\
&\left.\frac{1}{a-1}\left(\ln(1+\lambda_K+\lambda_{\rm fr}) - a\ln\left(1+\lambda_K+\frac{a-1}{a}\lambda_B+\frac{\lambda_\kappa}{a}\right)\right)\right]\\
&+\frac{C_R\alpha_s(\mu_{\rm fr})}{2\pi\alpha_s^2\beta_0^2}\frac{(\lambda_{\rm fr}-\lambda_\kappa)^2}{2(a-1)}. \qquad \text{(A.6)}
\end{aligned}$$

JHEP07(2021)020

Case $a = 1$. It is straightforward, albeit tedious, to take the limit $a \to 1$ of the previous formulae. If $\lambda_\kappa > \lambda_{\rm fr}$,

$$
\begin{aligned}
-\ln \Delta_i(\kappa|1) = & \frac{C_i}{2\pi\alpha_s\beta_0^2}\left[-\lambda_B+\lambda_\kappa+(1+\lambda_B+\lambda_K)\ln\left(\frac{1+\lambda_B+\lambda_K}{1+\lambda_\kappa+\lambda_K}\right)\right] \\
& + \frac{C_i\beta_1}{4\pi\beta_0^3}\left[V(1+\lambda_B+\lambda_K) - \frac{2\lambda_B-2\lambda_\kappa+2(1+\lambda_B+\lambda_K)\ln(1+\lambda_\kappa+\lambda_K)}{1+\lambda_\kappa+\lambda_K}\right. \\
& \qquad \left. + \ln^2(1+\lambda_K+\lambda_\kappa)\right] \\
& - \frac{C_iK}{4\pi^2\beta_0^2}\left[\frac{\lambda_\kappa-\lambda_B}{1+\lambda_\kappa+\lambda_K} + \ln\left(\frac{1+\lambda_B+\lambda_K}{1+\lambda_\kappa+\lambda_K}\right)\right], \qquad \text{(A.7)}
\end{aligned}
$$

and if $\lambda_\kappa < \lambda_{\rm fr}$

$$
\begin{aligned}
-\ln \Delta_i(\kappa|1) = & \frac{C_i}{2\pi\alpha_s\beta_0^2}\left[-\lambda_B+\lambda_{\rm fr}+(1+\lambda_B+\lambda_K)\ln\left(\frac{1+\lambda_B+\lambda_K}{1+\lambda_{\rm fr}+\lambda_K}\right)\right] \\
& + \frac{C_i\beta_1}{4\pi\beta_0^3}\left[V(1+\lambda_B+\lambda_K) - \frac{2\lambda_B-2\lambda_{\rm fr}+2(1+\lambda_B+\lambda_K)\ln(1+\lambda_{\rm fr}+\lambda_K)}{1+\lambda_{\rm fr}+\lambda_K}\right. \\
& \qquad \left. + \ln^2(1+\lambda_K+\lambda_{\rm fr})\right] \\
& - \frac{C_iK}{4\pi^2\beta_0^2}\left[\frac{\lambda_{\rm fr}-\lambda_B}{1+\lambda_{\rm fr}+\lambda_K} + \ln\left(\frac{1+\lambda_B+\lambda_K}{1+\lambda_{\rm fr}+\lambda_K}\right)\right] \\
& + \frac{C_R\alpha_s(\mu_{\rm fr})}{2\pi\alpha_s^2\beta_0^2}\left[(\lambda_{\rm fr}-\lambda_\kappa)(\lambda_B-\lambda_{\rm fr}) + \frac{1}{2}(\lambda_{\rm fr}-\lambda_\kappa)^2\right]. \qquad \text{(A.8)}
\end{aligned}
$$

Case $a < 1$. For completeness, we provide also the formulae when $a < 1$. For some values of λ_κ, they can be related to the expression in the $a > 1$ case. If $\lambda_\kappa \geq a\lambda_{\rm fr} + (1-a)\lambda_B$, $\Delta(\kappa|a<1)$ is given by the expression of $\Delta(\kappa|a>1)$ when $\lambda_\kappa > \lambda_{\rm fr}$. In a similar way, when $\lambda_\kappa < \lambda_{\rm fr}$, $\Delta(\kappa|a<1)$ is given by the expression of $\Delta(\kappa|a>1)$ when $\lambda_\kappa \leq a\lambda_{\rm fr} + (1-a)\lambda_B$. In the remaining κ domain, $\lambda_{\rm fr} \leq \lambda_\kappa \leq a\lambda_{\rm fr} + (1-a)\lambda_B$, one finds

$$
\begin{aligned}
-\ln \Delta_i(\kappa|a) = & \frac{C_i}{2\pi\alpha_s\beta_0^2}\left[\frac{(1-a)\lambda_B+a\lambda_{\rm fr}-\lambda_\kappa}{a-1}\ln(1+\lambda_K+\lambda_{\rm fr}) + W(1+\lambda_K+\lambda_B)\right. \\
& \qquad \left. \frac{1}{a-1}(W(1+\lambda_K+\lambda_\kappa) - aW(1+\lambda_K+\lambda_{\rm fr}))\right] \\
& + \frac{C_i\beta_1}{4\pi\beta_0^3}\left[\frac{2((1-a)\lambda_B+a\lambda_{\rm fr}-\lambda_\kappa)}{a-1}\frac{(1+\ln(1+\lambda_K+\lambda_{\rm fr}))}{1+\lambda_K+\lambda_{\rm fr}} + V(1+\lambda_K+\lambda_B)\right. \\
& \qquad \left. \frac{1}{a-1}(V(1+\lambda_K+\lambda_\kappa) - aV(1+\lambda_K+\lambda_{\rm fr}))\right] \\
& - \frac{C_iK}{4\pi^2\beta_0^2}\left[\frac{(1-a)\lambda_B+a\lambda_{\rm fr}-\lambda_\kappa}{a-1}\frac{1}{1+\lambda_K+\lambda_{\rm fr}} + \ln(1+\lambda_K+\lambda_B)\right. \\
& \qquad \left. \frac{1}{a-1}(\ln(1+\lambda_K+\lambda_\kappa) - a\ln(1+\lambda_K+\lambda_{\rm fr}))\right] \\
& + \frac{C_R\alpha_s(\mu_{\rm fr})}{2\pi\alpha_s^2\beta_0^2}\left[\frac{((1-a)\lambda_B+a\lambda_{\rm fr}-\lambda_\kappa)^2}{2a(1-a)}\right]. \qquad \text{(A.9)}
\end{aligned}
$$

JHEP07(2021)020

The non-global term. On top of eq. (A.1), the Sudakov factor receives a contribution from soft non-global emissions of the form

$$-\ln(\Delta_i^{\rm NG}(\kappa|a)) = 2C_iC_A\left(\frac{\alpha_s}{2\pi}\right)^2\frac{\pi^2}{3}\int_0^{e^{-B}}\frac{\mathrm{d}z}{z}\ln\left(\frac{1}{\mu_K z}\right)\int_0^1\mathrm{d}\theta\,\Theta(z\theta^a-\kappa) \tag{A.10}$$

$$= C_iC_A\left(\frac{\alpha_s}{2\pi}\right)^2\frac{\pi^2}{3}\ln^2(\mu_K\kappa)+C_iC_A\left(\frac{\alpha_s}{2\pi}\right)^2\frac{\pi^2}{3}\Big[-\ln^2\left(\mu_K e^{-B}\right)$$
$$+2a^2\left(1-\kappa^{1/a}e^{B/a}\right)+2a\left(\ln(\kappa\mu_K)-\ln\left(e^{-B}\mu_K\right)\right)\kappa^{1/a}e^{B/a}\Big]\,. \tag{A.11}$$

In this expression, we have separated the single-log term from the pure α_s or power corrections in κ which can be neglected to N^{2}DL accuracy.

B The size of finite z corrections in the definition of κ

In this appendix, we evaluate the impact of not neglecting the $1-z$ factor on the definition of κ in eq. (1.1). From an analytic point of view, this is a sub-leading, non logarithmic correction and thus not needed to reach N^{2}DL. For example, at the level of the Sudakov $\Delta(\kappa|a)$, the $1-z$ in the definition of $k_\perp = z(1-z)p_T\theta$ for the running coupling scale induces a N^{3}DL correction of the form:

$$\delta\ln(\Delta(\kappa|a)) = \frac{4\alpha_s^2\beta_0 C_i}{\pi}\int_0^1\frac{\mathrm{d}z'}{z'}\int_0^1\frac{\mathrm{d}\theta'}{\theta'}\log(1-z')\Theta(z'\theta'^a-\kappa) \tag{B.1}$$

$$= \frac{4\alpha_s^2\beta_0 C_i}{\pi a}\left(\frac{\pi^2}{6}\ln(\kappa)+\zeta(3)+\mathcal{O}(\kappa)\right) \tag{B.2}$$

where we have used $\alpha_s(k_\perp)\simeq\alpha_s(zp_T\theta)(1-2\alpha_s\beta_0\ln(1-z))$ at our order of interest. In the same way, one can determine the magnitude of the leading correction induced by $1-z$ in the definition of $\kappa = z(1-z)(\Delta R/R)^a$ by using the double logarithmic formula (2.6) for the Sudakov, with the veto constraint including the $1-z$ factor:

$$\ln(\Delta(\kappa|a)) = -\frac{2\alpha_s C_i}{\pi}\int_0^1\frac{\mathrm{d}z'}{z'}\int_0^1\frac{\mathrm{d}\theta'}{\theta'}\Theta(z'(1-z')\theta'^a-\kappa) \tag{B.3}$$

$$= -\frac{\alpha_s C_i}{\pi a}\left(\ln^2(\kappa)-\frac{\pi^2}{3}+\mathcal{O}(\kappa)\right) \tag{B.4}$$

The correction to the double logarithmic result is therefore a sub-leading non logarithmic correction.

That said, we would like to understand its impact on Monte-Carlo results in the experimental setups explored in this paper. The results are presented in figure 11 while their implications are commented over the main text.

JHEP07(2021)020

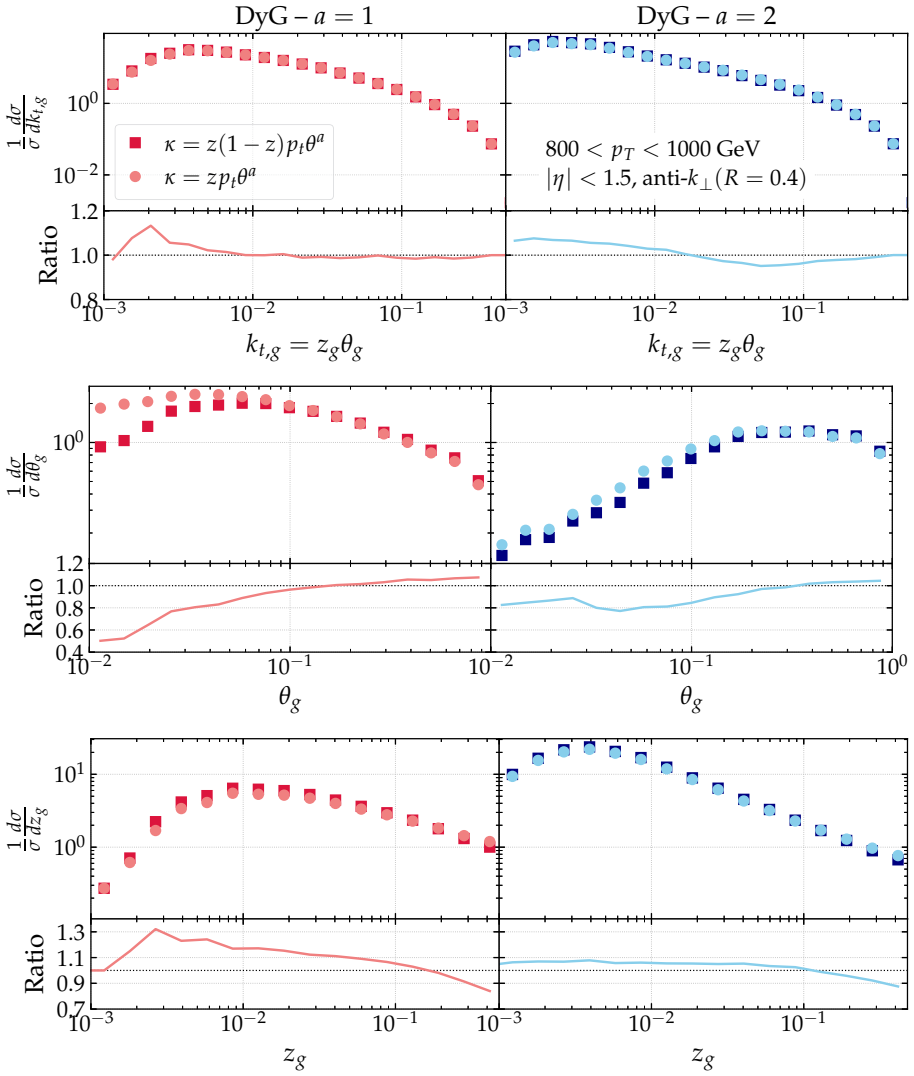


Figure 11. Impact of including or not the recoil factor in the hardness variable definition as a function of a for dijet events at parton level in Pythia with $\sqrt{s}=13$ TeV for $k_{t,g}$ (top), θ_g (center) and z_g (bottom).

JHEP07(2021)020

C Impact of jet clustering algorithms

Throughout the main text we have determined dynamically groomed substructure observables for jets found with an initial anti-k_t clustering and subsequently re-clustered with Cambridge/Aachen. This two-step process has advantages from an experimental point of view. However, from a theoretical perspective we have seen in section 2.4.1 that this two-step process induces boundary logarithms in the calculation. In this appendix, we would like to investigate at the Monte-Carlo level if the observables are modified when defining the jets only with C/A. This is shown in figure 12. Interestingly, we observe how the bump at large angles whose origin we have discussed in the main text disappears when clustering with C/A. Nevertheless, the impact of these two jet clustering strategies is mild for all cases.

D Non-perturbative corrections with Pythia and Herwig

In order to compare our analytic predictions with ALICE's experimental data, we add non-perturbative effects through a single parameter extracted from Monte-Carlo simulations. This factor, thoroughly explained in section 3.2, is provided in figure 13, where we took the ratio of MCs before and after hadronisation. Besides the default settings of Pythia and Herwig, we show two additional curves in which the parton shower cutoff, denoted as $\mu_{\rm NP}^{\rm parton}$, in Herwig is changed from its default value of 1 GeV to the number used in Pythia where $\mu_{\rm NP}^{\rm parton} = 0.5\,\text{GeV}$. Note that changing this factor does not necessarily imply a one-to-one correspondence between the two event generators. The value of this factor, like any other hadronisation-related parameter, is tuned to data. Then, one cannot vary it when running the Monte-Carlo at hadron level because its predictive power would be negatively affected. Therefore, we only vary this factor for the parton level result, that is, for the denominator of our non-perturbative factor.

The point of the variation of the parton shower stopping is to demonstrate the sensitivity of the dynamically groomed observables to that scale, and the limitations of this method for incorporating hadronisation corrections into analytic calculations. This is manifest in figure 13, where the hadron-to-parton ratio varies from 0.5 to 2.5 for those settings that share the same value of $\mu_{\rm NP}^{\rm parton}$, while it explodes for the default Herwig-AO and Herwig-Dip in the limit of non-perturbative values of $(k_{t,g}, z_g, \theta_g)$. In the latter case, the reason for the rapid growth of the non-perturbative factor, e.g. in the low k_t regime, is rooted in the fact that the parton-level shower does not generate splittings below $\mu_{\rm NP}^{\rm parton}$, while hadronization and underlying event populate this part of the phase-space. In terms of Lund planes, the area covered by the parton-level result and the hadron level one are clearly distinct in Herwig. This effect is less pronounced whenever $\mu_{\rm NP}^{\rm parton}$ is low, as in default Pythia. This is explicitly shown in figure 14.

As we have already mentioned, there is no preferred value of $\mu_{\rm NP}^{\rm parton}$ when running parton level simulations and the large variations encountered in the non-perturbative factor simply indicate that the parton-level results are out of their regime of applicability. Then, we decide to use the average of the Monte-Carlo generators with the same value of $\mu_{\rm NP}^{\rm parton}$ as the central value of the non-perturbative factor. The uncertainty band is obtained from the envelope of the five MC settings.

JHEP07(2021)020

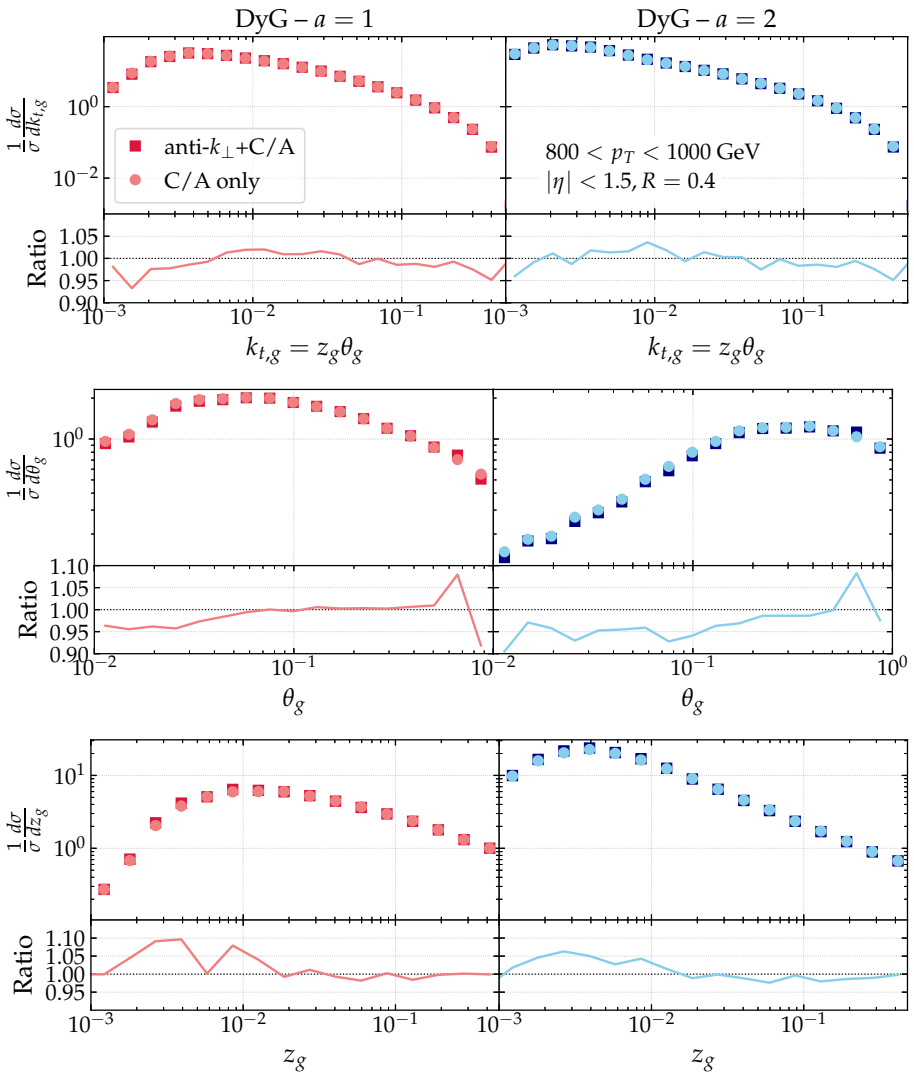


Figure 12. Impact of different clustering strategies in $k_{t,g}$ as a function of a for dijet events at parton level in PYTHIA with $\sqrt{s}$=13 TeV for $k_{t,g}$ (top), θ_g (center) and z_g (bottom).

JHEP07(2021)020

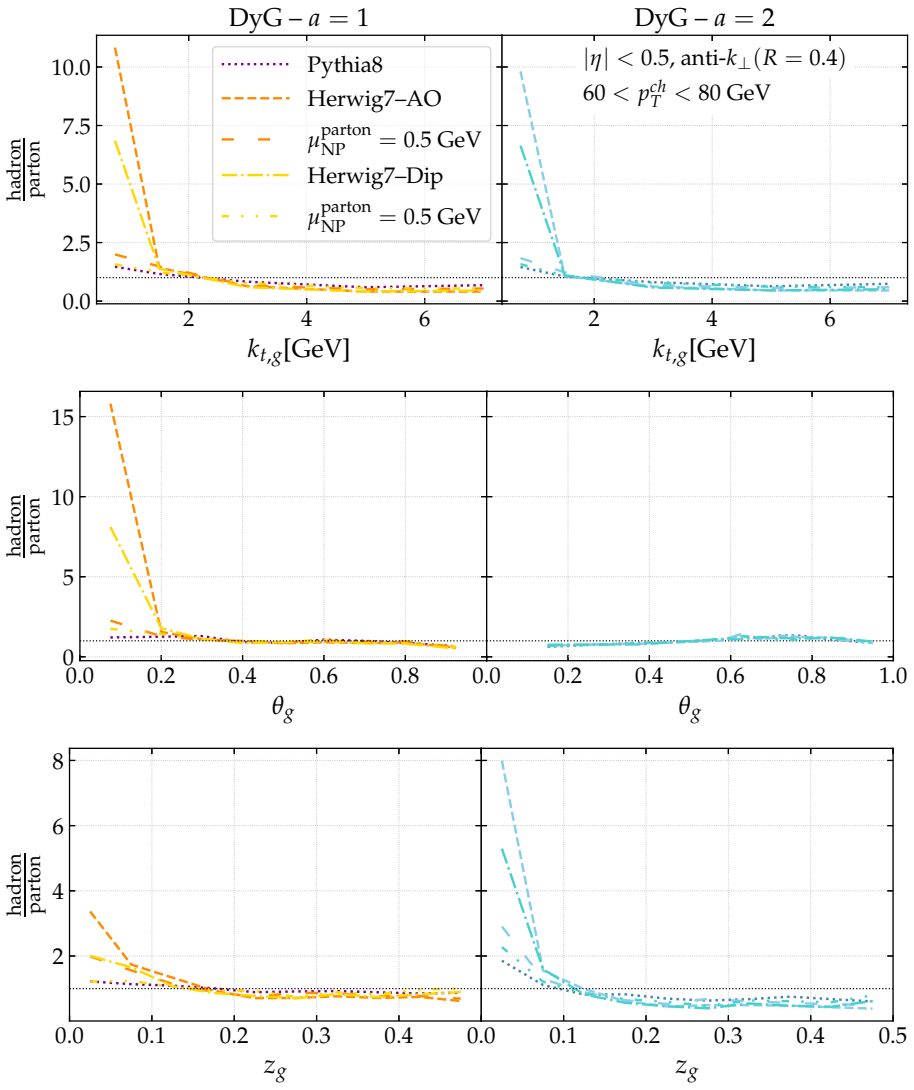


Figure 13. Ratio of hadron-to-parton level distributions for $k_{t,g}$ (top), θ_g (middle) and z_g (bottom) with five different Monte-Carlo settings: Pythia8 (dotted, purple), Herwig7-AO with default parameters (orange, dashed) and with the shower cut-off set to 0.5 GeV at parton level only (orange, loosely dashed), Herwig7-Dip with default parameters (gold, dotted dashed) and with the shower cut-off set to 0.5 GeV at parton level only (gold, loosely dotted dashed).

JHEP07(2021)020

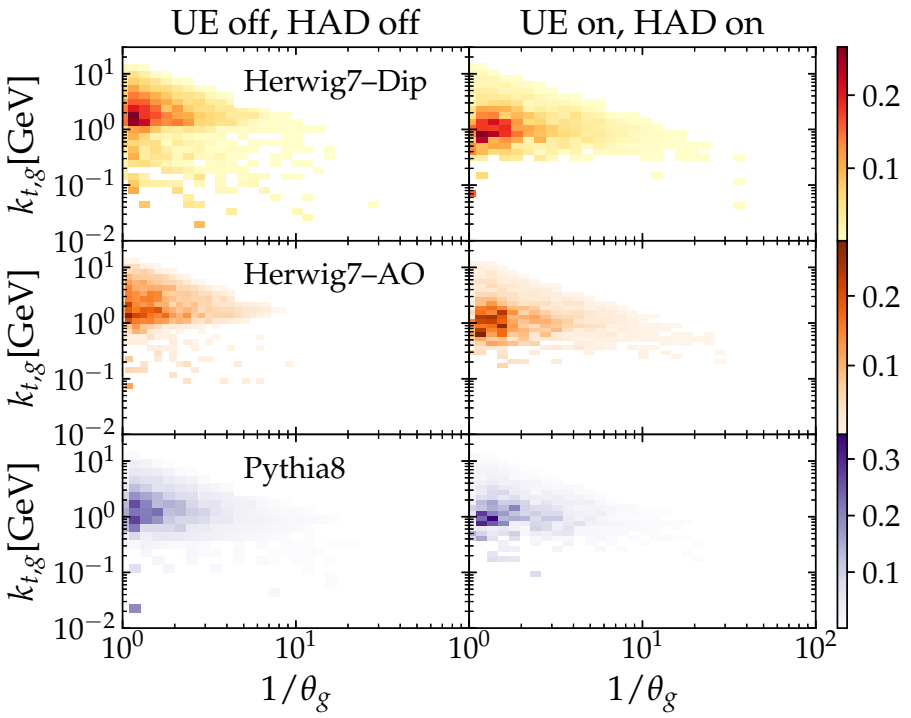


Figure 14. Lund planes generated by the three Monte-Carlo setups used in this work, Herwig7-Dip (top), Herwig-AO (center) and Pythia8 (bottom), with ALICE kinematics and $a=1$.

E Monte-Carlo description of $(z_g, \theta_g, k_{t,g})$ data

In this appendix we compare the three Monte-Carlo settings that we explore through this paper, i.e. Pythia8, Herwig7-AO and Herwig7-Dip, to the preliminary ALICE data. The results are shown in figure 15. Notice that through these comparisons we are testing simultaneously the parton shower, i.e. dipole-style or angular-ordered, and the hadronization mechanism, i.e. Lund string or cluster models. In the case of $k_{t,g}$, no significant differences are observed among all Monte-Carlos. For θ_g, Herwig7-Dip provides the best description of the data from small to large angles. All three Monte-Carlo settings are able to capture the data in the intermediate range of this measurement $0.4<\theta_g<0.7$ and differences only appear in the tails of figure 15, where the hadronization mechanism seems to dominate for $\theta_g<0.4$. Finally, all Monte-Carlos show a significant depletion at $0.2<z_g<0.3$ that is ameliorated for $a=1$. Pythia achieves the best theory-to-data ratio, but its not obvious for this observable to disentangle between parton-shower dominated differences and hadronization mechanisms.

JHEP07(2021)020

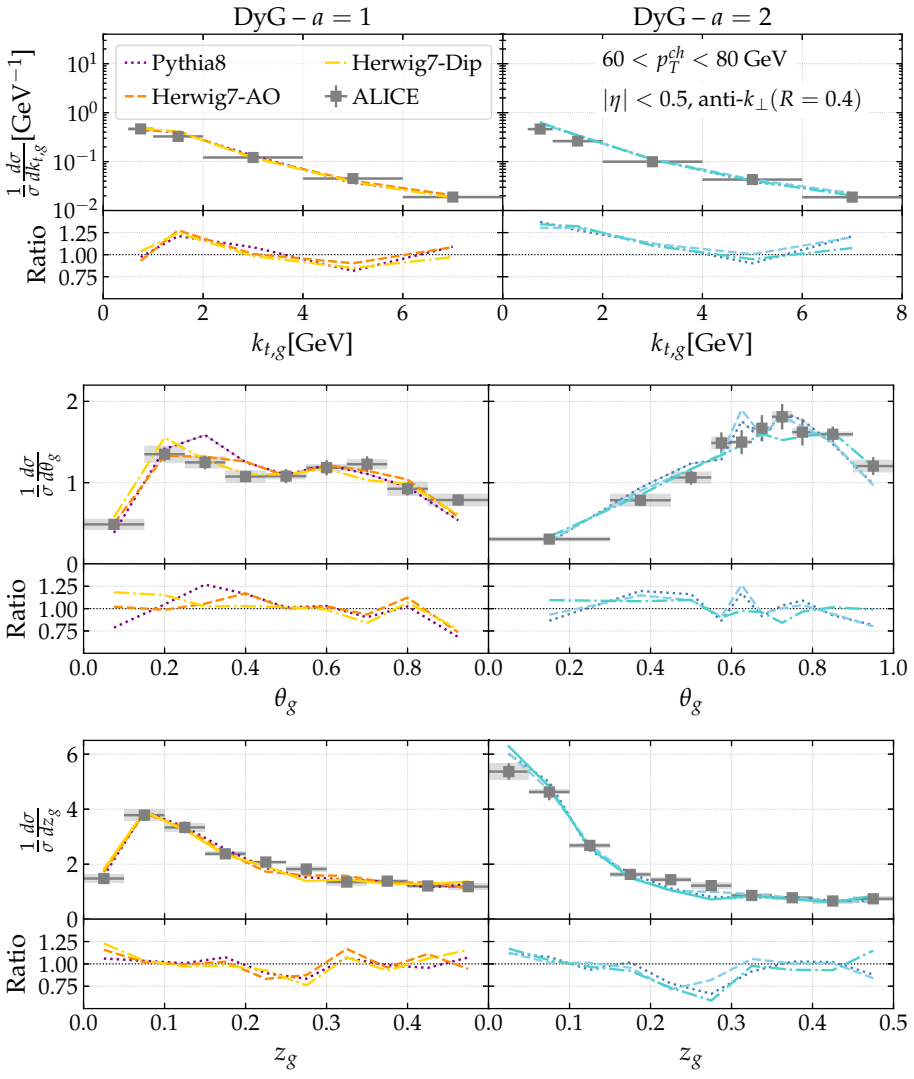


Figure 15. Monte-Carlo to data comparison of $k_{t,g}$ (top), θ_g (middle) and z_g (bottom) for $a=1$ (left) and $a=2$ (right) in the Dynamical Grooming condition, see eq. (1.1). In the bottom panels, the theory-to-data ratios, computed using the same binning as the data, are presented.

JHEP07(2021)020

References

[1] G.F. Sterman and S. Weinberg, *Jets from Quantum Chromodynamics*, *Phys. Rev. Lett.* **39** (1977) 1436 [INSPIRE].

[2] M. Connors, C. Nattrass, R. Reed and S. Salur, *Jet measurements in heavy ion physics*, *Rev. Mod. Phys.* **90** (2018) 025005 [arXiv:1705.01974] [INSPIRE].

[3] S. Marzani, G. Soyez and M. Spannowsky, *Looking inside jets: an introduction to jet substructure and boosted-object phenomenology*, vol. 958, Springer (2019) [DOI] [arXiv:1901.10342] [INSPIRE].

[4] A.J. Larkoski, S. Marzani, G. Soyez and J. Thaler, *Soft Drop*, *JHEP* **05** (2014) 146 [arXiv:1402.2657] [INSPIRE].

[5] A.J. Larkoski, S. Marzani and J. Thaler, *Sudakov Safety in Perturbative QCD*, *Phys. Rev. D* **91** (2015) 111501 [arXiv:1502.01719] [INSPIRE].

[6] J. Thaler and K. Van Tilburg, *Identifying Boosted Objects with N-subjettiness*, *JHEP* **03** (2011) 015 [arXiv:1011.2268] [INSPIRE].

[7] F.A. Dreyer, G.P. Salam and G. Soyez, *The Lund Jet Plane*, *JHEP* **12** (2018) 064 [arXiv:1807.04758] [INSPIRE].

[8] A. Lifson, G.P. Salam and G. Soyez, *Calculating the primary Lund Jet Plane density*, *JHEP* **10** (2020) 170 [arXiv:2007.06578] [INSPIRE].

[9] J.M. Butterworth, A.R. Davison, M. Rubin and G.P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, *Phys. Rev. Lett.* **100** (2008) 242001 [arXiv:0802.2470] [INSPIRE].

[10] D. Krohn, J. Thaler and L.-T. Wang, *Jet Trimming*, *JHEP* **02** (2010) 084 [arXiv:0912.1342] [INSPIRE].

[11] ATLAS collaboration, *Optimisation of large-radius jet reconstruction for the ATLAS detector in* 13 *TeV proton-proton collisions*, *Eur. Phys. J. C* **81** (2021) 334 [arXiv:2009.04986] [INSPIRE].

[12] M. Dasgupta, A. Fregoso, S. Marzani and G.P. Salam, *Towards an understanding of jet substructure*, *JHEP* **09** (2013) 029 [arXiv:1307.0007] [INSPIRE].

[13] M. Dasgupta, A. Fregoso, S. Marzani and A. Powling, *Jet substructure with analytical methods*, *Eur. Phys. J. C* **73** (2013) 2623 [arXiv:1307.0013] [INSPIRE].

[14] S. Marzani, L. Schunk and G. Soyez, *A study of jet mass distributions with grooming*, *JHEP* **07** (2017) 132 [arXiv:1704.02210] [INSPIRE].

[15] S. Marzani, L. Schunk and G. Soyez, *The jet mass distribution after Soft Drop*, *Eur. Phys. J. C* **78** (2018) 96 [arXiv:1712.05105] [INSPIRE].

[16] Z.-B. Kang, K. Lee, X. Liu and F. Ringer, *The groomed and ungroomed jet mass distribution for inclusive jet production at the LHC*, *JHEP* **10** (2018) 137 [arXiv:1803.03645] [INSPIRE].

JHEP07(2021)020

[17] M. Dasgupta, K. Khelifa-Kerfa, S. Marzani and M. Spannowsky, *On jet mass distributions in Z+jet and dijet processes at the LHC*, *JHEP* **10** (2012) 126 [arXiv:1207.1640] [INSPIRE].

[18] Z.-B. Kang, K. Lee, X. Liu, D. Neill and F. Ringer, *The soft drop groomed jet radius at NLL*, *JHEP* **02** (2020) 054 [arXiv:1908.01783] [INSPIRE].

[19] J. Baron, D. Reichelt, S. Schumann, N. Schwanemann and V. Theeuwes, *Soft-drop grooming for hadronic event shapes*, arXiv:2012.09574 [INSPIRE].

[20] D. Anderle, M. Dasgupta, B.K. El-Menoufi, J. Helliwell and M. Guzzi, *Groomed jet mass as a direct probe of collinear parton dynamics*, *Eur. Phys. J. C* **80** (2020) 827 [arXiv:2007.10355] [INSPIRE].

[21] Z.-B. Kang, K. Lee, X. Liu and F. Ringer, *Soft drop groomed jet angularities at the LHC*, *Phys. Lett. B* **793** (2019) 41 [arXiv:1811.06983] [INSPIRE].

[22] A. Kardos, A.J. Larkoski and Z. Trócsányi, *Groomed jet mass at high precision*, *Phys. Lett. B* **809** (2020) 135704 [arXiv:2002.00942] [INSPIRE].

[23] A. Kardos, G. Somogyi and Z. Trócsányi, *Soft-drop event shapes in electron-positron annihilation at next-to-next-to-leading order accuracy*, *Phys. Lett. B* **786** (2018) 313 [arXiv:1807.11472] [INSPIRE].

[24] C. Frye, A.J. Larkoski, M.D. Schwartz and K. Yan, *Factorization for groomed jet substructure beyond the next-to-leading logarithm*, *JHEP* **07** (2016) 064 [arXiv:1603.09338] [INSPIRE].

[25] Y. Mehtar-Tani, A. Soto-Ontoso and K. Tywoniuk, *Dynamical grooming of QCD jets*, *Phys. Rev. D* **101** (2020) 034004 [arXiv:1911.00375] [INSPIRE].

[26] Y. Mehtar-Tani, A. Soto-Ontoso and K. Tywoniuk, *Tagging boosted hadronic objects with dynamical grooming*, *Phys. Rev. D* **102** (2020) 114013 [arXiv:2005.07584] [INSPIRE].

[27] A. Soto-Ontoso, *Dynamical grooming at work: from p+p to Pb+Pb*, arXiv:2008.05797 [INSPIRE].

[28] Y.L. Dokshitzer, G.D. Leder, S. Moretti and B.R. Webber, *Better jet clustering algorithms*, *JHEP* **08** (1997) 001 [hep-ph/9707323] [INSPIRE].

[29] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040 [hep-ph/0409146] [INSPIRE].

[30] A.J. Larkoski and J. Thaler, *Unsafe but Calculable: Ratios of Angularities in Perturbative QCD*, *JHEP* **09** (2013) 137 [arXiv:1307.1699] [INSPIRE].

[31] ALICE collaboration, *Jet substructure measurements in* pp *and Pb-Pb collisions at* $\sqrt{s_{\rm NN}} = 5.02$ *TeV with ALICE*, arXiv:2009.07172 [INSPIRE].

[32] ALICE collaboration, *Investigating Hard Splittings via Jet Substructure in pp and Pb-Pb Collisions at* $\sqrt{s_{\rm NN}} = 5.02$ *TeV with ALICE*, in *10th International Conference on Hard and Electromagnetic Probes of High-Energy Nuclear Collisions: Hard Probes 2020*, (2020) [arXiv:2009.12247] [INSPIRE].

[33] J. Mulligan and M. Ploskon, *Identifying groomed jet splittings in heavy-ion collisions*, *Phys. Rev. C* **102** (2020) 044913 [arXiv:2006.01812] [INSPIRE].

[34] T. Sjöstrand, S. Mrenna and P.Z. Skands, *A Brief Introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852 [arXiv:0710.3820] [INSPIRE].

JHEP07(2021)020

[35] M. Dasgupta, F.A. Dreyer, K. Hamilton, P.F. Monni and G.P. Salam, *Logarithmic accuracy of parton showers: a fixed-order study*, *JHEP* **09** (2018) 033 [*Erratum ibid.* **03** (2020) 083] [arXiv:1805.09327] [INSPIRE].

[36] K. Hamilton, R. Medves, G.P. Salam, L. Scyboz and G. Soyez, *Colour and logarithmic accuracy in final-state parton showers*, arXiv:2011.10054 [INSPIRE].

[37] S. Catani, L. Trentadue, G. Turnock and B.R. Webber, *Resummation of large logarithms in e^+e^- event shape distributions*, *Nucl. Phys. B* **407** (1993) 3 [INSPIRE].

[38] A. Banfi, G.P. Salam and G. Zanderighi, *Principles of general final-state resummation and automated implementation*, *JHEP* **03** (2005) 073 [hep-ph/0407286] [INSPIRE].

[39] S. Catani, B.R. Webber and G. Marchesini, *QCD coherent branching and semiinclusive processes at large x*, *Nucl. Phys. B* **349** (1991) 635 [INSPIRE].

[40] M. Dasgupta and G.P. Salam, *Resummation of nonglobal QCD observables*, *Phys. Lett. B* **512** (2001) 323 [hep-ph/0104277] [INSPIRE].

[41] A. Banfi, M. Dasgupta, K. Khelifa-Kerfa and S. Marzani, *Non-global logarithms and jet algorithms in high-pT jet shapes*, *JHEP* **08** (2010) 064 [arXiv:1004.3483] [INSPIRE].

[42] Y. Delenda and K. Khelifa-Kerfa, *On the resummation of clustering logarithms for non-global observables*, *JHEP* **09** (2012) 109 [arXiv:1207.4528] [INSPIRE].

[43] A. Banfi, G.P. Salam and G. Zanderighi, *Phenomenology of event shapes at hadron colliders*, *JHEP* **06** (2010) 038 [arXiv:1001.4082] [INSPIRE].

[44] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079 [arXiv:1405.0301] [INSPIRE].

[45] H.-L. Lai et al., *New parton distributions for collider physics*, *Phys. Rev. D* **82** (2010) 074024 [arXiv:1007.2241] [INSPIRE].

[46] M. Cacciari, S. Frixione, M.L. Mangano, P. Nason and G. Ridolfi, *The $t\bar{t}$ cross-section at 1.8 TeV and 1.96 TeV: A Study of the systematics due to parton densities and scale dependence*, *JHEP* **04** (2004) 068 [hep-ph/0303085] [INSPIRE].

[47] M. Cacciari, G.P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, *JHEP* **04** (2008) 063 [arXiv:0802.1189] [INSPIRE].

[48] K. Khelifa-Kerfa, *Non-global logs and clustering impact on jet mass with a jet veto distribution*, *JHEP* **02** (2012) 072 [arXiv:1111.2016] [INSPIRE].

[49] ATLAS collaboration, *Measurement of soft-drop jet observables in pp collisions with the ATLAS detector at $\sqrt{s} = 13$ TeV*, *Phys. Rev. D* **101** (2020) 052007 [arXiv:1912.09837] [INSPIRE].

[50] A. Tripathee, W. Xue, A. Larkoski, S. Marzani and J. Thaler, *Jet Substructure Studies with CMS Open Data*, *Phys. Rev. D* **96** (2017) 074003 [arXiv:1704.05842] [INSPIRE].

[51] P. Caucal, A. Takacs, A. Soto-Ontoso and G. Soyez, *SD+ DyG: A two-step groomer for QCD jets*, in preparation.

[52] A. Banfi, G.P. Salam and G. Zanderighi, *Infrared safe definition of jet flavor*, *Eur. Phys. J. C* **47** (2006) 113 [hep-ph/0601139] [INSPIRE].

[53] CMS collaboration, *Measurement of the Splitting Function in pp and Pb-Pb Collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV*, *Phys. Rev. Lett.* **120** (2018) 142302 [arXiv:1708.09429] [INSPIRE].

JHEP07(2021)020

[54] STAR collaboration, *Measurement of groomed jet substructure observables in p+p collisions at $\sqrt{s} = 200$ GeV with STAR*, *Phys. Lett. B* **811** (2020) 135846 [arXiv:2003.02114] [INSPIRE].

[55] J. Bellm et al., *HERWIG 7.0/HERWIG++ 3.0 release note*, *Eur. Phys. J. C* **76** (2016) 196 [arXiv:1512.01178] [INSPIRE].

[56] S. Gieseke, P. Stephens and B. Webber, *New formalism for QCD parton showers*, *JHEP* **12** (2003) 045 [hep-ph/0310083] [INSPIRE].

[57] S. Platzer and S. Gieseke, *Coherent Parton Showers with Local Recoils*, *JHEP* **01** (2011) 024 [arXiv:0909.5593] [INSPIRE].

[58] M. Cacciari, G.P. Salam and G. Soyez, *FastJet User Manual*, *Eur. Phys. J. C* **72** (2012) 1896 [arXiv:1111.6097] [INSPIRE].

[59] S. Catani and M.H. Seymour, *A General algorithm for calculating jet cross-sections in NLO QCD*, *Nucl. Phys. B* **485** (1997) 291 [*Erratum ibid.* **510** (1998) 503] [hep-ph/9605323] [INSPIRE].

[60] J. Bellm, G. Nail, S. Plätzer, P. Schichtel and A. Siódmok, *Parton Shower Uncertainties with HERWIG 7: Benchmarks at Leading Order*, *Eur. Phys. J. C* **76** (2016) 665 [arXiv:1605.01338] [INSPIRE].

[61] M. Dasgupta, F.A. Dreyer, K. Hamilton, P.F. Monni, G.P. Salam and G. Soyez, *Parton showers beyond leading logarithmic accuracy*, *Phys. Rev. Lett.* **125** (2020) 052002 [arXiv:2002.11114] [INSPIRE].

[62] M. Arratia, Y. Song, F. Ringer and B.V. Jacak, *Jets as precision probes in electron-nucleus collisions at the future Electron-Ion Collider*, *Phys. Rev. C* **101** (2020) 065204 [arXiv:1912.05931] [INSPIRE].

[63] P. Caucal, A. Takacs and A. Soto-Ontoso, *Dynamically groomed jet radius as a test for color coherence in heavy-ion collisions*, in preparation.

JHEP07(2021)020

Article II.

5.2 Quarks and gluons in the Lund plane

Frédéric A. Dreyer, Gregory Soyezb and Adam Takacs

JHEP, **08**, 177 (2022)

Paper about quark and gluon jet classification in pp collisions. Keywords: jets, jet substructure, q/g jet tagging, machine learning, event generators.

Published for SISSA by Springer

Received: February 2, 2022
Revised: May 30, 2022
Accepted: June 22, 2022
Published: August 18, 2022

Quarks and gluons in the Lund plane

Frédéric A. Dreyer,[a] **Gregory Soyez**[b] **and Adam Takacs**[c]

[a] *Rudolf Peierls Centre for Theoretical Physics, Parks Road, Oxford OX1 3PU, U.K.*

[b] *IPhT, Université Paris-Saclay, CNRS UMR 3681, CEA Saclay, Orme des Merisiers, Bât 774, F-91191 Gif-sur-Yvette, France*

[c] *Department of Physics and Technology, University of Bergen, Allégaten 55, 5007 Bergen, Norway*

E-mail: frederic.dreyer@physics.ox.ac.uk, gregory.soyez@ipht.fr, adam.takacs@uib.no

Abstract: Discriminating quark and gluon jets is a long-standing topic in collider phenomenology. In this paper, we address this question using the Lund jet plane substructure technique introduced in recent years. We present two complementary approaches: one where the quark/gluon likelihood ratio is computed analytically, to single-logarithmic accuracy, in perturbative QCD, and one where the Lund declusterings are used to train a neural network. For both approaches, we either consider only the primary Lund plane or the full clustering tree. The analytic and machine-learning discriminants are shown to be equivalent on a toy event sample resumming exactly leading collinear single logarithms, where the analytic calculation corresponds to the exact likelihood ratio. On a full Monte Carlo event sample, both approaches show a good discriminating power, with the machine-learning models usually being superior. We carry out a study in the asymptotic limit of large logarithm, allowing us to gain confidence that this superior performance comes from effects that are subleading in our analytic approach. We then compare our approach to other quark-gluon discriminants in the literature. Finally, we study the resilience of our quark-gluon discriminants against the details of the event sample and observe that the analytic and machine-learning approaches show similar behaviour.

Keywords: Jets and Jet Substructure, Resummation

ArXiv ePrint: 2112.09140

https://doi.org/10.1007/JHEP08(2022)177

JHEP08(2022)177

Contents

1 Introduction **1**

2 Lund plane(s) and baseline discriminants **3**

3 Analytic discriminants in the Lund-plane with strong angular ordering **5**
- 3.1 Generic considerations 5
- 3.2 Optimal discriminant for primary Lund declusterings 7
- 3.3 Extension to the full clustering tree 9

4 Beyond strong angular ordering: including clustering logarithms **10**
- 4.1 Generic considerations 10
- 4.2 Clustering logarithm with the full Lund tree 11
- 4.3 Clustering logarithms with primary radiation only 14

5 Machine learning approaches **15**
- 5.1 Primary Lund plane and LSTM 15
- 5.2 Full Lund tree and Lund-Net 16

6 Validation in a pure-collinear (toy) parton shower **16**

7 Full Monte Carlo simulations **20**
- 7.1 Setup 20
- 7.2 Tagging performance 21
- 7.3 Resilience 24
- 7.4 Comparison with other approaches 27
- 7.5 Effect of clustering logarithms and of azimuthal angles 30
- 7.6 Asymptotic single-logarithmic limit 31

8 Conclusions **33**

A Sudakov factors with exact splitting functions **35**

1 Introduction

Quarks and gluons, the constituents of the proton, are fundamental entities of essential relevance for physics at the Large Hadron Collider (LHC) at CERN. While these particles are ubiquitous at hadron colliders, they are never observed directly, but rather fragment and hadronise immediately into collimated sprays of colourless hadrons. These decay products

JHEP08(2022)177

are referred to as jets and are generally defined through the application of a sequential recombination algorithm (see e.g. [1, 2]).

As most of the jets observed at collider experiments arise from the fragmentation of a light parton, a detailed understanding of their properties is crucial for experimental analyses. In this context, many experimental studies make use of tools that reliably identify the flavour of jets, e.g. to enhance signals from new physics (decaying predominantly to quarks) from QCD backgrounds (producing predominantly gluon jets). Since quark and gluon branch into one another, it is highly non-trivial to even define what is meant by a "quark jet" or a "gluon jet" (see, for example, the discussion in ref. [3], as well as refs. [4, 5]). As a direct consequence, it is delicate to introduce a properly-defined flavoured jet algorithm [6].

Over the past decade, jet substructure, the study of the internal dynamics of jets, has proven a useful approach to study the decay of heavy particles at and above the electroweak scale, providing a promising avenue to search for signs of new physics beyond the Standard Model (see [2, 7, 8] for recent reviews). While jet substructure has applications in many directions including for example precision measurements in QCD and the study of the quark-gluon plasma produced in heavy-ion collisions, recent years have seen an increasing interest in leveraging progress in deep learning to a range of jet tagging problems [9–18].

Several jet substructure techniques have been introduced to address the question of quark/gluon discrimination. This includes jet-shape based observables like jet angularities [19, 20], energy-energy correlation functions [21] or the jet charge [22–24], counting observables like the charged track multiplicity or the Iterative Soft Drop multiplicity [25], as well as a series of recent deep-learning-based approaches using a range of network architectures and inputs [26–30]. Other techniques, such as jet topics [4, 31, 32], are based on a statistical ensemble of events and are directly meant to obtain separate distributions for quarks and gluons. These are not discussed here as we instead target quark/gluon discriminants working on individual jets.

Recently, the Lund Jet Plane has been introduced [33] as a powerful technique to tackle a wide range of jet substructure applications. For example, the primary Lund plane density has been measured by the ATLAS [34] and ALICE [35] collaborations highlighting, for example, differences between general-purpose Monte-Carlo event generators. This Lund plane density is amenable to precision calculations in perturbative QCD [36], showing an agreement with the ATLAS measurement. Finally, Lund-plane variables can be used as inputs to machine-learning tagger [30, 33].

In this paper, we will use the Lund plane approach to study quark/gluon discrimination. We will do this using both an analytic approach and machine-learning tagging methods. In both cases, we will build two taggers: one based on information from primary Lund declusterings only, and a second based on the full Lund declustering tree. Our analytic approach is based on a resummed calculation of the likelihood ratio at the single-logarithmic accuracy, i.e. matching the logarithmic accuracy obtained in ref. [36] for the primary Lund plane density. We note that likelihood ratios have already been relied upon in the context of boosted-jet discrimination, for example, shower deconstruction [37–39]. Our machine-learning taggers follow the guidelines from refs. [30, 33]. One of the main novelties of this work is that we will aim to gain a first-principles understanding of the

JHEP08(2022)177

behaviour of the neural network by comparing it with our analytic discriminants in specific limits where the analytic approach is known to be optimal. This provides cross-validation of both approaches, shedding light on the importance of subleading effects in the analytic tagger and providing information on the convergence of deep-learning methods.

The paper is organised as follows. In section 2 we describe the Lund plane which will serve as framework for this study. We describe the analytic strongly angular-ordered Lund-plane discriminant in section 3, both for the primary Lund plane and for the full clustering tree. We discuss the inclusion of clustering logarithms in section 4. The baseline machine learning models used in our comparisons are described in section 5, and we provide a validation of the analytic discriminants against these models using a toy shower in section 6. Finally, we perform an in-depth comparison of the performance and resilience of a wide range of methods on full Monte Carlo simulations in section 7, showing that our approaches are either better or on par with state-of-the-art methods both in terms of discriminating power and in terms of resilience.

The code implementing our analytic quark-gluon discriminant based on Lund declusterings is available at https://gitlab.com/gsoyez/lund-quark-gluon.

2 Lund plane(s) and baseline discriminants

In order to fix once and for all the notations to be used throughout this paper, we briefly remind the reader of how the primary Lund plane declusterings are constructed [33]. We also introduce a generalisation beyond the primary plane that instead keeps the full *declustering tree* that we exploit later in this paper.

Primary Lund declusterings. For a given jet, we first recluster its constituents with the Cambridge/Aachen algorithm [40, 41].[1] We then build the list of primary declusterings as follows:

1. start with j being the full reclustered jet;
2. undo the last step of the clustering, $j \to j_1 + j_2$, giving two subjets j_1 and j_2. We assume without loss of generality that j_1 is the "harder branch" i.e. that $p_{t1} > p_{t2}$.
3. We define the set of coordinates $\mathcal{T} = \{\Delta, k_t, z, \psi, \dots\}$ for this branching:

$$\begin{aligned} \Delta &= \sqrt{(y_1 - y_2)^2 + (\phi_1 - \phi_2)^2}, & k_t &= p_{t2}\Delta, \\ z &= \frac{p_{t2}}{p_{t1} + p_{t2}}, & \psi &= \tan^{-1}\frac{y_2 - y_1}{\phi_2 - \phi_1}. \end{aligned} \tag{2.1}$$

4. Iterate by going back to step 2 with $j \leftarrow j_1$.

This produces a tuple, ordered from the first declustering to the last,

$$\mathcal{L}_{\text{primary}} = [\mathcal{T}_1, \dots, \mathcal{T}_i, \dots, \mathcal{T}_n] \tag{2.2}$$

that we refer to as the primary Lund declusterings associated with the jet j.

[1] The reason why we use the Cambridge/Aachen algorithm instead of other algorithms of the generalised-k_t family [42] algorithm, is discussed in section 2.4 of ref. [33].

JHEP08(2022)177

The averaged primary Lund plane density is then simply defined as the average number of declusterings for a given $\ln\Delta$ and $\ln k_t$:

$$\rho(\Delta, k_t) = \frac{1}{N_{\text{jets}}} \frac{d^2N}{d\ln\Delta\, d\ln k_t}. \tag{2.3}$$

This quantity has been measured by the ATLAS collaboration in ref. [34] (see also ref. [43] for a preliminary ALICE measurement) and studied analytically in ref. [36].

Lund declustering tree. Instead of focusing only on following the hardest branch through the declustering, one can retain the full Cambridge/Aachen tree structure yielding an associated tree of Lund variables:

$$\mathcal{L}_{\text{tree}}(j) = [(\mathcal{T}(j), \mathcal{L}_{\text{tree}}(j_1), \mathcal{L}_{\text{tree}}(j_2))] \tag{2.4}$$

where the tree, $\mathcal{L}_{\text{tree}}(j)$, associated with a jet j has a set $\mathcal{T}(j)$ of Lund coordinates associated with the branching $j \to j_1 + j_2$, with $p_{t1} > p_{t2}$, as well as sub-trees $\mathcal{L}_{\text{tree}}(j_1)$ and $\mathcal{L}_{\text{tree}}(j_2)$ associated with j_1 and j_2 respectively. Note that this structure can be flattened into a tuple

$$\mathcal{L}_{\text{tree}}(j) = [(\mathcal{T}_1, i_{\text{hard},1}, i_{\text{soft},1}), \ldots, (\mathcal{T}_i, i_{\text{hard},i}, i_{\text{soft},i}), \ldots, (\mathcal{T}_n, i_{\text{hard},n}, i_{\text{soft},n})], \tag{2.5}$$

where $i_{\text{hard},i}$ (resp. $i_{\text{soft},i}$) indicate the index in the tuple for the next branching along the hard (resp. soft) branch, or 0 in the lack thereof.

Iterated Soft Drop multiplicity. The iterated Soft Drop multiplicity [25] can be straightforwardly defined from the tuple of primary Lund declusterings as the number of declusterings satisfying a given Soft Drop [44, 45] condition $z_i > z_{\text{cut}}(\Delta_i/R)^\beta$ with R the jet radius and z_{cut} and β the Soft Drop parameters, with $\beta < 0$. A standard choice is to take $\beta = -1$ so as to effectively use Soft Drop to impose a cut on $z\Delta$ which is similar to a k_t scale. In this paper, we define the iterated Soft Drop multiplicity as the number of primary declusterings above a fixed k_t cut. This is motivated by the fact that a dimensionful k_t cut is more adequate than a cut on $z\Delta$ to separate between a perturbative region ($k_t > k_{t,\text{cut}}$) and a non-perturbative region ($k_t < k_{t,\text{cut}}$).[2] The use of a k_t cut will also be used with all the other methods introduced in this paper, allowing for a direct comparison of their performance.

Analytically, one can show that the Iterated Soft Drop multiplicity is the optimal quark-gluon discriminant in the double-logarithmic approximation (see below for a proof).

Baseline discriminant: the average Lund-plane density. The baseline approach we will consider throughout this paper is the one that was introduced in ref. [33]. We first compute the average primary Lund plane densities $\rho_{q,g}(\Delta, k_t)$ separately for the quarks and gluons samples, respectively. For a given jet with Lund declusterings $\{(\Delta_i, k_{t,i})\}$, we then define a likelihood ratio

$$\mathbb{L}_{\text{density}} = \prod_i \frac{\rho_g(\Delta_i, k_{t,i})}{\rho_q(\Delta_i, k_{t,i})}. \tag{2.6}$$

In practice, the average densities $\rho_{q,g}(\Delta, k_t)$ are computed in bins of $\ln\Delta$ and $\ln k_t$.

[2] The fundamental physics motivation behind this choice is that, in a resummed calculation in perturbative QCD, the scale entering the strong coupling is typically the k_t of the emission.

JHEP08(2022)177

This approach has already shown to give good results when applied to discriminating boosted W bosons from QCD jets. The application to quark-gluon discrimination considered in this paper is actually simpler as it does not involve a separate isolation and treatment of a hard two-prong decay as was the case in [33] for boosted W tagging.

Eq. (2.6) would be an optimal discriminant if the Lund plane declusterings were independent. We know that this is not the case in practice due to effects such as the energy lost by the leading parton, flavour changes and clustering effects, as discussed in [36]. However, we still expect this approach to yield a better performance than the Iterated Soft Drop multiplicity as it captures some aspects of the quark and gluon radiation patterns beyond the soft-collinear approximation. Throughout this paper, we will consider quark-gluon tagging using the average Lund plane density as a reference approach from which we want to build more performant discriminants.

3 Analytic discriminants in the Lund-plane with strong angular ordering

3.1 Generic considerations

In this section, we introduce a series of quark-gluon discriminants based on a first-principles treatment of the Lund plane declusterings in perturbative QCD. The performance of these new tools will be assessed later in sections 6 and 7.

The core idea is to explicitly compute the likelihood ratio

$$\mathbb{L} = \frac{p_g(\mathcal{L})}{p_q(\mathcal{L})}\,, \tag{3.1}$$

for a set of Lund declusterings $\mathcal{L}$ — either primary-only or including the full tree — where $p_{q,g}(\mathcal{L})$ denotes the probability to observe the given set of Lund declusterings assuming the jet is either initiated by a quark or by a gluon.[3] For this procedure to be infrared-and-collinear safe, we only consider emissions above a given (relative) transverse momentum cut, i.e. require $k_t \geq k_{t,\text{cut}}$.

In the (double-logarithmic) soft-collinear limit, emissions are independent and the single-emission probability for quarks and gluons only differ by the overall colour factor (C_F for a quark, C_A for a gluon). For n primary emissions $\mathcal{T}_i$, one therefore has

$$\mathbb{L}_{LL} = \prod_i \frac{p_g(\mathcal{T}_i)}{p_q(\mathcal{T}_i)} = \left(\frac{C_A}{C_F}\right)^n. \tag{3.2}$$

This shows that the likelihood ratio is only a (monotonic) function of n and hence that the iterated Soft Drop multiplicity is the optimal discriminant at leading (double) logarithmic accuracy. In this limit, additional, non-primary, declusterings in the full Lund tree all come with a factor C_A and therefore do not contribute to eq. (3.2).

In what follows, we want to extend this result to single-logarithmic accuracy, as what was done in [36] for the average primary Lund plane density. For this, several single-logarithmic effects have to be taken into account: (i) corrections to the running of the

[3] Throughout this paper, we do not distinguish between quarks and anti-quarks.

JHEP08(2022)177

strong coupling, (ii) collinear effects stemming either from splittings where the flavour of the leading parton changes or from finite z splittings, (iii) clustering effects where the exact Cambridge/Aachen clustering has to be taken into account for multiple soft emissions at commensurate angles, and (iv) effects of multiple soft emissions at large angles.

In the (soft-collinear) limit where emissions are independent, running coupling effects — (i) in the above list — do not change the double-logarithmic result in eq. (3.2) as both the quark and gluon probabilities are multiplied by α_s taken at the same scale, namely the (relative) k_t of the emission. Running-coupling corrections will nevertheless appear together with the other single-logarithmic contributions and this is discussed below.

Next, the effect of soft-wide-angle radiation — (iv) in the above list — depends on the details of the hard process that underlines the samples of quark and gluon jets. These contributions would, for example, be different in "quark jets" in $qg \to Zq$ and $qq \to qq$ events (see e.g. the discussion in refs. [3, 46]). In this paper, we will focus on universal aspects in the collinear limit and therefore neglect these contributions which scale like the square of the jet radius. (For phenomenological applications in section 7, we will therefore use $R = 0.4$.) Beyond the small-radius limit, soft-wide-angle effects would have to be included. In this case, one should consider the approach where one has to discriminate a specific "quark-enriched" signal process from a specific "gluon-enriched" background process. This study goes beyond the scope of this paper. It would nevertheless be interesting, in a follow-up study, to investigate if the analytic techniques developed in this paper could be used to assess the process-dependence of quark-gluon tagging, potentially in combination with the concept of jet topics [31].

In a similar spirit, fixed-order (e.g. NLO) corrections to the underlying hard process would also be process-specific and, as such, fall beyond the scope of this paper. That said, if we were to address a process-specific tagging using our analytic techniques, NLO effects would become relevant in two cases. Firstly, in the context of our resummed approach, the would start contributing one order beyond our single-logarithmic approach. Secondly, exact fixed-order corrections would impact the region of large quark and gluon tagging efficiencies (see e.g. the discussion in ref. [5]).

Clustering effects are delicate to handle in an analytic calculation as even in the large-N_c limit they, in principle, require the full matrix angular dependence for an arbitrary number of emissions strongly-ordered in energy. Since we can expect that collinear effects, and flavour-changing contributions in particular, are numerically dominant in the context of quark-gluon discrimination, we will as a first step neglect clustering effects. In other words, in this section we work in the regime where emissions are strongly ordered in angle and derive a quark-gluon discriminant either using only primary emissions, section 3.2, or using the full clustering tree, section 3.3.

We come back to the question of clustering logarithms in section 4. We will see explicitly in our Monte Carlo simulations in section 7 that clustering effects have a smaller numerical impact on quark-gluon discrimination than the collinear enhancements.

JHEP08(2022)177

3.2 Optimal discriminant for primary Lund declusterings

We start by considering only the primary Lund plane declusterings $\{(\Delta_i, k_{t,i})\}$ with $k_t \geq k_{t,\text{cut}}$. For these, we want to compute the likelihood ratio

$$\mathbb{L}_{\text{primary}} = \frac{p_g(\{\Delta_i, k_{t,i}, z_i, \dots\})}{p_q(\{\Delta_i, k_{t,i}, z_i, \dots\})}, \tag{3.3}$$

at single logarithmic accuracy in perturbative QCD, in the limit where the emissions are strongly ordered in angle, i.e. that $\Delta_1 \gg \Delta_2 \gg \cdots \gg \Delta_n$. In this limit, we should include in eq. (3.3) the contributions associated either with the running of the strong QCD coupling effects, or with any hard-collinear effect.

The quark and gluon probability distributions can be computed iteratively starting from the first (largest-angle) splitting. A key point to take into account is the fact that collinear branchings can change the flavour of the leading branch, either through a $g \to q\bar{q}$ splitting, or through a $q \to qg$ splitting where the emitted gluon is harder than the final quark. At each splitting, we should therefore keep track of the flavour of the leading parton as well as of its splitting fraction z and its relative transverse momentum k_t. It is convenient to introduce a matrix

$$p_{ab}^{(i)} \equiv \begin{pmatrix} p(q_i|q_0) & p(q_i|g_0) \\ p(g_i|q_0) & p(g_i|g_0) \end{pmatrix}, \tag{3.4}$$

where $p(b_i|a_0)$ denotes the probability that the harder branch has flavour b after the i^{th} declustering, given that it started (at step "0") with a jet of flavour a. This matrix is initialised as $p_{ab}^{(0)} = \delta_{ab}$ and is recursively constructed from step $i-1$ to step i for each of the $i = 1, \dots, n$ Lund declusterings.

Assuming that just before branching i the jet has flavour a, the probability after branching i should include two effects: (i) the probability that the splitting has the observed kinematic properties $\Delta_i, z_i, \dots$, potentially including a change of the leading flavour, and (ii) a Sudakov factor implementing the fact that no emission has occurred between the previous angle Δ_{i-1} and Δ_i (with $\Delta_0 \equiv R$), and with $k_t > k_{t,\text{cut}}$. This Sudakov resums the virtual corrections between Δ_{i-1} and Δ_i. This leads to the recursion

$$p_{ab}^{(i)} = \frac{\alpha_s(k_{ti})}{\pi \Delta_i} \begin{pmatrix} \tilde{P}_{qq}(z_i) & \tilde{P}_{qg}(z_i) \\ \tilde{P}_{gq}(z_i) & \tilde{P}_{gg}(z_i) \end{pmatrix} \begin{pmatrix} S_q^{(i-1,1)} & 0 \\ 0 & S_g^{(i-1,1)} \end{pmatrix} p_{ab}^{(i-1)}. \tag{3.5}$$

In this expression, the splitting kernels $\tilde{P}_{ab}$ are directly related to the Altarelli-Parisi splitting functions with the extra requirement that since the declustering procedure follows the

JHEP08(2022)177

hardest branch one should impose $z_i < \frac{1}{2}$:[4]

$$\tilde{P}_{qq} = P_{gq}(z)\Theta\left(z<\frac{1}{2}\right) = C_F\frac{1+(1-z)^2}{z}\Theta\left(z<\frac{1}{2}\right), \tag{3.6a}$$

$$\tilde{P}_{gq} = P_{qq}(z)\Theta\left(z<\frac{1}{2}\right) = C_F\frac{1+z^2}{1-z}\Theta\left(z<\frac{1}{2}\right), \tag{3.6b}$$

$$\tilde{P}_{qg} = [P_{qg}(z)+P_{qg}(1-z)]\Theta\left(z<\frac{1}{2}\right) = 2n_f T_R[z^2+(1-z)^2]\Theta\left(z<\frac{1}{2}\right), \tag{3.6c}$$

$$\tilde{P}_{gg} = [P_{gg}(z)+P_{gg}(1-z)]\Theta\left(z<\frac{1}{2}\right) = 2C_A\left[\frac{1-z}{z}+\frac{z}{1-z}+z(1-z)\right]\Theta\left(z<\frac{1}{2}\right). \tag{3.6d}$$

The Sudakov factors, $S_{q,g}^{(i-1,i)}$, between the angle of the last splitting Δ_{i-1} and the angle of the current splitting Δ_i is computed as

$$S_f^{(i-1,i)} = \exp\left[-\int_{\Delta_i}^{\Delta_{i-1}}\frac{d\Delta}{\Delta}\int dz\frac{\alpha_s(p_{ti}z\Delta)}{\pi}P_f(z_i)\Theta(p_{ti}z\Delta > k_{t,\text{cut}})\right], \tag{3.7}$$

with P_f the total splitting function for a parton of flavour f and p_{ti} the transverse momentum (with respect to the beam) of parton i before splitting. The k_t of the emission is taken as $p_{ti}z\Delta$ which is equivalent to our definition in eq. (2.1) in the collinear limit.[5] This Sudakov is evaluated at next-to-leading logarithmic (NLL) accuracy with $\Delta_i \ll \Delta_{i-1}$, and we find

$$\begin{aligned} S_f^{(i-1,i)} = \exp\Bigg\{ &-\frac{C_f}{2\pi\alpha_s\beta_0^2}\Bigg[(1-\lambda_{i-1})\ln\frac{1-\lambda_{i-1}}{1-\lambda_{\text{cut}}} - (1-\lambda_i)\ln\frac{1-\lambda_i}{1-\lambda_{\text{cut}}} - \lambda_i + \lambda_{i-1} \\ &- \frac{\alpha_s\beta_1}{\beta_0}\left(\frac{1}{2}\ln^2(1-\lambda_i) - \frac{1}{2}\ln^2(1-\lambda_{i-1}) + \frac{\lambda_i-\lambda_{i-1}}{1-\lambda_{\text{cut}}}\ln(1-\lambda_{\text{cut}})\right) \\ &+ \left(\frac{\alpha_s K}{2\pi} - \frac{\alpha_s\beta_1}{\beta_0}\right)\left(\frac{\lambda_i-\lambda_{i-1}}{1-\lambda_{\text{cut}}} - \ln\frac{1-\lambda_{i-1}}{1-\lambda_i}\right)\Bigg]\Bigg\}, \end{aligned} \tag{3.8}$$

with $\alpha_s \equiv \alpha_s(p_{t,\text{jet}}R)$,

$$\lambda_{i-1} = 2\alpha_s\beta_0\left(\ln\frac{R}{x\Delta_{i-1}} - B_f\right), \qquad \beta_0 = \frac{11C_A - 4n_f T_R}{12\pi}, \tag{3.9a}$$

$$\lambda_i = 2\alpha_s\beta_0\left(\ln\frac{R}{x\Delta_i} - B_f\right), \qquad \beta_1 = \frac{17C_A^2 - 5C_A n_f - 3C_F n_f}{24\pi^2}, \tag{3.9b}$$

$$\lambda_{\text{cut}} = 2\alpha_s\beta_0\ln\frac{p_{t,\text{jet}}R}{k_{t,\text{cut}}}, \qquad K = \left(\frac{67}{18} - \frac{\pi^2}{6}\right)C_A - \frac{5}{9}n_f, \tag{3.9c}$$

$$B_q = -\frac{3}{4} \qquad B_g = -\frac{11C_A - 4n_f T_R}{12C_A}, \tag{3.9d}$$

[4] We have chosen notations where the indices of the $\tilde{P}$ kernels represent the flavour of the hard branch, so as to make the matrix product in eq. (3.5) more obvious. As a consequence, these indices do not always match with the standard indices in the Altarelli-Parisi kernels where the indices instead refer to the flavour of the emitted parton with momentum fraction z. Finally, our probability distributions are taken differentially in Δ_i and z_i. The specific choice of variables is however irrelevant for the problem of quark-gluon classification as it cancels in the likelihood ratio.

[5] Conversely, the value of p_{ti} can be deduced from Δ_i, k_{ti} and z_i using $p_{ti} = k_{ti}/(z_i\Delta_i)$.

JHEP08(2022)177

and x defined as the momentum fraction of the total jet momentum carried by the subjet j before branching i. We note that in the rare occurrences where the Lund declusterings are not ordered in angle — which cannot be ruled out with the Cambridge/Aachen declustering procedure — we set the Sudakov to $S_f = 1$. We also point out that the contribution from hard-collinear splitting to the above expressions have been computed by setting an upper bound e^{B_f} on the z integration in (3.7). This is correct at NLL accuracy. Although it has the drawback to insert uncontrolled subleading corrections — compared to the traditional expression which can be recovered by keeping only the first non-trivial term in B_f — it has the advantage of having a clean endpoint, i.e. $S_f = 1$ for $\lambda_{i-1} \leq \lambda_{\text{cut}}$.

If we introduce the short-hand notations $\tilde{P}_{ab}^{(i)}$ and $S_{ab}^{(i,i-1)} \equiv \delta_{ab} S_a^{(i,i-1)}$ for the full splitting matrix and Sudakov matrix, the probabilities after including all the Lund declusterings takes the form

$$p^{(\text{final})} = S^{(n+1,n)} \tilde{P}^{(n)} S^{(n,n-1)} \dots \tilde{P}^{(i)} S^{(i,i-1)} \dots \tilde{P}^{(1)} S^{(1,0)} p^{(0)}, \tag{3.10}$$

where the leftmost factor in the right-hand side takes into account the fact that there are no more emissions between the angle of the last declustering, Δ_n, and the smallest angle accessible after the last splitting: $\Delta_{n+1} \equiv \Delta_{\text{min}} = k_{t,\text{cut}}/p_{tn}$ with p_{tn} the transverse momentum (with respect to the beam) of the leading parton after the last declustering. Eq. (3.10) has the simple physical interpretation of successive primary branchings, producing the factors $\tilde{P}^{(i)}$, interleaved with Sudakov factors, $S^{(i,i-1)}$, which resum virtual corrections between two primary emissions. Finally, the probabilities associated with an initial quark or gluon jet are given by

$$p_q(\{\Delta_i, k_{t,i}, z_i, \dots\}) = p^{(\text{final})}(q|q_0) + p^{(\text{final})}(g|q_0), \tag{3.11a}$$

$$p_g(\{\Delta_i, k_{t,i}, z_i, \dots\}) = p^{(\text{final})}(q|g_0) + p^{(\text{final})}(g|g_0), \tag{3.11b}$$

translating the fact that we are inclusive over all flavours of the final leading parton.

The probabilities in eqs. (3.11) can be directly inserted in (3.3) to obtain a quark-gluon discriminant. It is, by construction, the optimal discriminant at single-logarithmic accuracy in the limit where the declusterings are strongly ordered in angle. Since the above procedure keeps track of the flavour and momentum fraction x of the leading parton at each step, it takes into account the possible correlations between the different declusterings, hence going beyond the independent-emission assumption used with the average Lund plane density (section 2, eq. (2.6)).

3.3 Extension to the full clustering tree

While non-primary (secondary, tertiary, ...) declusterings have no impact at leading-logarithmic accuracy, they start carrying information at our single-logarithmic accuracy. Generalising the approach from the previous section to the full clustering tree is mostly a technical step. This time, we therefore settle to compute

$$\mathbb{L}_{\text{tree}} = \frac{p_g(\mathcal{L}_{\text{tree}})}{p_q(\mathcal{L}_{\text{tree}})} \tag{3.12}$$

in the strongly angular-ordered limit.

JHEP08(2022)177

This is again done recursively over the full (de-)clustering tree. For this, consider a declustering $j_{\text{parent}} \to j_{\text{hard}} + j_{\text{soft}}$, with kinematic variables $\mathcal{T}$, i.e. with an angle Δ, a soft momentum fraction z and relative transverse momentum k_t. The probabilities associated with the parent jet can be deduced from those of the subjets as follows:

$$p_q(\mathcal{L}_{\text{parent}}) = S_q(\Delta_{\text{prev}}, \Delta) \left[\tilde{P}_{qq}(z) p_q(\mathcal{L}_{\text{hard}}) p_g(\mathcal{L}_{\text{soft}}) + \tilde{P}_{gq}(z) p_g(\mathcal{L}_{\text{hard}}) p_q(\mathcal{L}_{\text{soft}})\right] \quad (3.13a)$$

$$p_g(\mathcal{L}_{\text{parent}}) = S_g(\Delta_{\text{prev}}, \Delta) \left[\tilde{P}_{gg}(z) p_g(\mathcal{L}_{\text{hard}}) p_g(\mathcal{L}_{\text{soft}}) + \tilde{P}_{qg}(z) p_q(\mathcal{L}_{\text{hard}}) p_q(\mathcal{L}_{\text{soft}})\right] \quad (3.13b)$$

where Δ_{prev} is the angle at which the last declustering before the one under consideration happened (with $\Delta_{\text{prev}} = R$ for the first declustering). As in section 3.2, $S(\Delta_{\text{prev}}, \Delta)$ is a Sudakov factor imposing that no other emission with $k_t > k_{t,\text{cut}}$ occurred since the last declustering at an angle Δ_{prev}, cf. eq. (3.8).[6] The splitting kernels $\tilde{P}_{ab}$ are the same as in eq. (3.6). These expressions have the same form as eq. (3.5) except that, at each step, they also include the probability for the soft branch.

This recursion is applied until each branch can no longer be declustered in which case, if the last splitting has occurred at an angle Δ_{last}, one then just includes a factor

$$p_q(\mathcal{L} = \emptyset) = S_q(\Delta_{\text{last}}, \Delta_{\text{min}}), \quad (3.14a)$$

$$p_g(\mathcal{L} = \emptyset) = S_g(\Delta_{\text{last}}, \Delta_{\text{min}}), \quad (3.14b)$$

where, as for the primary case, $\Delta_{\text{min}} = k_{t,\text{cut}}/p_t$ for a final branch of momentum p_t.

4 Beyond strong angular ordering: including clustering logarithms

4.1 Generic considerations

We conclude this section on analytic methods by discussing the inclusion of clustering logarithms in our approach. These logarithms arise from situations where we have at least two emissions with commensurate angles and the exact Cambridge/Aachen clustering has to be considered in order to label the emissions as primary, secondary, ternary, etc. When the emissions at commensurate angles are strongly ordered in energy, this leads to single-logarithmic contributions (see e.g. [36]).

In practice, the Cambridge/Aachen clustering can produce clusterings which are not in agreement with the naive physical expectation. Consider for example a quark-initiated jet with two gluon emissions. The harder emission is emitted from the quark and comes with a colour factor C_F. The softer emission can either be seen as emitted from the quark, with a colour factor C_F, or as emitted from the gluon, with a colour factor C_A. When the two emission angles are similar, the actual Cambridge/Aachen clustering will sometimes cluster the second gluon in the C_F^2 contribution with the first gluon, yielding a secondary Lund declustering, or, conversely, cluster the second gluon from the $C_F C_A$ contribution with the hard quark, yielding a primary declustering.

[6] Although we have only made explicit the angular dependence, the Sudakov factors also depend on the prong momenta.

JHEP08(2022)177

In order to compute these contributions, we need the full angular structure of the matrix elements for an arbitrary number of emissions at commensurate angles. When computing the average Lund plane density, this can be addressed, at least in the large-N_c limit, by a Monte Carlo integration similar to the one used to resum non-global logarithms in [47]. (See [48] for an approach valid at finite N_c.)

In our quark-gluon-discrimination application, one would have to keep track of all the possible colour configuration with which an emission can be radiated by the full set of emissions at larger transverse momentum (or, at our accuracy, at larger energies). This is beyond what can be practically achieved. Instead, we will adopt a simplified approach where we apply a matrix-element correction which only describes correctly situations where (any number of) pairs of emissions are at commensurate angles. This is similar in spirit to the NODS scheme introduced in [49] to implement subleading-N_c corrections in parton showers.[7]

4.2 Clustering logarithm with the full Lund tree

Since clustering logarithms have an explicit dependence on radiation in/from different leaves, we first consider the situation where the quark-gluon tagging is done using the full Lund declustering tree. The case where only primary radiation is considered will be discussed in section 4.3 below.

In our approximation where we only allow for two emissions at commensurate angles, we then consider two declusterings $\mathcal{T}_1 \equiv \{\Delta_1, k_{t1}, z_1, \psi_1\}$ and $\mathcal{T}_2 \equiv \{\Delta_2, k_{t2}, z_2, \psi_2\}$, with $\Delta_1 \sim \Delta_2 \ll 1$. Since clustering corrections happen for two emissions at similar angles and we only aim at describing the configurations where we have only pairs of particles at commensurate angles, we can assume that $\mathcal{T}_1$ and $\mathcal{T}_2$ correspond to consecutive Lund declusterings and that all the other emissions are at widely different angles. We can further assume that $\mathcal{T}_1$ happens before $\mathcal{T}_2$ in the sequence of declusterings, i.e. $\Delta_1 > \Delta_2$. Our approach is to modify the emission probability in eq. (3.12) for $\mathcal{T}_2$ to include corrections due to the presence of $\mathcal{T}_1$.

There are two main kinematic configurations to consider: either $\mathcal{T}_1$ and $\mathcal{T}_2$ are both reconstructed as consecutive "primary" emissions from the same hard branch, or $\mathcal{T}_2$ is reconstructed as a "secondary" emission from $\mathcal{T}_1$.[8] In the "primary" case, we can either have $z_2 \ll z_1$ ($k_{t2} \ll k_{t1}$) or $z_2 \gg z_1$ ($k_{t2} \gg k_{t1}$), while in the "secondary" case we can assume $z_2 \ll z_1$ ($k_{t2} \ll k_{t1}$). This is illustrated by the Lund diagrams in figure 1. At single-logarithmic accuracy, the clustering correction is computed in the flavour channel where both emissions are gluons. The distinction between the primary and secondary cases is decided by the Cambridge/Aachen clustering. In both cases, if C_R is the colour factor of the common hard branch, the matrix element corresponding to a given clustering

[7] Note the key difference that the NODS method produces the correct behaviour at large-N_c for any number of emissions at commensurate angles. The matrix-element correction only applies to subleading-N_c corrections. In our case, the correct behaviour is only guaranteed for pairs of emissions at commensurate angles even in the large-N_c limit.

[8] Primary and secondary are here counted from the hard branch common to both emissions, even if this one can be anywhere in the Lund tree.

JHEP08(2022)177

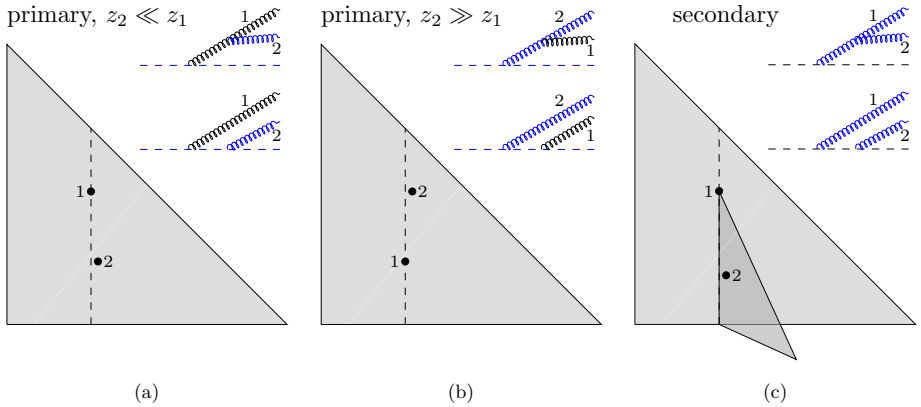


Figure 1. Schematic representation of the three kinematic configurations affected by clustering logarithms. We consider two consecutive declusterings $\mathcal{T}_1$ and $\mathcal{T}_2$ with $\mathcal{T}_2$ occurring after $\mathcal{T}_1$ in the Lund sequence. If $\mathcal{T}_2$ is reconstructed as "primary" (i.e. in the same plane as $\mathcal{T}_1$) it can either be much softer (case (a)) or much harder (case (b)) than $\mathcal{T}_1$. Case (c) describes the situation where $\mathcal{T}_2$ is reconstructed as a secondary emission from $\mathcal{T}_1$, and hence can be considered much softer.

sequence will have a contribution proportional to C_R^2 and one proportional to $C_R C_A$. In the strongly-angular-ordered limit, only the first term (C_R^2) contributes to the "primary" clustering and only the second term ($C_R C_A$) to the "secondary" clustering. The gluon-emission diagrams shown in figure 1 represent the two contributions for each clustering case. The first particles to cluster are highlighted in blue.

Let us first handle the case where $\mathcal{T}_1$ and $\mathcal{T}_2$ are both "primary" emissions. Say the parent parton has a colour factor C_R. If $z_1 \gg z_2$, the $C_R\,\mathrm{d}^2\Delta_2/\Delta_2^2$ behaviour which corresponds to the collinear limit in section 3.3 should be replaced by the full soft-gluon radiation squared matrix element

$$\left[\frac{C_A}{2}\frac{1}{\Delta_{12}^2} + \frac{C_A}{2}\frac{\Delta_1^2}{\Delta_{12}^2\Delta_2^2} + \left(C_R - \frac{C_A}{2}\right)\frac{1}{\Delta_2^2}\right] d^2\Delta_2, \tag{4.1}$$

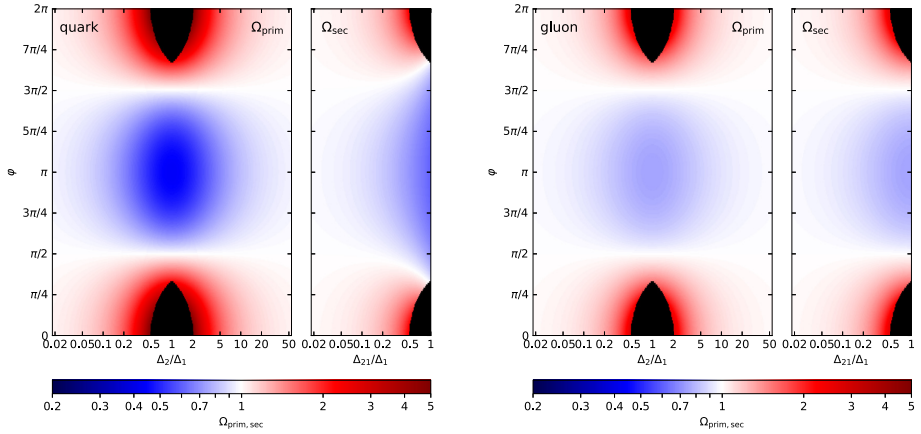
where $\Delta_{12}^2 = \Delta_1^2 + \Delta_2^2 - 2\Delta_1\Delta_2\cos(\psi_2 - \psi_1)$ is the angle between the two emitted gluons. This means that, in the gluon emission part of eq. (3.5), we should apply a correction factor

$$\Omega_{\text{prim}} = 1 + \frac{C_A}{2C_R}\left(\frac{\Delta_2^2}{\Delta_{12}^2} + \frac{\Delta_1^2}{\Delta_{12}^2} - 1\right). \tag{4.2}$$

It is straightforward to show that the "primary" case with $z_1 \ll z_2$, gives the same correction Ω_{prim}. As expected, $\Omega_{\text{prim}} \to 1$ when $\Delta_1 \gg \Delta_2$ (or when $\Delta_1 \ll \Delta_2$) so that the strongly-ordered limit is recovered. Since both emissions are primary, we never have $\Delta_{12} \ll \Delta_1, \Delta_2$.

We now turn to the "secondary" case where $\mathcal{T}_2$ is emitted from the soft branch of $\mathcal{T}_1$. Here, Δ_2 is the angle between the emissions $\mathcal{T}_1$ and $\mathcal{T}_2$, i.e. $\Delta_2 \equiv \Delta_{12}$, and we denote by

JHEP08(2022)177



(a) (b)

Figure 2. Clustering correction factors Ω_{prim} (primary) and Ω_{sec} (secondary) for quark (left) and gluon (right) leading partons. The black regions correspond to the kinematic regions where emissions are not clustered in the indicated Lund plane.

$\Delta_{02} = \Delta_1^2 + \Delta_2^2 - 2\Delta_1\Delta_2\cos(\psi_2 - \psi_1)$ the angle between $\mathcal{T}_2$ and the hard branch of $\mathcal{T}_1$. In the collinear limit this would correspond to a factor C_A/Δ_2^2 which has to be replaced by the full angular structure

$$\frac{C_A}{2}\frac{1}{\Delta_2^2} + \frac{C_A}{2}\frac{\Delta_1^2}{\Delta_{02}^2\Delta_2^2} + \left(C_R - \frac{C_A}{2}\right)\frac{1}{\Delta_{02}^2}, \tag{4.3}$$

yielding a correction factor

$$\Omega_{\text{sec}} = \frac{1}{2}\left(1 + \frac{\Delta_1^2}{\Delta_{02}^2} + \frac{\Delta_2^2}{\Delta_{02}^2}\right) + \frac{C_R}{C_A}\frac{\Delta_2^2}{\Delta_{02}^2}. \tag{4.4}$$

Without surprise, $\Omega_{\text{sec}} \to 1$ if $\Delta_2 \ll \Delta_1$, recovering the strongly-ordered case. The correction factors Ω_{prim} and Ω_{sec} are plotted in figure 2. We see that they are indeed localised around the region where both emissions have commensurate angles. They tend to be larger for quarks than for gluons.

It is interesting to notice that the above correction which accounts for clustering logarithms introduces a dependence on the azimuthal angle ψ. It is the only dependence on ψ at the single logarithmic accuracy.

In principle, the Sudakov factors should also receive single-logarithmic corrections due to clustering effects. Since clustering logarithms only affect flavour-diagonal emissions of two soft gluons (at least within our approximations), it is however relatively straightforward to convince oneself that these corrections only lead to a reshuffling of some contributions between different factors in the overall probability distribution and can therefore be neglected. This can be understood as follows. Suppose one has a parton emitting a gluon "1" at an

JHEP08(2022)177

angle Δ_1. In our approach, that the previous emission occurred at an angle Δ_0 and that the following emissions on the hard and soft branches occur at angles $\Delta_{2,\text{hard}}$ and $\Delta_{2,\text{soft}}$ respectively. In the strongly-ordered limit, emission "1" is involved in three contributions to the Sudakov: one between Δ_1 and Δ_0, one between $\Delta_{2,\text{hard}}$ and Δ_1, and one between $\Delta_{2,\text{soft}}$ and Δ_1. Within our approximation, if we want to compute the corrections to the Sudakov at angles commensurate with Δ_1, we can assume that all the other emissions are at widely separate angles, i.e. $\Delta_{2,\text{hard}}, \Delta_{2,\text{soft}} \ll \Delta_1 \ll \Delta_0$. To compute the overall Sudakov factor, summed over the three regions described above, one should integrate the exact matrix element, including the full angular structure. Within our approximation where we only target correctness for two emissions at commensurate angles, this integral is proportional to $C_R \log(\Delta_0/\Delta_{2,\text{hard}}) + C_A \log(\Delta_1/\Delta_{2,\text{soft}})$, which is the same as the strongly-ordered limit.

In practice, inserting corrections factors due to clustering logarithms in eq. (3.13) requires some care as it depends whether the Lund declustering, $\mathcal{T}$, that it implements comes from following the hard or the soft branch at the previous declustering. First, at our single-logarithmic accuracy, clustering corrections are only non-trivial for two successive gluon emissions. Then, say that the previous branching, happening at an angle Δ_{prev} is denoted by $\mathcal{T}_{\text{prev}}$. If $\mathcal{T}$ follows $\mathcal{T}_{\text{prev}}$ along the harder branch, we only apply a correction for the contributions where $\mathcal{T}_{\text{prev}}$ did not have a flavour change. The correction is then applied only for the flavour-diagonal contribution with a colour factor C_R being C_F or C_A depending on the flavour of the hard parton. Conversely, if $\mathcal{T}$ follows $\mathcal{T}_{\text{prev}}$ along the softer branch, only the flavour-diagonal term in $p_g(\mathcal{L}_{\text{parent}})$ receives a correction with a colour factor C_R given by the flavour of the hard branch at the branching $\mathcal{T}_{\text{prev}}$.

It is interesting to note that the correction factors Ω_{prim} and Ω_{sec} explicitly depend on the azimuthal angles of the declusterings, which is new compared to the strongly-angular-ordered case. If we want to consider only the Δ, z and k_t variables for each Lund declusterings (see the discussion in section 7 below), we can integrate out the ϕ dependence, averaging over the domain allowed by the fact that the declustering $\mathcal{T}_1$ is undone before $\mathcal{T}_2$. We denote these azimuthally-averaged correction factors by $\bar{\Omega}_{\text{prim}}$ and $\bar{\Omega}_{\text{sec}}$. They only depend on the ratio $x = \Delta_2/\Delta_1$ and are found to be

$$\bar{\Omega}_{\text{prim}} \overset{x<1/2}{=} 1 + \frac{C_A}{C_R}\frac{x^2}{1-x^2} \tag{4.5a}$$

$$\overset{x>1/2}{=} 1 + \frac{C_A}{2C_R}\left[\frac{1+x^2}{|1-x^2|}\frac{1-\frac{2}{\pi}\tan^{-1}\left(\frac{1+x}{|1-x|}\sqrt{\frac{2x-1}{2x+1}}\right)}{1-\frac{2}{\pi}\tan^{-1}\left(\sqrt{\frac{2x-1}{2x+1}}\right)}\right] \tag{4.5b}$$

$$\bar{\Omega}_{\text{sec}} \overset{x<1/2}{=} 1 + \frac{C_R}{C_A}\frac{x^2}{1-x^2} \tag{4.5c}$$

$$\overset{x>1/2}{=} \frac{1}{2} + \left[\frac{\text{sgn}(1-x^2)}{2} + \frac{C_R}{C_A}\frac{x^2}{|1-x^2|}\right]\frac{1-\frac{2}{\pi}\tan^{-1}\left(\frac{1+x}{|1-x|}\sqrt{\frac{2x-1}{2x+1}}\right)}{1-\frac{2}{\pi}\tan^{-1}\left(\sqrt{\frac{2x-1}{2x+1}}\right)}. \tag{4.5d}$$

4.3 Clustering logarithms with primary radiation only

Our last analytic step is to include the effect of clustering logarithms in the Lund quark-gluon discriminant which only uses primary declusterings. We do this in an approximation where we only allow for pairs of emissions to be at commensurate angles.

JHEP08(2022)177

As for the full Lund tree, two types of corrections should be included: corrections to the matrix element for the radiation of two soft gluons at commensurate angles, and potential corrections to the Sudakov factor. Corrections to the real radiation are trivial: the splitting factors $\tilde{P}_{qq}(z_i)$ and $\tilde{P}_{gg}(z_i)$ in eq. (3.5) should be multiplied by a factor Ω_{prim} (eq. (4.2)), respectively with $C_R = C_F$ and $C_R = C_A$.

Since we are no longer including a Sudakov factor for the soft branch, our previous argument saying that the overall Sudakov factor was not affected by clustering effects no longer holds. Let us therefore again consider a gluon emission "1" at an angle Δ_1 and relative transverse momentum k_{t1}. In the soft-gluon limit, the total Sudakov factor at (relative) transverse momentum scales smaller than k_{t1} should use the full matrix element for radiation from the system including both the parent parton and emission "1", i.e.

$$-\log S_f^{(i-1,i)} = \int_{k_{t,\text{cut}}}^{k_{t1}} \frac{\alpha_s(k_t)}{\pi^2} \int d^2\Delta_2 \left[\left(C_R - \frac{C_A}{2} \right) \frac{1}{\Delta_2^2} + \frac{C_A}{2} \frac{1}{\Delta_{12}^2} + \frac{C_A}{2} \frac{\Delta_1^2}{\Delta_2^2 \Delta_{12}^2} \right] \times \left[1 - \Theta(\Delta_{12} < \Delta_1)\Theta(\Delta_{12} > \Delta_2) \right], \tag{4.6}$$

where the square bracket in the second line imposes that "2" is clustered as a primary emission. This gives a correction compared to the strongly-angular-ordered case which is found to be[9]

$$-\delta \log S_f^{(i-1,i)} = (C_A - C_R)\xi \left[\int_{k_{t,\text{cut}}}^{k_{t1}} \frac{\alpha_s(k_t)}{\pi} \right], \tag{4.7}$$

with $\xi = 0.323006$. Note that this contribution happens to vanish when the parent parton is a gluon ($C_R = C_A$).

5 Machine learning approaches

5.1 Primary Lund plane and LSTM

A natural approach to adopt is the Deep-Learning technique used in the original study of the primary Lund Plane [33], which was already showing an excellent discriminating power in the context of boosted W tagging. Here we only consider the long short-term memory (LSTM) [50] network architecture as it showed the best performance in [33].

In practice, we input the list of Lund declusterings $\{(\ln \Delta_i, \ln k_{ti})\}$ to an LSTM of dimension 128 connected to a dropout layer (with rate 20%), with a final dense layer of dimension two and softmax activation. The network weights are initialised with a He uniform variance scaling initialiser [51], and the training is performed using an Adam optimisation algorithm [52] with a batch size of 128, a learning rate of 0.0005 and a categorical cross-entropy loss function. Our model is implemented using `TensorFlow` v2.1.0.

The data sample is split into a training sample (80%), a validation sample (10%) and a testing sample (10%). We train over a maximum of 50 epochs, with an early stopping when the performance does not increase over four epochs.

[9] This result is essentially the same as the $\mathcal{O}(\alpha_s^2)$ clustering logarithms contribution to the primary Lund plane density found in [36] (see eq. (3.25) there).

JHEP08(2022)177

For each configuration, we have run five independent trainings. For the quality measures reported below, the central value is obtained by averaging over the five runs and the uncertainty band is taken as their envelope.

5.2 Full Lund tree and Lund-Net

In order to take full advantage of the information contained in secondary leaves of the Lund plane, we consider the Lund-Net model introduced in ref. [30] and its associated code [53].

As input, we transform the tree of Lund declusterings into a graph, with the kinematic variables $\mathcal{T}$ of a declustering serving as attributes of a node on the graph. The Cambridge/Aachen clustering sequence is used to form bidirectional edges along the nodes connected in the Lund tree.

The graph architecture uses an EdgeConv operation [54], which applies a multi-layer perceptron (MLP) to each incoming edge of a node, using combined features of the node pairs as inputs, producing a learned edge feature. This initial shared MLP consists of two layers, each consisting of a dense network with batch normalisation [55] and ReLU activation [56], which are followed by an aggregation step taking an element-wise average of the learned edge features of the incoming edges as well as a shortcut connection [57]. The same MLP is applied to all nodes, leading to updated node features but keeping the structure of the graph unchanged. The Lund-Net architecture consists of six such EdgeConv blocks stacked together, and the number of channels of the MLPs are $(32, 32)$, $(32, 32)$, $(64, 64)$, $(64, 64)$, $(128, 128)$ and $(128, 128)$. Their output is concatenated for each node, and processed by a MLP with 384 channels, to which a global average pooling is applied to extract information from all nodes in the graph. This is followed by a final fully connected layer with 256 units and a dropout layer with rate 10%, with a softmax output giving the result of the classification. The Lund-Net model is implemented with the Deep Graph Library 0.5.3 [58] using the PyTorch 1.7.1 [59] backend, and training is performed for 30 epochs, using an Adam optimiser [52] to minimise the cross entropy loss. An initial learning rate of 0.001 is used, which is lowered by a factor 10 after the 10th and 20th epochs. As for the LSTM approach, the data sample is randomly split in 80/10/10% training/validation/testing samples, and we take the average and envelope of five runs.

In this paper, the inputs for each Lund declustering include, by default, $\ln \Delta$, $\ln k_t$, $\ln z$ and ψ. In section 6 which probes the collinear limit of our discriminants, the azimuthal angle is irrelevant and therefore not included in any of our approaches. Furthermore, in section 7.4 we discuss the effect of adding particle-ID information to the inputs, and in section 7.5 we discuss the effect of the azimuthal angle ψ. When imposing a cut on the (relative) transverse momentum, only the Lund declusterings with k_t above the cut are included in the data sample.

6 Validation in a pure-collinear (toy) parton shower

Before turning to a full Monte Carlo-based assessment of the discriminating performance of the tools introduced in the previous sections, we provide a cross-validation between the analytic and deep-learning approaches. To do this, we use a setup in which our analytic

JHEP08(2022)177

approach in sections 3.2 and 3.3 corresponds to the exact likelihood-ratio discriminant. This is achieved by generating events directly in the strong-angular-ordered limit, where our analytic approach from section 3 becomes exact: for simplicity, we use a fixed-coupling approximation, a fixed initial jet p_t of 1 TeV (with R=1)[10] and a fixed cut on emissions $k_{t,\text{cut}} = 1\,\text{GeV}$. We generate pure-quark and pure-gluon samples by starting the simulation with either an initial hard quark or an initial hard gluon. We simulate branchings using the full Altarelli-Parisi splitting functions, keeping track at each emission of the angle Δ and energy fraction z of the emission. In the strict collinear limit, a parton of momentum p_t branches in two partons of momenta $(1-z)p_t$ an zp_t, so the transverse momentum of each parton in the cascade — or, equivalently, its fraction of the initial jet p_t — can be deduced from the angles (Δ_i) and momentum fractions (z_i) at each branching. In practice, we have used (a slightly adapted version of)[11] the `microjet` code [60] to simulate events with strong angular ordering. The Lund declusterings are taken directly from the event trees, without any reclustering with the Cambridge/Aachen jet algorithm. This guarantees the absence of clustering logarithms.

Our analytic approaches to quark-gluon discrimination are applied as described in sections 3.2 and 3.3 except for two details: (i) they have been adapted to use a fixed-coupling approximation, and (ii) the Sudakov factor in eq. (3.7) has been computed keeping the full splitting function so as to guarantee that the resulting probability distributions match exactly that of the generated sample, including corrections strictly beyond our single-logarithmic approximation. With a fixed-coupling approximation, the calculation of the Sudakov exponent is relatively straightforward and expressions are given in appendix A for completeness.

With this setup in mind, we want to check that the machine learning (ML) approach using an LSTM network trained on primary Lund declusterings (section 5.1) converges to the same performance as what is given by the analytic approach in section 3.2. Similarly, we expect that the Lund-Net approach from section 5.2, trained on full Lund trees, gives the same performance as that of the analytic discriminant based on the full Lund tree in section 3.3. We also want to check that these new tools offer a better discriminating power than what is obtained using either the Iterated Soft Drop multiplicity or the average primary Lund plane density (see section 2).

In practice, we use a sample of 10^6 events generated with our adapted version of the `microjet` code, with α_s fixed either to 0.1 or to 0.5. These samples are either used to compute the analytic discriminant,[12] eqs. (3.3) or (3.12) or as inputs to train/validate/test our neural-network-based models. For the methods using machine-learning, the event sample is split in 80/10/10% training/validation/testing samples. This is repeated five

[10] Since we work in the pure collinear limit, the jet radius just plays the role of a reference scale for the logarithms of the angles, i.e. large-angle corrections scaling like powers of R are absent. The precise value of R is therefore irrelevant and we chose $R = 1$ for simplicity.

[11] Our adaptation compared to the original work in [60] mostly consists in imposing a k_t cutoff (instead of a small cut on z, as well as to keep the full tree of the generated cascade rather than just the final particles.

[12] For $\alpha_s = 0.1$, our analytic results have been obtained with a sample size of 10^7 events instead of the default sample of 10^6 events. This shows no visually observable differences on the results presented here.

JHEP08(2022)177

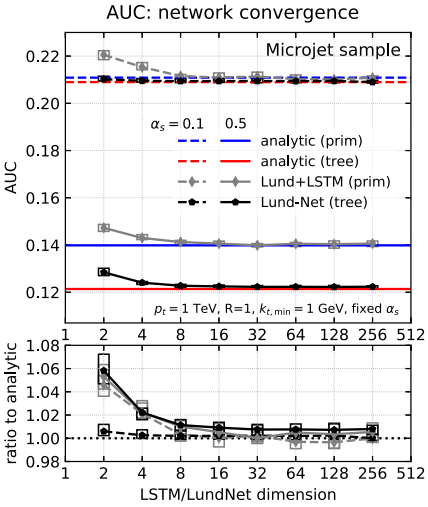


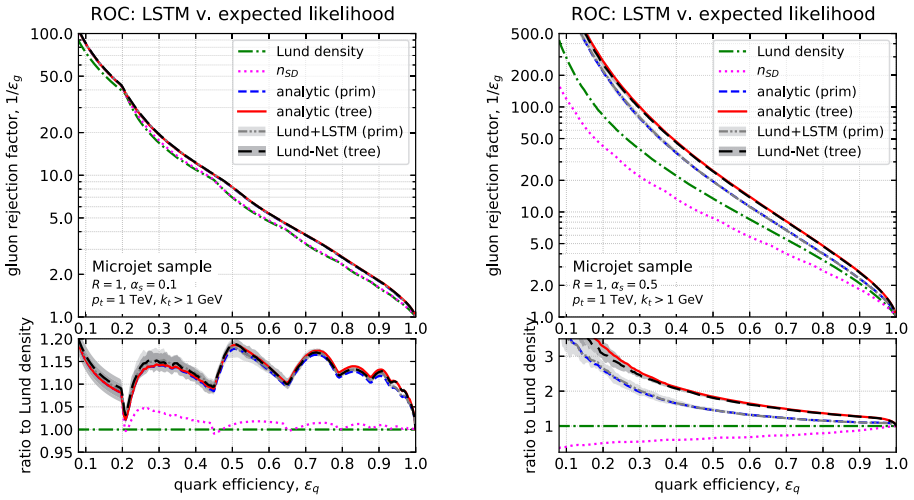
Figure 3. Area under the ROC curve (AUC) as a function of dimension of the network. The networks converge to the optimal (analytic) discriminant above $\approx$ 32 nodes, proving that the networks learned all the features of the `microjet` samples.

times for 5 different subdivisions and we take the average and envelope of these five runs respectively as an estimate of the performance and of the associated uncertainty. The quark and gluon efficiencies, ε_q and ε_g, are defined as the fraction of the events in the quark and gluon samples, respectively, which pass a cut on the analytic likelihood ratio or on the network discriminant.

Since the analytic models reproduce the exact likelihood ratio for these collinear samples, they are expected to provide the optimal discriminants. We first study how the area under the ROC curve (AUC) evolves as a function of the dimension of the LSTM or EdgeConv block in our machine-learning setup, varied between 2 and 256 nodes,[13] compared to the expected exact analytic result. This is shown in figure 3. Here, the AUC is defined as the area under the $(\varepsilon_q, \varepsilon_g)$ curve, meaning that a lower AUC means a better discriminating power. It is remarkable that for a network dimension of 32 or above, the neural network is able to reproduce the expected optimal discriminant to within at most 1%, for both values of α_s. If we look at the full ROC curves, figure 4, we see again the same level of agreement. The larger uncertainty at smaller quark efficiencies is expectable as only a fraction of the background events pass the tagger. Based on figures 3 and 4, we note a hierarchy between the classifiers with Lund-based methods performing better than the Iterated SoftDrop multiplicity and, among the Lund-based methods, the ones using the full tree information performing better that the ones using only primary declusterings. Improving the (logarithmic) accuracy of the analytic approach and exploiting more jet substructure

[13] For Lund-Net, the dimension refers to the size of the first MLP in the initial EdgeConv block, keeping the scaling of the successive layers identical to the one in section 5.2.

JHEP08(2022)177



(a) $\alpha_s = 0.1$.

(b) $\alpha_s = 0.5$.

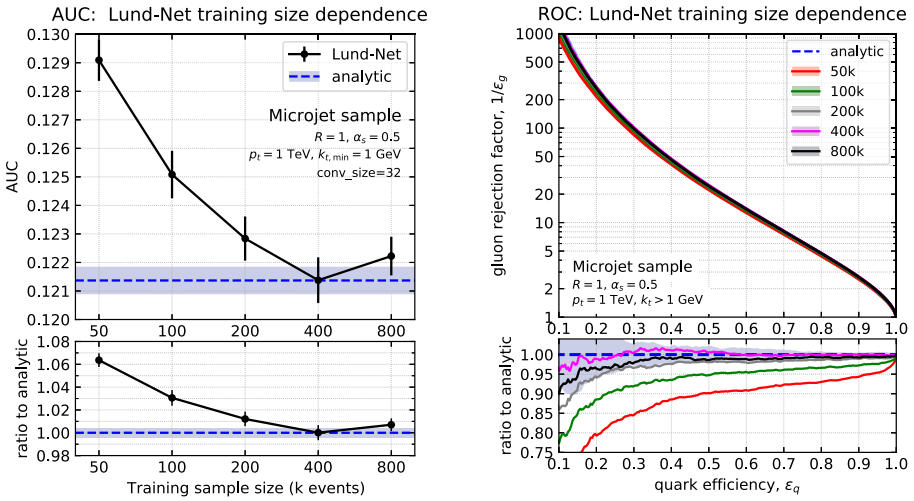
Figure 4. The ROC curves for the analytic and ML classifiers for the `microjet` sample. For this sample, our analytic approach is exact, showing that the neural networks capture the full (single-logarithmic) information in the training sample.

information both lead to a performance increase. Furthermore, the performance differences are enhanced as one opens up the phase space to include more emissions (or increase α_s).

To further investigate the $\lesssim 1\%$ difference between the Lund-Net and analytic results for $\alpha_s = 0.5$, we show in figure 5 the performance this time as a function of the size of the training sample (keeping the size of the validation and testing samples to 10^5 jets). The shaded band around the analytic expectation represent the statistical fluctuations obtained by splitting the full 10^6 event sample in 10 subsamples of 10^5 events, running our analytic discriminant independently on each subsample. For a testing sample size of 10^5 jets and $\varepsilon_q = 0.1$, we only keep $\sim 0.1\%$, i.e. ~ 100, of the gluon jets which is compatible with the $\sim 10\%$ observed statistical uncertainty. One sees that within the statistical uncertainties, the performance of the Lund-Net approach matches that of the analytic expectation for a training sample of 4×10^5 events or more. This is seen both for the AUC, figure 5a and for the ROC curves, figure 5b. In the latter case, the convergence is slightly slower at small $\varepsilon_q = 0.1$, as one could have expected.

Before we close this section, we note that additional tests targetting the asymptotic single-logarithmic limit of full Monte Carlo simulations will be carried on in section 7.6.

JHEP08(2022)177



(a) AUC.

(b) ROC curves.

Figure 5. Convergence of the ML-based Lund taggers as a function of the training sample size for the AUC (a) and ROC curve (b). The ML testing phase is always performed on the same 100k events and the shaded band around the analytic results is the fluctuations across different samples of 100k events. The optimal performance is reached for training sample sizes of 400k events and above.

7 Full Monte Carlo simulations

7.1 Setup

We now move to testing the performance of Lund-plane-based quark-gluon discriminants with full Monte Carlo samples. For our reference quark and gluon samples, we simulated respectively $Z+q$ and $Z+g$ events with Pythia v8.24 [61, 62] with multi-parton interactions enabled with the Monash13 tune [63]. For the $Z + q$ sample, only light quark flavours (u, d and s) have been included. The Z boson is set to decay to invisible neutrinos. Jets are then reconstructed on the remaining final-state using the anti-k_t algorithm [42] with $R = 0.4$, as implemented in FastJet [64, 65]. We select at most the two hardest jets within $|y| < 2.5$ and keep only the ones with $500 < p_t < 550\,\text{GeV}$. For each selected jet, we recluster its constituents with the Cambridge/Aachen algorithm and we construct the Lund declusterings following the recipe described in section 2. The studies described below are performed with quarks and gluon samples of 10^6 jets each.

To probe the resilience of our quark-gluon discriminants against various effects we have generated additional event samples. The first one uses the same setup as above with hadronisation and multi-parton interactions switched off, hence probing the influence of non-perturbative effects. The second uses dijet events, $qq \to qq$ with light quark flavours and $gg \to gg$, and is meant to probe the dependence on the hard process. The third one uses the same setup as the reference sample (Z+jet with non-perturbative effects enabled),

JHEP08(2022)177

this time generated with Herwig v7.2.0 [66, 67] so as to probe the dependence on the Monte Carlo generator.

Since our goal in this paper is to address "universal" aspects of quark-gluon tagging (recall the discussion in section 2), one can consider that varying the Monte Carlo generator and the hard process yielding the quark and gluon samples provide a picture of how much freedom there is beyond these universal aspects. For a practical analysis where one would consider specific signal and background processes, this discussion would have to be revisited and it would be interesting to include samples including full NLO matrix elements matched with parton shower, using tools such as MC@NLO [68], Powheg [69–71] or Sherpa [72, 73].

We test a total of six quark-gluon discriminants: the Iterated Soft Drop multiplicity (n_{SD}) and the discriminant based on the average Lund plane density (Lund density), both described in section 2, our new analytic discriminants using either the primary declusterings only (analytic(prim), sections 3.2 and 4.3) or the full declustering tree (analytic(tree), sections 3.3 and 4.2), and the deep-learning approaches using either only the primary Lund declusterings (Lund+LSTM (prim)) or the full Lund tree (Lund-Net (tree)) both described in section 5.

For the analytic models, clustering contributions are included with their dependence on the azimuthal angle ψ. We further discuss the influence of clustering logarithms and of the azimuthal angle ψ in section 7.5 below. Our analytic models are only considered in the presence of a k_t cut on the Lund declusterings, guaranteeing infrared-and-collinear safety.

As in the previous section, for the methods using machine learning, the event sample is subdivided into 8×10^5 training jets, 10^5 evaluation jets and 10^5 testing jets. We use five different subdivisions of the full sample to assess the uncertainties on the performance. For the discriminant based on the average Lund plane density, we use the first 9×10^5 events to build a (binned) estimate of $\rho_{q.g}(\Delta, k_t)$ and the 10^5 remaining events as a testing sample.

7.2 Tagging performance

We first look at the performance of our taggers. In this section we use our reference Monte Carlo sample, i.e. Z+jet events generated with Pythia with hadronisation and multi-parton interactions enabled. Figure 6 shows the area under the ROC curve as a function of the k_t cut applied on Lund declusterings, and figure 7 shows the ROC curves themselves for two specific choices of the cut: no cut (figure 7a) or $k_t > 1\,\text{GeV}$ (figure 7b).

Leaving aside for now the methods based on deep learning, we see the expected pattern. First, the average Lund density brings a small improvement compared to the Iterated SoftDrop approach. It is interesting to notice that while the performance of the Lund density approach flattens as the k_t cut is lowered, that of ISD gets worse at small k_t cuts. Since the Iterated SoftDrop multiplicity and our analytic approaches are based on perturbative QCD arguments, one might have anticipated their performance degradation for low values of the k_t cut. In particular, since our analytic models include the running of the strong coupling with k_t, they become unstable as we approach the Landau pole. The average Lund density approach however directly uses the Pythia sample to estimate the likelihood and is therefore free of these effects.

JHEP08(2022)177

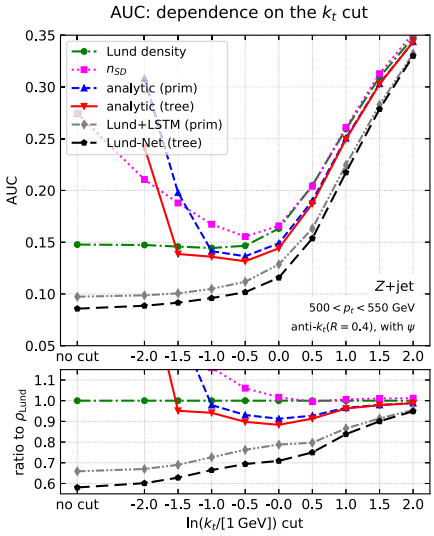


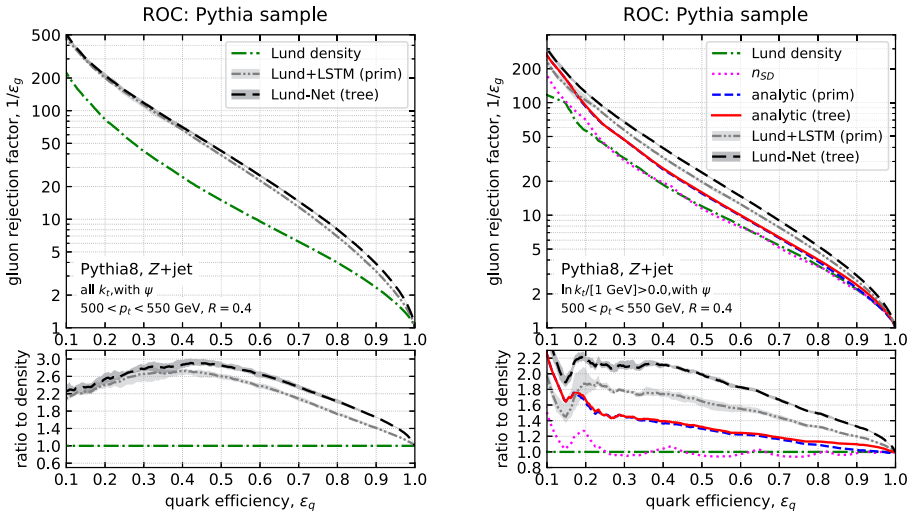
Figure 6. Area under the ROC curve (AUC) obtained using Z+jet events simulated with Pythia. All the Lund-plane based quark-gluon discriminants studied in the paper are shown: the Iterated Soft Drop multiplicity (n_{SD}), the likelihood based on the average primary Lund-plane density, and our new analytic and ML-based discriminants using either primary Lund branchings of the full Lund declustering tree. The AUC is plotted as a function of the minimum transverse momentum, k_t, cut imposed on the Lund declusterings.

Let us now focus on our analytic discriminants. Compared to the average density we see a $\sim$ 5% improvement in the AUC when using only the primary declusterings, reaching 9% for $k_{t,\text{cut}} = 1\,\text{GeV}$. As visible in figure 7b, this improvement increases towards smaller quark efficiency where it can reach 30−50% for ε_q in the 0.2−0.5 range. Adding the information from the full clustering tree, this improvement in AUC increases slightly, reaches e.g. $\sim$ 12% for $k_{t,\text{cut}} = 1\,\text{GeV}$. Looking at the ROC curve, this improvement is seen mostly at large ε_q with limited impact at smaller ε_q. As for ISD, the performance of our analytic models worsens for small k_t cuts, below 1 GeV. This is most likely due to a breakdown of the perturbative approach.

If we now turn to the Lund methods using deep learning, we see a clear improvement in discriminating power for all k_t cuts and across all values of the quark efficiency. The AUC is reduced (i.e. improved) by 20−40% for a cut on k_t below 1 GeV and the gluon rejection factor is improved by a factor between 2 and 3 for ε_q in the 0.2−0.5 range.

A striking feature of the machine-learning-based approaches in figures 6 and 7, especially compared to the results shown in the collinear sample in section 6, is that they show a substantial performance improvement compared to the analytic models. There can be several explanations for this. Of course, since our analytic approach is purely perturbative, differences can be of non-perturbative origin. This is certainly the case at very small

JHEP08(2022)177



(a) All k_t.

(b) $k_t > 1\,\text{GeV}$.

Figure 7. ROC curves corresponding to the AUC showed in figure 6. The left plot corresponds to all Lund declustering being included while for the right plot a k_t cut of 1 GeV has been applied. Since the methods anchored in perturbative QCD are meant to be effective in the region where perturbative QCD applies, they are only shown on the right plot.

values of $k_{t,\text{cut}}$ where our perturbative approach breaks down when the performance of the machine-learning-based approaches keeps improving. However, the gain is already visible at values of the k_t cut close to 10 GeV, where non-perturbative corrections are relatively small. From a pure perturbative perspective, there are at last three possible explanations as to why the machine learning approaches may outperform our analytic discriminant.

First, our treatment of clustering logarithms is only correct for pairs of emissions at commensurate angles so we should expect corrections even in the single-logarithmic limit.

Secondly, our analytic discriminant works in the limit of small angles, where quark-gluon discrimination can be thought of as universal (at least within our single-logarithmic approximation). The deep-learning methods will learn additional information, starting at the single-logarithmic accuracy, from radiation at large angle. Since this information is process-dependent, one should expect this gain in performance to come at the expense of an enlarged sensitivity to the hard process. We will come back to this point in section 7.3.

Lastly, there can be effects of subleading perturbative corrections that are not included in our analytic approach. These can either be subleading logarithmic corrections beyond single logarithms, or finite, fixed-order, corrections which would induce additional correlations between the Lund declusterings that are neglected at our analytic accuracy but that the neural network training would pick. In section 7.6, we show that if we take a more asymptotic regime, the gap between the analytic and deep-learning approaches shrink,

JHEP08(2022)177

strongly suggesting that the differences seen in figures 6 and 7 is not dominated by our simplified treatment of clustering logarithms.

We should also point out that subleading logarithmic corrections, or fixed-order corrections, are not fully included in Monte-Carlo event generators like Pythia. The improvement seen with deep-learning approaches should therefore be taken with caution.[14]

Of course, to this list of perturbative effects, one should also add non-perturbative corrections which, while beyond the reach of our analytic approach, are captured by the neural networks.

7.3 Resilience

The discriminating power of a quark-gluon tagger is not necessarily the only quality feature we may want to require. Indeed, our extraction of the tagging performance is obtained for a specific event sample which can have its own limitations or simply be different from the event sample used in later practical applications.

Ideally, one would want a tagger to be *resilient*, i.e. to show a degree of insensitivity to potential mismodelling aspects or to specific details of an event sample. For example, if we want to be able to describe a tagger from first-principles perturbative QCD, we would want to limit its sensitivity to the modelling of non-perturbative effects. In a similar spirit, we want to limit the sensitivity of a tagger to the details of the event generator used to obtain the event sample. More specifically, in the context of quark-gluon discrimination, we want our taggers to be insensitive to the details of the (hard) processes contributing to the signal and background(s) we try to separate. This last point is intimately related to the intrinsic ill-defined nature of quark-gluon tagging. In this context, resilience can be seen as a measure of universality.[15]

In this section, we therefore investigate the resilience of our tagger against the three effects listed above: (i) non-perturbative effects, (ii) the choice of the hard process and, (iii) the choice of a Monte Carlo event generator. The first is probed by comparing our reference sample to a sample generated at parton level, i.e. with hadronisation and multi-parton interactions switched off (see section 7.1 for details). For the second we use either our reference $Z+$jet sample or a sample of dijet event, and for the third, we compare our default Pythia8 sample to a Herwig7 sample.

For the analytic models, we apply them directly to the different event samples, obtaining in each case the quark and gluon efficiency as a function of the cut on the model's output, i.e. either the Iterated Soft Drop multiplicity, or the analytic likelihood ratio for the "Lund density" approach or for our new primary or full analytic discriminants. For

[14] It is tempting to argue that Monte Carlo event generators implement a more precise kinematics than the approximate one used in our analytic approach. For example, the Sudakov factor in the analytic calculations only retains the contributions up to single logarithms. For the fixed-coupling toy microjet sample used in section 6, we had instead kept the full z dependence of the splitting function in the Sudakov factor. We have checked that the effect of keeping the full splitting instead of keeping only the terms relevant at the single-logarithmic accuracy is, at most, 0.5%. This is clearly insufficient to explain the differences between the analytic and machine-learning approaches observed here.

[15] The idea of being resilient against details of the hard process however extends to tagging applications beyond quark-gluon discrimination.

JHEP08(2022)177

machine-learning approaches, we have trained the networks on our reference Pythia8 Z+jet sample with hadronisation and multi-parton interactions, and applied the resulting network to the other event samples.

For a given fully-specified tagger, i.e. a tagging method and cut on its output (a.k.a. a working point), one obtains quark and gluon efficiencies $\varepsilon_{q,g}^{(\mathrm{ref})}$ and $\varepsilon_{q,g}^{(\mathrm{alt})}$, respectively for the reference and alternative event samples. The resilience of the taggers is then simply defined as the inverse of the relative change between the two samples:

$$\zeta = \left[\left(\frac{2(\varepsilon_q^{(\mathrm{alt})} - \varepsilon_q^{(\mathrm{ref})})}{\varepsilon_q^{(\mathrm{alt})} + \varepsilon_q^{(\mathrm{ref})}} \right)^2 + \left(\frac{2(\varepsilon_g^{(\mathrm{alt})} - \varepsilon_g^{(\mathrm{ref})})}{\varepsilon_g^{(\mathrm{alt})} + \varepsilon_g^{(\mathrm{ref})}} \right)^2 \right]^{-1} . \tag{7.1}$$

With this choice, a bigger ζ corresponds to a more resilient tagger. We should then select a working point at which we evaluate the performance and resilience of a tagger. In practice, we take the point at which the tagging performance, defined as the significance $\varepsilon_q/\sqrt{\varepsilon_g}$, is maximal for the reference sample. This is typically realised for $\varepsilon \sim 0.3{-}0.5$. The resulting performance (significance) and resilience are denoted by Π_{best} and ζ_{best}, respectively. Ideally, we therefore seek for a tagger with large Π_{best} and ζ_{best}. We have tested that selecting instead a fixed quark efficiency, e.g. $\varepsilon_q = 0.5$, produces similar results.[16] That said, while quantitative arguments can be made about the relative discriminating performance of our taggers, it is more delicate to reach such a precise quantitative discussion of resilience. The discussion below therefore tries to remain mostly at a qualitative level, i.e. noting that taggers with larger resilience are likely to have less modelling uncertainties. It would be interesting — and clearly beyond the scope of this paper — to perform a dedicated study of resilience.

We want to study how resilience and performance behave for our discriminants, varying the k_t cut on Lund declusterings. Our results are presented in figure 8, for the three types of resilience we want to investigate: resilience against the specifics of the hard process (figure 8a), resilience against non-perturbative effects (figure 8b), and resilience against the choice of the event generator (figure 8c). To guide the eye, the results corresponding to a k_t cut of 1 GeV are represented with filled symbols, with all the other results using open symbols. For all three resiliences, the usual trade-off is observed: as we increase the k_t cut, performance decreases and resilience increases. Overall, our analytic models and our Deep Learning results show a similar behaviour, although a given performance-resilience point is achieved for a different k_t cut for different taggers. Our analytic models however appear as slightly more resilient to non-perturbative effects than their machine-learning equivalents.

Compared to the Iterative Soft Drop and Lund density approaches, one sees that the analytic model typically bring a gain in performance without sacrificing in resilience.

Focusing on the results with a k_t cut of 1 GeV, it is interesting to see that the machine-learning-based techniques reach a larger performance, as already seen in figures 6 and 7, at the expense of having a smaller resilience. This hints towards the interpretation that this

[16] In general, one can argue that the lower ε_q values should be ignored because they are subject to large statistical fluctuations. The large and low ε_q values are also impractical because they do not yield a large discriminating power.

JHEP08(2022)177

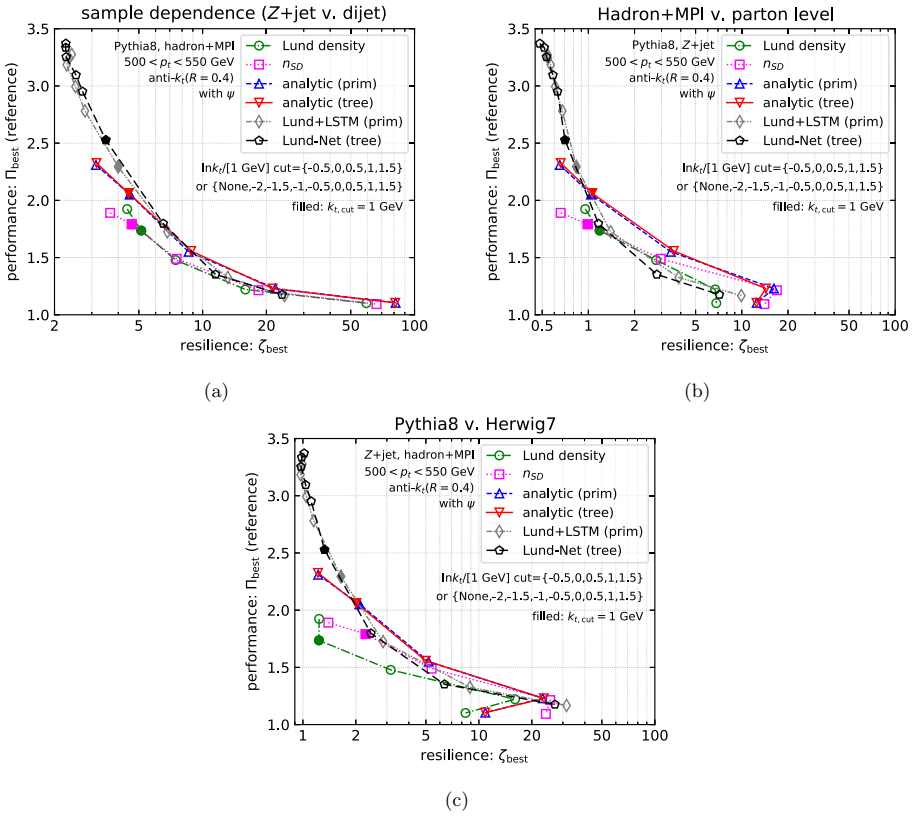


Figure 8. Plots of performance as a function of resilience for different discriminants. The curves are obtained by scanning over a range of accessible $k_{t,\text{cut}}$ values. The filled symbols correspond to a k_t cut of 1 GeV. Different plots correspond to different probes of resilience: (a) probes the sample dependence (replacing the Z+jet sample with a dijet sample, (b) probes non-perturbative effects (using parton-level simulations instead of full simulations with hadronisation and multi-parton interactions) and (c) probes the effect of the Monte Carlo generator (using Herwig 7 instead of Pythia 8).

gain in performance is obtained by the neural networks exploiting information (i) going beyond the "universal" collinear behaviour (worse resilience against the choice of hard process), (ii) in non-perturbative effects (worse resilience against hadronisation and MPI), and (iii) specific to the modelling of the events (worse resilience against the choice of event generator). In all three cases, increasing the k_t cut by a few hundred MeVs would result in a behaviour very similar to the one of the analytic model, both in terms of performance and in terms of resilience.

Finally, if all one cares about is performance, machine-learning discriminants using the full information in the Lund tree show the best result, albeit at the expense of a poor resilience. This should at least be kept in mind when using a quark-gluon discriminant for potentially different applications, or when assessing uncertainties associated with a tagger.

JHEP08(2022)177

7.4 Comparison with other approaches

In this section, we compare the performance of our Lund-plane-based taggers to that of other existing taggers.

We therefore select our taggers based on the full Lund tree: the Lund-Net tagger with no k_t cut and the analytic discriminant ("analytic(tree)" in previous figures) with a k_t cut of 1 GeV (referred to as "Lund NLL" in this section), and compare them to a series of pre-existing discriminants. We first consider benchmark jet shapes:

- angularities [19, 20], defined as the following sum over the jet constituents $\lambda_\alpha = (\sum_{i\in\text{jet}} p_{t,i}\Delta R^\alpha_{i,\text{jet}})/(R^\alpha \sum_{i\in\text{jet}} p_{t,i})$. We work with $\alpha = 1$, sometimes referred to as width or girth.

- energy-energy correlation functions [21], defined as the following sum over pairs of jet constituents $\text{EEC}_\beta = (\sum_{i,j\in\text{jet}} p_{t,i}p_{t,j}\Delta R^\beta_{ij})/[R^\beta(\sum_{i\in\text{jet}} p_{t,i})^2]$. In this case, we will set $\beta = 1/2$.

- to probe the effect of a k_t cut similar to the one we introduce in the Lund plane techniques, we have considered the case where the angularities and energy-correlation function are defined on the Lund declusterings (primary and secondary) above a given k_t cut. We recall that in this case, we expect that the Iterative Soft Drop multiplicity and our analytic Lund-tree discriminant are respectively optimal at leading and next-to-leading logarithmic accuracy in QCD.

In all cases, a cut is applied on one of these jet shapes and the full ROC curve is obtained by varying the cut. We then consider a series of recent machine-learning-based quark-gluon discriminants (for which we also apply a cut on the network output):

- Particle-Net described in ref. [28], based on point clouds. In practice, we have directly used the `ParticleNet` code available from [74], modifying the provided `keras` example to use our event sample. We have used a batch size of 1000 and kept the best model over a training of 50 epochs. Note that this model includes the particle ID in the network inputs.

- Particle-Flow Networks (PFN) from ref. [27]. This includes the rapidity, azimuth (both relative to the jet axis) and transverse momentum information of each jet constituent. Each particle is mapped into a per-particle latent space. The sum over all particles of these spaces is then mapped onto a final discriminating variable. We have also considered the so-called PFN-ID approach where the particle IDs are also included. In practice, we have used the code provided in the `EnergyFlow` package [75], modifying the examples to use our event samples and training over 60 epochs.

- Energy-Flow Networks (EFN) also from [27] and again only adapted from the example given in the `EnergyFlow` package to our needs. The approach is similar to that of the PFN above except that the latent space uses IRC-safe information through a weighting proportional to the p_t of each particle.

JHEP08(2022)177

Finally, we also consider the Lund-Net approach, labelled "Lund-Net(+ID)" where each Lund tuple $(\ln\Delta, \ln k_t, \ln z, \Psi)$ is supplemented by one additional integers for each of the two subjets j_1 and j_2 and determined as follows: if the subjet has a single constituent we use the PDG ID of the constituent, otherwise we set the integer to 0. The idea is similar to having particle identification added from the PFN to the PFN-ID approach.[17]

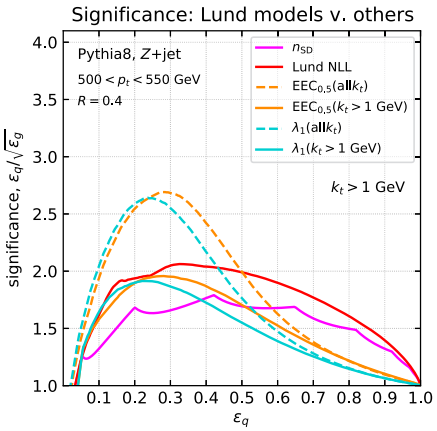
Our findings are presented in figure 9, for the signal significance (top row), and for the trade-off between significance and resilience against the choice of Monte-Carlo event generator (bottom row). Focusing first on the signal significance for analytic discriminants, figure 9a, we see relatively different patterns between the shape-based observables λ_1 and $\mathrm{EEC}_{0.5}$ and the Lund-based observables, with the performance peaking at larger values of the quark efficiency in the latter case, with shape-based observables reaching a better overall performance. However, computing the EEC and λ using the Lund declusterings with a k_t above 1 GeV, i.e. with the same input information in all cases, we see that our analytic Lund discriminant shows indeed an improvement at all quark efficiencies compared to the shape-based discriminants. A similar pattern is seen in the resilience plot on figure 9c with the Lund analytic model with a 1-GeV k_t cut being intermediate between the jet shape with a 1-GeV k_t cut and with no k_t cut. It is interesting to notice that the addition of particle ID information to the Lund-Net approach improves the performance at low k_t cut, or with no k_t cut at all, but changes neither the performance nor the resilience once a larger k_t cut is applied. This is most likely due to the fact that all the input subjets have more than one constituent and hence the ID information is 0. This contrasts with the findings in ref. [30], where the addition of the jet mass had a negative impact on resilience at large k_t cuts. At low resilience (large significance), the jet shapes give a slightly better performance vs. significance behaviour than our Lund-plane approach. Focusing instead on the AUC — the bottom-right table in figure 9 — we see that our analytic Lund-tree approach does a better job than the other jet shapes (i.e. a lower AUC), including jet shapes computed on the full set of constituents.

Moving to machine-learning-based models, figure 9b, we see a significance pattern mostly similar across different models. The performance of Lund-based models is on par with the one obtained from Particle-Net. Compared to energy/particle flow approaches, our Lund-based results show a slightly better performance than the EFN and PFN results, but fall slightly lower than the performance of PFN-ID. Adding the particle ID information using our Lund-Net(+ID) approach recovers a performance similar to the PFN-ID approach, although with a marginally smaller average peak performance. If we instead look at the AUC, we see that the Lund-Net(+ID) reaches the best performance (lowest AUC), marginally better that the PFN-ID and Lund-Net approaches, then followed by the Particle-Net model. While the PFN-ID method shows a small performance improvement at mid signal efficiency, the Lund-Net(+ID) setup has a small advantage at small and large signal efficiencies.[18] The observed differences are however of a size similar to the statisti-

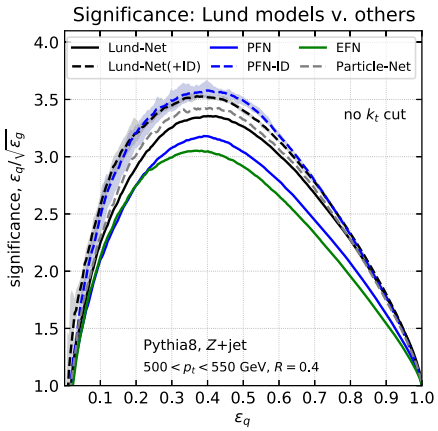
[17] In an experimental context, complete particle-ID information would not be available. One could however separate charged tracks from neutral energy deposits in calorimeters, with potential additional information such as heavy quarks or electromagnetic v. hadronic calorimeters.

[18] Including the particle ID information in a more coherent way, e.g. as a separate information that is fed to the final dense layers, one might be able to make up the difference with PFN-ID at mid quark efficiencies.

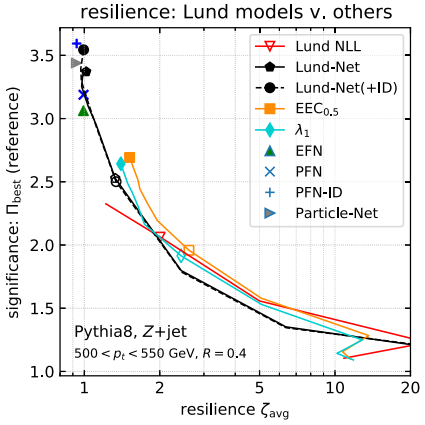
JHEP08(2022)177



(a)



(b)



(c)

model	AUC
n_{SD} ($k_t > 1$ GeV)	0.1658
Lund NLL ($k_t > 1$ GeV)	**0.1441**
EEC$_{0.5}$ (all k_t)	0.2074
EEC$_{0.5}$ ($k_t > 1$ GeV)	0.2150
λ_1 (all k_t)	0.2270
λ_1 ($k_t > 1$ GeV)	0.2371
Lund-Net	0.0858 ± 0.0007
Lund-Net(+ID)	**0.0835 ± 0.0005**
Particle-Net	0.0871 ± 0.0009
PFN	0.0994 ± 0.0009
PFN-ID	0.0853 ± 0.0005
EFN	0.1080 ± 0.0010

Figure 9. Comparison of the Lund-plane-based approaches with other models. Explicit plots of the signal significance $\varepsilon_q/\sqrt{\varepsilon_g}$ are shown in the upper plots, first for analytic discriminants, figure (a), then for machine-learning-based approaches, figure (b). The bottom panel, figure (c), shows the corresponding performance v. resilience plot, where the resilience is measured with respect to the choice of Monte Carlo generator (cf. section 7.3). The table on the bottom-right corner gives the area under the ROC curve (AUC) for the different models (lower is batter). For the ML-based models, the uncertainty is half the difference between the minimal and maximal values obtained over 5 different runs.

JHEP08(2022)177

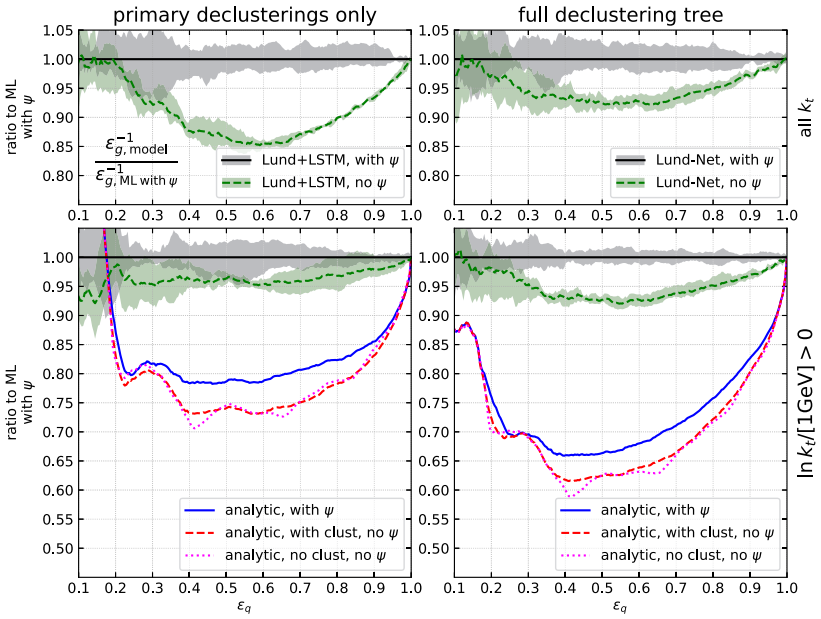


Figure 10. Plots showing the ratio of the ROC curve with and without azimuthal-angle dependence and/or clustering logarithms. The plots of the left column include only primary declusterings while the plots of the right column include the full declustering tree. The bottom plots include declusterings above a 1-GeV k_t cut, while the top plots include all relative transverse momenta. In all cases, the ROC curve is normalised to the ML Results including azimuthal-angle dependence.

cal fluctuations observed in our simulations (only shown, on the significance plot, for the Lund-Net(+ID) and PFN-ID for the sake of readability).

Finally, from figure 9c, we see a similar degree of resilience for all machine-learning-based approaches. Again, it would be interesting to train energy/particle flow networks or the ParticleNet network on Lund declusterings above a certain k_t cut (or using another cut definition) to study the performance versus resilience trade-off in a broader perspective.

7.5 Effect of clustering logarithms and of azimuthal angles

All the results presented so far have included the dependence on the Lund azimuthal angles ψ_i. Since it is known that these are not properly described at single-logarithmic accuracy by the standard dipole showers (including Pythia8) [76], we want to briefly investigate their impact on discriminating power. Additionally, from our analytic perspective, the dependence on the Lund azimuthal angles ψ_i only comes in through the clustering logarithms. This therefore gives us an explicit opportunity to investigate the numerical impact of including clustering logarithms into our analytic discriminants.

In figure 10, we show the background rejection, $1/\varepsilon_g$, including either only primary declusterings (left column) of the full declustering tree (right column), as obtained using

JHEP08(2022)177

our reference Pythia 8 Z+jet sample with hadronisation and multi-parton interactions. The top row, with only machine-learning results, corresponds to the case without a k_t cut while the bottom row, with both ML and analytic results, includes only the declusterings with $k_t \geq 1\,\text{GeV}$. To increase readability, we show in all cases, the ratio relative to what is obtained with the corresponding ML model — LSTM or Lund depending on whether only the primary or all the declusterings are used — including the ψ dependence.

We first discuss the ML results, presented in figure 10 either with (solid, black) or without (dashed, green) ψ information. We see that including the ψ information brings a 5−15% performance gain, mostly at intermediate quark efficiency. This gain is larger at lower k_t where non-perturbative effects are larger. When a 1-GeV k_t cut is imposed, we also show the results of our analytic quark-gluon taggers, again using either primary-only information (left column), or using the full declustering tree (right column). In each case, three results are given: including the ψ angles (solid, blue), not including the ψ angles but including the (ψ-averaged) clustering logarithms from eq. (4.5) (dashed, red), or including neither the ψ angles, nor the clustering logarithms (dotted, magenta). We again see a $\sim 10\%$ increase in performance brought by the inclusion of azimuthal angles. The fact that these performance gains are of similar magnitude in the deep-learning and analytic approaches indicates that our simplified treatment of clustering logarithms is a decent approximation.

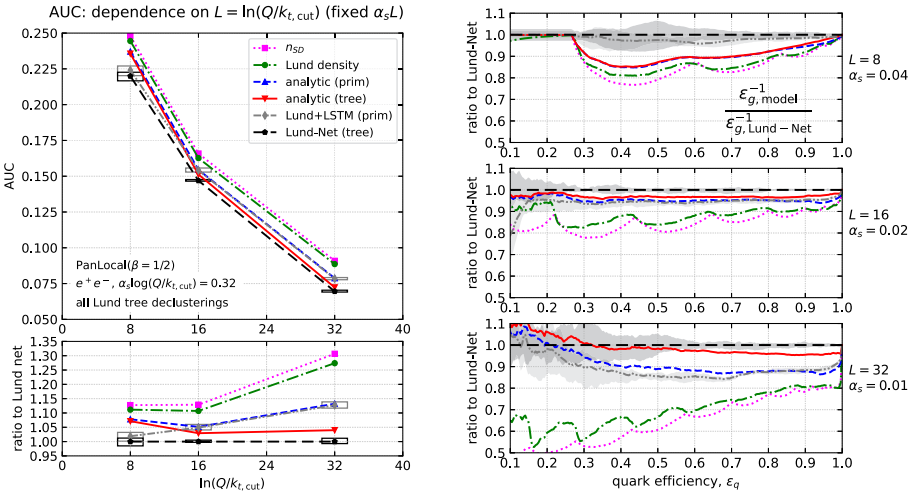
Finally, when the azimuthal angles are not included, we see that the influence of clustering logarithms is small.

7.6 Asymptotic single-logarithmic limit

In this final study, we want to further study the differences between the analytic and ML results. Our aim is here to take a limit where subleading effects decrease. Since our analytic approach technically resums double and single logarithms of $\log(p_t R/k_{t,\text{cut}})$, we want to proceed in a similar way as for the NLL-accuracy tests in [76], i.e, take the limit $\alpha_s(p_t R) \to 0$, $\log(p_t R/k_{t,\text{cut}}) \to \infty$ while keeping $\alpha_s(p_t R)\log(p_t R/k_{t,\text{cut}})$ constant. The main idea behind this limit is that subleading-logarithmic contributions as well as fixed-order contributions are suppressed as $\alpha_s \to 0$.

Generating and analysing events over an exponentially increasing range of scales poses a series of numerical challenges which, in practice, make it unreachable for standard Monte Carlo event generators like Pythia8. We have therefore used instead the PanScales e^+e^- code developed precisely to overcome these challenges in ref. [76]. We have therefore generated $e^+e^- \to Z \to q\bar{q}$ (quark) events and $e^+e^- \to H \to gg$ (gluon) events with a centre-of-mass energy Q, fixing $\alpha_s(Q)\log(Q/k_{t,\text{cut}}) = 0.32$ and taking $\alpha_s(Q)$ to be either 0.04, 0.02, 0.01, corresponding to $L = \log(Q/k_{t,\text{cut}})$ of either 8, 16, or 32. In all cases, we have used the PanLocal shower in its antenna variant with the β parameter set to 1/2. Subleading colour corrections are included using the Nested Ordered Double-Soft (NODS) scheme as described in [49]. This produces event samples with a single-logarithmic accuracy with the exception of the subleading-N_c corrections for which the NODS method only guarantees the correct behaviour for (any number of) pairs of emissions at commensurate angles. We reconstruct the e^+e^- Lund declusterings in each of the event hemispheres. The e^+e^- reconstruction follows an almost-trivial adaptation of the hadronic collisions

JHEP08(2022)177



(a) (b)

Figure 11. Dependence of (a) the AUC, and (b) the ROC curve on $L = \log(Q/k_{t,\mathrm{cut}})$ fixing $\alpha_s L = 0.32$. The results are obtained with an e^+e^- parton-level setup using the PanLocal PanScales shower which is NLL-accurate. Different curves show different methods and we see that the analytic and ML models converge to one another as we tale the asymptotic limit $L \to \infty$, $\alpha_s \to 0$, $\alpha_s L =$ constant. For the plots on the right column, we take the ratio of the ROC curves to the Lund-Net results.

technique described in section 2, except perhaps for the reconstruction of the azimuthal angle ψ which is described in details in [76].

Both our analytic methods and the approaches based on Machine Learning can be straightforwardly applied to these new sets of events. We therefore study the same set of quark-gluon discriminants as in section 7.2. The contribution from clustering logarithms is included in our analytic models and information from the Lund azimuthal angles is included both in the analytic and ML approaches.

Our results are presented in figure 11 for the AUC, figure 11a and for the ROC curves, figure 11b. For the latter, we have normalised the gluon rejection rate ε_g^{-1} to the Lund-Net rejection rate. One can see from these plots that the difference between the analytic and Machine-Learning-based methods decrease when increasing L (decreasing α_s) at fixed $\alpha_s L$. This is true separately for the approaches using only primary declusterings ("analytic(prim)" and "Lund+LSTM") and for the approaches using the full declustering tree ("analytic(tree)" and "Lund-Net"). At the same time, the performance gain compared to the Iterated SoftDrop multiplicity increases.

Before closing this section, we want to address a last point about the azimuthal angle dependence and single-logarithmic accuracy. It has been pointed out in ref. [76] that, due to non-physical recoil effects, dipole showers such as Pythia or Dire, would generate a spurious dependence on ψ, potentially biasing the assessment of quark-gluon classification. Using

JHEP08(2022)177

the machinery described in this section, we have studied potential differences between the Pythia shower and the PanScales showers which are free of this effect [76]. Within the few-percent accuracy of our studies we have not been able to isolate a clearly-visible impact of this effect.

8 Conclusions

This paper addresses the question of quark/gluon discrimination using the Lund-plane approach to characterise the substructure of jets. Our main result is that it is possible to compute the quark-gluon likelihood ratio from the first principles in QCD. The calculation is done at the single-logarithmic accuracy, including all collinear contributions as well as clustering effects for any number of pairs of emissions at commensurate angles. This automatically provides us with an optimal quark/gluon tagger at the same accuracy.

As expectable, this tagger shows an improved performance either compared to using the average Lund plane density to build the likelihood ratio, or compared to the Iterated Soft Drop multiplicity which corresponds to the optimal quark/gluon discriminant at leading (double) logarithmic accuracy. Most of the improvement ($\lesssim 10\%$ for the AUC compared to n_{SD}) is already captured when including only primary declusterings, but the effect of additional declustering in subsidiary Lund planes ($\sim 3\%$) is clearly visible, especially at larger quark efficiencies. The gain in performance can be attributed to the better treatment of the kinematics of each emission, e.g. through the full Altarelli-Parisi splitting functions, through the full antenna pattern for emissions at commensurate angles, and to a better treatment of the correlations between emissions, e.g. by taking into account the energy of the emitting parton or by including clustering effects. Furthermore, this gain in performance is accompanied by a gain in resilience against effects beyond our perturbative calculation. In this context, we have studied three specific effects: the dependence against non-perturbative effects, the dependence against the specific choice of quark/gluon enriched samples used as benchmarks, and the choice of Monte Carlo event generator.

In section 7.4, we have compared our Lund-based approach to other typical quark/gluon taggers using jet substructure, like angularities or energy correlation functions. Focusing for simplicity on the case where all the taggers are applied to Lund declusterings with a k_t above 1 GeV to reduce non-perturbative effects, we see that the Lund-based likelihood approach gives a gain in performance, especially at large quark efficiency, while maintaining a similar degree of resilience.[19]

The second set of results in this paper is the extensive study of quark/gluon taggers using deep-learning techniques combined with Lund declustering information. When applied to the full set of declusterings in a jet, our ML-based tagger reaches a performance (and resilience) comparable to that obtained with Particle-Flow networks [27], and marginally better than what is achieved by Particle-Net [28]. Compared to the analytic results, our ML-tagger gives a clearly visible performance gain, even when considering declusterings above a given k_t cut. This gain in performance, however, comes at a price in all the forms

[19] If the jet shapes are computed on the full jet, they yield a larger significance at smaller quark efficiency at the expense of a reduced resilience against non-perturbative effects.

JHEP08(2022)177

of resilience we have studied, especially the sensitivity to hadronisation and multi-parton interactions.

One of the key points of this paper is the direct comparison between the analytic and deep-learning approaches. Since our tagger targets the optimal discriminant in the single-logarithmic approximation one can directly compare its performance with that of the deep-learning models. We first did that in the strongly angular-ordered limit where our analytic calculation is exact. The results in section 6 indicate a convergence of the deep-learning taggers to the optimal performance as long as the size of the network is taken large enough. Beyond the collinear limit, where our analytic treatment is only approximate, the deep-learning approach shows a better performance. However, if we move progressively to the single-logarithmic asymptotic limit ($\alpha_s \to 0$ at fixed $\alpha_s \log(Q/k_{t,\text{cut}})$ the difference between the two approaches drastically reduces as we showed in section 7.6. At the same time, the gain in performance compared to leading-logarithmic-accurate taggers — i.e. the average Lund density and the Iterated Soft Drop multiplicity — increases.

The above observations strongly suggest that the gain in performance observed for deep-learning taggers (in addition to our analytic tagger) in phenomenological Monte Carlo applications come predominantly either from subleading effects (beyond single logarithms), from large-angle soft emissions (not included in our analytic calculation), or from non-perturbative effects.

This points towards several interesting physics considerations. First, subleading logarithmic effects, albeit present in data, are not properly included in any parton shower Monte Carlo generator today. Conclusions regarding subdominant logarithmic effects should therefore be taken carefully. In this context, it would be interesting in the future to further investigate potential differences between standard dipole showers (like Pythia) which are known to have failures at the single-logarithmic accuracy, or even at leading-log for subleading colour effects, and the PanScales showers which are NLL-accurate.

Then, large-angle soft emissions are process-dependent and should therefore be treated carefully when applied outside the configurations where they have been tested and calibrated. In the future, it would be interesting to see if an analytic treatment similar to the one adopted in this paper could allow for quantitative assessment of the process-dependence of quark/gluon discrimination (see also ref. [46] for a Monte-Carlo-based study).

Finally, non-perturbative effects come with non-negligible modelling uncertainties and should therefore also be taken carefully. The ability to progressively reduce non-perturbative effects by increasing the k_t cut-off on Lund declusterings could help further investigating the impact of non-perturbative effects, and the associated systematic uncertainties, in a practical context.

In conclusion, we have seen that Lund-plane declusterings were useful to define a variety of quark/gluon discriminants, bridging regions targetting high discriminating performance and regions where a high-precision degree of control can be reached from first-principles QCD.

JHEP08(2022)177

Acknowledgments

G.S. is especially grateful to Ben Nachman and Eric Metodiev for stimulating discussions. G.S. and F.D. are also grateful to Gavin Salam for earlier work on the Lund jet plane without which this paper would never have been possible. A.T. wishes to thank the Institut de Physique Theorique for the hospitality. This work has been supported in part by the French Agence Nationale de la Recherche, under grant ANR-15-CE31-0016 (G.S. and A.T.), by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 788223, PanScales) (G.S.), by the Starting Grant from Trond Mohn Foundation (BFS2018REK01) (A.T.), and by a Royal Society University Research Fellowship (URF\R1\211294) (F.D.).

A Sudakov factors with exact splitting functions

In section 6, we have used an event sample generated in the strong-angular-ordered limit to compare the performance of our analytic discriminants (also in the limit of strong angular ordering, see section 3 to that of deep-learning-based discriminants. To guarantee that the analytic approach reproduces the exact likelihood ratio, we have kept the full Altarelli-Parisi splitting functions in the Sudakov (using a fixed-coupling approximation). If we have a hard parton of momentum xp_t (with p_t the initial transverse momentum of the jet) and flavour f, and compute the Sudakov factor between an angle Δ_{i-1} and Δ_i, we find:

$$-\log S_f^{(i-1,i)} = \frac{2\alpha_s C_f}{pi}\left[\frac{\log^2 x_1}{2} - \frac{\log^2 x_2}{2} + B_f \log\frac{x_2}{x_1} + \mathrm{Li}_2(x_1) - \mathrm{Li}_2(x_2) + \delta R_f\right], \quad \text{(A.1)}$$

with

$$x_1 = \frac{k_{t,\mathrm{cut}}}{x\Delta_{i-1}p_t}, \qquad B_q = -\frac{3}{4}, \quad \text{(A.2)}$$

$$x_2 = \frac{k_{t,\mathrm{cut}}}{x\Delta_i p_t}, \qquad B_g = -\frac{11C_A - 4n_f T_R}{12C_A}, \quad \text{(A.3)}$$

and

$$\delta R_q = \frac{3}{2}(x_2 - x_1), \quad \text{(A.4)}$$

$$\delta R_g = \frac{3}{2}(x_2 - x_1) + \left(\frac{1}{2} - \frac{n_f T_R}{C_A}\right)\left[(x_2 - x_1) - \frac{1}{2}(x_2^2 - x_1^2) + \frac{2}{9}(x_2^3 - x_1^3)\right]. \quad \text{(A.5)}$$

JHEP08(2022)177

References

[1] G.P. Salam, *Towards jetography*, *Eur. Phys. J. C* **67** (2010) 637 [arXiv:0906.1833] [INSPIRE].

[2] S. Marzani, G. Soyez and M. Spannowsky, *Looking inside jets: an introduction to jet substructure and boosted-object phenomenology*, *Lect. Notes Phys.* **958** (2019) 1 [arXiv:1901.10342] [INSPIRE].

[3] P. Gras et al., *Systematics of quark/gluon tagging*, *JHEP* **07** (2017) 091 [arXiv:1704.03878] [INSPIRE].

[4] P.T. Komiske, E.M. Metodiev and J. Thaler, *An operational definition of quark and gluon jets*, *JHEP* **11** (2018) 059 [arXiv:1809.01140] [INSPIRE].

[5] A.J. Larkoski and E.M. Metodiev, *A theory of quark vs. gluon discrimination*, *JHEP* **10** (2019) 014 [arXiv:1906.01639] [INSPIRE].

[6] A. Banfi, G.P. Salam and G. Zanderighi, *Infrared safe definition of jet flavor*, *Eur. Phys. J. C* **47** (2006) 113 [hep-ph/0601139] [INSPIRE].

[7] A.J. Larkoski, I. Moult and B. Nachman, *Jet substructure at the Large Hadron Collider: a review of recent advances in theory and machine learning*, *Phys. Rept.* **841** (2020) 1 [arXiv:1709.04464] [INSPIRE].

[8] R. Kogler et al., *Jet substructure at the Large Hadron Collider: experimental review*, *Rev. Mod. Phys.* **91** (2019) 045003 [arXiv:1803.06991] [INSPIRE].

[9] L. de Oliveira, M. Kagan, L. Mackey, B. Nachman and A. Schwartzman, *Jet-images — deep learning edition*, *JHEP* **07** (2016) 069 [arXiv:1511.05190] [INSPIRE].

[10] P. Baldi, K. Bauer, C. Eng, P. Sadowski and D. Whiteson, *Jet substructure classification in high-energy physics with deep neural networks*, *Phys. Rev. D* **93** (2016) 094034 [arXiv:1603.09349] [INSPIRE].

[11] G. Kasieczka, T. Plehn, M. Russell and T. Schell, *Deep-learning top taggers or the end of QCD?*, *JHEP* **05** (2017) 006 [arXiv:1701.08784] [INSPIRE].

[12] A. Butter et al., *The machine learning landscape of top taggers*, *SciPost Phys.* **7** (2019) 014 [arXiv:1902.09914] [INSPIRE].

[13] H. Qu and L. Gouskos, *ParticleNet: jet tagging via particle clouds*, *Phys. Rev. D* **101** (2020) 056019 [arXiv:1902.08570] [INSPIRE].

[14] E.A. Moreno et al., *JEDI-net: a jet identification algorithm based on interaction networks*, *Eur. Phys. J. C* **80** (2020) 58 [arXiv:1908.05318] [INSPIRE].

[15] CMS collaboration, *Performance of the DeepJet b tagging algorithm using* 41.9 fb^{-1} *of data from proton-proton collisions at* 13 *TeV with phase* 1 *CMS detector*, CMS-DP-2018-058, CERN, Geneva, Switzerland (2018).

[16] Y.-L. Du, D. Pablos and K. Tywoniuk, *Jet tomography in heavy-ion collisions with deep learning*, *Phys. Rev. Lett.* **128** (2022) 012301 [arXiv:2106.11271] [INSPIRE].

[17] Y.-L. Du, D. Pablos and K. Tywoniuk, *Deep learning jet modifications in heavy-ion collisions*, *JHEP* **03** (2021) 206 [arXiv:2012.07797] [INSPIRE].

[18] L. Apolinário, N.F. Castro, M. Crispim Romão, J.G. Milhano, R. Pedro and F.C.R. Peres, *Deep learning for the classification of quenched jets*, *JHEP* **11** (2021) 219 [arXiv:2106.08869] [INSPIRE].

JHEP08(2022)177

[19] C.F. Berger, T. Kucs and G.F. Sterman, *Event shape/energy flow correlations*, *Phys. Rev. D* **68** (2003) 014012 [hep-ph/0303051] [INSPIRE].

[20] L.G. Almeida, S.J. Lee, G. Perez, G.F. Sterman, I. Sung and J. Virzi, *Substructure of high-p_T jets at the LHC*, *Phys. Rev. D* **79** (2009) 074017 [arXiv:0807.0234] [INSPIRE].

[21] A.J. Larkoski, G.P. Salam and J. Thaler, *Energy correlation functions for jet substructure*, *JHEP* **06** (2013) 108 [arXiv:1305.0007] [INSPIRE].

[22] R.D. Field and R.P. Feynman, *A parametrization of the properties of quark jets*, *Nucl. Phys. B* **136** (1978) 1 [INSPIRE].

[23] D. Krohn, M.D. Schwartz, T. Lin and W.J. Waalewijn, *Jet charge at the LHC*, *Phys. Rev. Lett.* **110** (2013) 212001 [arXiv:1209.2421] [INSPIRE].

[24] Z.-B. Kang, X. Liu, S. Mantry, M.C. Spraker and T. Wilson, *Dynamic jet charge*, *Phys. Rev. D* **103** (2021) 074028 [arXiv:2101.04304] [INSPIRE].

[25] C. Frye, A.J. Larkoski, J. Thaler and K. Zhou, *Casimir meets Poisson: improved quark/gluon discrimination with counting observables*, *JHEP* **09** (2017) 083 [arXiv:1704.06266] [INSPIRE].

[26] G. Louppe, K. Cho, C. Becot and K. Cranmer, *QCD-aware recursive neural networks for jet physics*, *JHEP* **01** (2019) 057 [arXiv:1702.00748] [INSPIRE].

[27] P.T. Komiske, E.M. Metodiev and J. Thaler, *Energy flow networks: deep sets for particle jets*, *JHEP* **01** (2019) 121 [arXiv:1810.05165] [INSPIRE].

[28] H. Qu and L. Gouskos, *ParticleNet: jet tagging via particle clouds*, *Phys. Rev. D* **101** (2020) 056019 [arXiv:1902.08570] [INSPIRE].

[29] J.S.H. Lee, S.M. Lee, Y. Lee, I. Park, I.J. Watson and S. Yang, *Quark gluon jet discrimination with weakly supervised learning*, *J. Korean Phys. Soc.* **75** (2019) 652 [arXiv:2012.02540] [INSPIRE].

[30] F.A. Dreyer and H. Qu, *Jet tagging in the Lund plane with graph networks*, *JHEP* **03** (2021) 052 [arXiv:2012.08526] [INSPIRE].

[31] E.M. Metodiev and J. Thaler, *Jet topics: disentangling quarks and gluons at colliders*, *Phys. Rev. Lett.* **120** (2018) 241602 [arXiv:1802.00008] [INSPIRE].

[32] J. Brewer, J. Thaler and A.P. Turner, *Data-driven quark and gluon jet modification in heavy-ion collisions*, *Phys. Rev. C* **103** (2021) L021901 [arXiv:2008.08596] [INSPIRE].

[33] F.A. Dreyer, G.P. Salam and G. Soyez, *The Lund jet plane*, *JHEP* **12** (2018) 064 [arXiv:1807.04758] [INSPIRE].

[34] ATLAS collaboration, *Measurement of the Lund jet plane using charged particles in* 13 *TeV proton-proton collisions with the ATLAS detector*, *Phys. Rev. Lett.* **124** (2020) 222002 [arXiv:2004.03540] [INSPIRE].

[35] ALICE collaboration, *Physics preliminary summary: measurement of the primary Lund plane density in pp collisions at* $\sqrt{s} = 13$ *TeV with ALICE*, ALICE-PUBLIC-2021-002, CERN, Geneva, Switzerland (2021).

[36] A. Lifson, G.P. Salam and G. Soyez, *Calculating the primary Lund jet plane density*, *JHEP* **10** (2020) 170 [arXiv:2007.06578] [INSPIRE].

[37] D.E. Soper and M. Spannowsky, *Finding physics signals with shower deconstruction*, *Phys. Rev. D* **84** (2011) 074002 [arXiv:1102.3480] [INSPIRE].

JHEP08(2022)177

[38] D.E. Soper and M. Spannowsky, *Finding top quarks with shower deconstruction*, *Phys. Rev. D* **87** (2013) 054012 [arXiv:1211.3140] [INSPIRE].

[39] D. Ferreira de Lima, P. Petrov, D. Soper and M. Spannowsky, *Quark-gluon tagging with shower deconstruction: unearthing dark matter and Higgs couplings*, *Phys. Rev. D* **95** (2017) 034001 [arXiv:1607.06031] [INSPIRE].

[40] Y.L. Dokshitzer, G.D. Leder, S. Moretti and B.R. Webber, *Better jet clustering algorithms*, *JHEP* **08** (1997) 001 [hep-ph/9707323] [INSPIRE].

[41] M. Wobisch and T. Wengler, *Hadronization corrections to jet cross-sections in deep inelastic scattering*, in *Workshop on Monte Carlo generators for HERA physics (plenary starting meeting)*, (1998), p. 270 [hep-ph/9907280] [INSPIRE].

[42] M. Cacciari, G.P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, *JHEP* **04** (2008) 063 [arXiv:0802.1189] [INSPIRE].

[43] ALICE collaboration, *QCD dynamics studied with jets in ALICE*, in 55$^{\text{th}}$ *rencontres de Moriond on QCD and high energy interactions*, (2021) [arXiv:2105.10523] [INSPIRE].

[44] A.J. Larkoski, S. Marzani, G. Soyez and J. Thaler, *Soft drop*, *JHEP* **05** (2014) 146 [arXiv:1402.2657] [INSPIRE].

[45] M. Dasgupta, A. Fregoso, S. Marzani and G.P. Salam, *Towards an understanding of jet substructure*, *JHEP* **09** (2013) 029 [arXiv:1307.0007] [INSPIRE].

[46] S. Bright-Thonney and B. Nachman, *Investigating the topology dependence of quark and gluon jets*, *JHEP* **03** (2019) 098 [arXiv:1810.05653] [INSPIRE].

[47] M. Dasgupta and G.P. Salam, *Resummation of nonglobal QCD observables*, *Phys. Lett. B* **512** (2001) 323 [hep-ph/0104277] [INSPIRE].

[48] Y. Hatta and T. Ueda, *Resummation of non-global logarithms at finite N_c*, *Nucl. Phys. B* **874** (2013) 808 [arXiv:1304.6930] [INSPIRE].

[49] K. Hamilton, R. Medves, G.P. Salam, L. Scyboz and G. Soyez, *Colour and logarithmic accuracy in final-state parton showers*, *JHEP* **03** (2021) 041 [arXiv:2011.10054] [INSPIRE].

[50] S. Hochreiter and J. Schmidhuber, *Long short-term memory*, *Neural Comput.* **9** (1997) 1735.

[51] K. He, X. Zhang, S. Ren and J. Sun, *Delving deep into rectifiers: surpassing human-level performance on ImageNet classification*, arXiv:1502.01852 [INSPIRE].

[52] D.P. Kingma and J. Ba, *Adam: a method for stochastic optimization*, 12, 2014 [arXiv:1412.6980] [INSPIRE].

[53] F.A. Dreyer and H. Qu, *fdreyer/LundNet: v*1.0.0, Zenodo, (2021).

[54] Y. Wang, Y. Sun, Z. Liu, S.E. Sarma, M.M. Bronstein and J.M. Solomon, *Dynamic graph CNN for learning on point clouds*, *ACM Trans. Graph.* **38** (2019) 1.

[55] S. Ioffe and C. Szegedy, *Batch normalization: accelerating deep network training by reducing internal covariate shift*, in *Proceedings of the* 32$^{\text{nd}}$ *international conference on machine learning*, Lille, France, 7–9 July 2015, *Proc. Machine Learn. Res.* **37** (2015) 448.

[56] X. Glorot, A. Bordes and Y. Bengio, *Deep sparse rectifier neural networks*, in *Proceedings of the* 14$^{\text{th}}$ *international conference on artificial intelligence and statistics*, Fort Lauderdale, FL, U.S.A., 11–13 April 2011, *Proc. Machine Learn. Res.* **15** (2011) 315.

[57] K. He, X. Zhang, S. Ren and J. Sun, *Deep residual learning for image recognition*, in 2016 *IEEE conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, U.S.A., IEEE, (2016), p. 770 [arXiv:1512.03385] [INSPIRE].

JHEP08(2022)177

[58] M. Wang et al., *Deep graph library: a graph-centric, highly-performant package for graph neural networks*, arXiv:1909.01315.

[59] A. Paszke et al., *Pytorch: an imperative style, high-performance deep learning library*, in *Advances in neural information processing systems* 32, H. Wallach et al. eds., Curran Associates Inc., (2019), p. 8024.

[60] M. Dasgupta, F. Dreyer, G.P. Salam and G. Soyez, *Small-radius jets to all orders in QCD*, *JHEP* **04** (2015) 039 [arXiv:1411.5182] [INSPIRE].

[61] T. Sjöstrand, S. Mrenna and P.Z. Skands, *PYTHIA* 6.4 *physics and manual*, *JHEP* **05** (2006) 026 [hep-ph/0603175] [INSPIRE].

[62] T. Sjöstrand et al., *An introduction to PYTHIA* 8.2, *Comput. Phys. Commun.* **191** (2015) 159 [arXiv:1410.3012] [INSPIRE].

[63] P. Skands, S. Carrazza and J. Rojo, *Tuning PYTHIA* 8.1*: the Monash* 2013 *tune*, *Eur. Phys. J. C* **74** (2014) 3024 [arXiv:1404.5630] [INSPIRE].

[64] M. Cacciari and G.P. Salam, *Dispelling the* N^3 *myth for the* k_t *jet-finder*, *Phys. Lett. B* **641** (2006) 57 [hep-ph/0512210] [INSPIRE].

[65] M. Cacciari, G.P. Salam and G. Soyez, *FastJet user manual*, *Eur. Phys. J. C* **72** (2012) 1896 [arXiv:1111.6097] [INSPIRE].

[66] M. Bahr et al., *HERWIG++ physics and manual*, *Eur. Phys. J. C* **58** (2008) 639 [arXiv:0803.0883] [INSPIRE].

[67] J. Bellm et al., *HERWIG* 7.0*/HERWIG++* 3.0 *release note*, *Eur. Phys. J. C* **76** (2016) 196 [arXiv:1512.01178] [INSPIRE].

[68] S. Frixione and B.R. Webber, *Matching NLO QCD computations and parton shower simulations*, *JHEP* **06** (2002) 029 [hep-ph/0204244] [INSPIRE].

[69] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040 [hep-ph/0409146] [INSPIRE].

[70] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070 [arXiv:0709.2092] [INSPIRE].

[71] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043 [arXiv:1002.2581] [INSPIRE].

[72] E. Gerwick, S. Hoeche, S. Marzani and S. Schumann, *Soft evolution of multi-jet final states*, *JHEP* **02** (2015) 106 [arXiv:1411.7325] [INSPIRE].

[73] SHERPA collaboration, *Event generation with Sherpa* 2.2, *SciPost Phys.* **7** (2019) 034 [arXiv:1905.09127] [INSPIRE].

[74] H. Qu and L. Gouskos, *Particle-net*, https://github.com/hqucms/ParticleNet, accessed 21 June 2021.

[75] P. Komiske, E. Metodiev and J. Thaler, *Energy flow*, https://energyflow.network, accessed 21 June 2021.

[76] M. Dasgupta, F.A. Dreyer, K. Hamilton, P.F. Monni, G.P. Salam and G. Soyez, *Parton showers beyond leading logarithmic accuracy*, *Phys. Rev. Lett.* **125** (2020) 052002 [arXiv:2002.11114] [INSPIRE].

JHEP08(2022)177

Article III.

5.3 Quenching effects in the cumulative jet spectrum

Adam Takacs and Konrad Tywoniuk

JHEP, **10**, 038 (2021)

Paper about jet suppression in heavy-ion collisions. Keywords: jets, jet quenching, jet energy-loss, nuclear modification, medium-induced emissions, out-of-cone radiation, quantile, topics, quark/gluon jet classification.

Published for SISSA by Springer

Received: *April 8, 2021*
Revised: *August 20, 2021*
Accepted: *September 19, 2021*
Published: *October 5, 2021*

Quenching effects in the cumulative jet spectrum

Adam Takacs[a,b] **and Konrad Tywoniuk**[a]

[a] *Department of Physics and Technology, University of Bergen, 5007 Bergen, Norway*

[b] *Department of Astronomy and Theoretical Physics, Lund University, S-223 62 Lund, Sweden*

E-mail: adam.takacs@uib.no, konrad.tywoniuk@uib.no

Abstract: The steeply falling jet spectrum induces a bias on the medium modifications of jet observables in heavy-ion collisions. To explore this effect, we develop a novel analytic framework to study the quenched jet spectrum and its cumulative. We include many energy-loss-related effects, such as soft and hard medium induced emissions, broadening, elastic scattering, jet fragmentation, cone size dependence, and coherence effects. We show that different observables, based on the jet spectrum, are connected, e.g., the nuclear modification, spectrum shift, and the quantile procedure. We present the first predictions for the nuclear modification factor and the quantile procedure with cone size dependence. As a concrete example, we compare dijet and boson+jet events to unfold the spectrum bias effects, and improve quark-, and gluon-jet classification using arguments based on the cumulative. Besides pointing out its flexibility, finally, we apply our framework to other energy loss models such as the hybrid weak/strong-coupling approach.

Keywords: Heavy Ion Phenomenology, Jets

ArXiv ePrint: 2103.14676

https://doi.org/10.1007/JHEP10(2021)038

JHEP10(2021)038

Contents

1 Introduction **1**

2 Quenching effects in the spectrum **3**
- 2.1 Constructing the quenching weight 5
- 2.2 The medium induced gluon spectrum at finite cone 5
- 2.3 Single-parton contribution to out-of-cone energy loss 9
- 2.4 Quenching the whole jet 11

3 Quenching effects in the cumulative spectrum **13**
- 3.1 The quantile ratio 14
- 3.2 Statistical advantage of the cumulative spectrum 16
- 3.3 Other types of observables based on the spectrum 17

4 Predictions **19**
- 4.1 Including the realistic spectrum and elastic energy loss 19
- 4.2 Numerical results for dijet events 20
- 4.3 Differences between dijets and Z+jets 23
  - 4.3.1 Difference in the R_{AA} 23
  - 4.3.2 Improving quark-gluon discrimination 24

5 Conclusion **26**

A Corrections to the Laplace transformation and to the running power **28**

B p_T-scaling of the out-of-cone emission quenching factor **29**

C Generating and parametrizing the jet spectrum **30**

D Other models **32**
- D.1 Momentum shift and elastic energy loss 32
- D.2 Hybrid weak-, strong-coupling 33

1 Introduction

At high-energy particle colliders, such as the LHC at CERN, the exchanged momentum is large enough to resolve and scatter fundamental partonic constituents of the matter. These violent scatterings deviate the partonic constituents of nucleons and allow for intense bremsstrahlung radiation that ultimately result in collimated bunches of hadronic particles and energy. These so-called jets open a new perspective on the understanding of strong interactions at intermediate energy scales between the scale of the hard partonic scattering

JHEP10(2021)038

and the hadronization scale. Jets can be calculated with high-precision within perturbative QCD both in electron-positron and proton-proton collisions [1–3].

For accelerated heavy nuclei, processes involving large momentum exchanges between the incoming partons have additionally the potential to be well-calibrated probes of the hot and dense quark-gluon plasma (QGP) that is formed in the aftermath of such violent collisions. In this context, jets are particularly interesting since their typical formation time-scales overlap with the time-scales governing the creation and evolution of the QGP, suggesting potentially substantial jet-medium interactions. While the strong separation of the medium scale and the jet scale motivates a perturbative description of hard jet-medium interactions, many aspects of this processes are in the realm of non-perturbative physics and have to be modeled on the phenomenological level. A satisfactory description of jet production in heavy-ion collisions would therefore allow us to separate perturbative from non-perturbative phenomena. Besides, a well-controlled scale separation is an indispensable insight when studying, e.g., jet-medium coupling, thermalization, or medium modifications of hadronization.

Jet studies have a rich history and a wide selection of observables have been discussed. This includes measurements of fully reconstructed jets and their substructure, for recent reviews see [1, 2, 4]. A fundamental observable is the momentum spectrum of jets for different reconstruction parameters R [5, 6]. The nuclear modification factor R_{AA}, compares the spectrum in heavy-ion collisions (AA) to proton-proton (pp) at the same reconstructed jet p_T. However, jets that interact with a surrounding medium lose energy and end up with smaller p_T. Therefore — and it is not emphasized enough — the jet selection for R_{AA} compares two jet populations originated at different p_T. The equal p_T selection induces a bias in the observables because the probability of creating a jet is steeply falling with p_T [7, 8]. This bias is explored in great detail below.

Recently, there has been efforts toward mitigating such bias effects by investigating novel observables or by using machine learning techniques [9]. One alternative, that we will investigate in detail in the current work, is to introduce a quantile procedure [10] to reconstruct a p_T that is closer to the initial jet p_T before quenching sets in. In contrast to R_{AA}, the quantile procedure uses the tail-cumulative of the jet spectrum and momentum ratio to reduce the bias coming from the steepness of the spectrum. We demonstrate the properties of the quantile procedure within a versatile framework to incorporate quenching effects and explain its robustness for the first time.

To improve the calibration of hard probes in heavy-ion collisions, new measurements have been suggested, e.g., involving boson+jet events [11–14]. Bosons suffer little modification in the medium, and their momenta are strongly correlated with the initiator of the recoiling jet. The jet spectrum in boson+jet events is slightly different from inclusive QCD jets, and we will use it to illustrate the bias on the quenching. Moreover, quark- and gluon-jet contributions in the inclusive and boson+jet samples differ. This can be used for quark-gluon jet discrimination in a model-independent fashion [15]. Using arguments on the cumulative, we improve the statistics of the classification task. Parallel with the works mentioned above, some numerical studies has also appeared using Bayesian and machine learning techniques to extract the energy loss properties from data [16].

JHEP10(2021)038

During the evolution of jets inside a QGP, their constituents scatter elastically and inelastically on the medium. The scatterings redistribute energy to larger angles out of the jet cone, resulting in energy loss. The inelastic (or radiative) part describes the emissions induced by the medium (medium-induced emissions, or MIE for short). The MIE has well-known limits: (i) in the multiple soft scattering limit, the scattering centers act coherently, resulting in suppression of emissions (QCD analog of the Landau-Pomeranchuk-Migdal effect), captured by the BDMPSZ formula, which describes the induced emission of soft gluons [17–19]; whil (ii) the single hard scattering limit is captured by the GLV formula, which describes emissions of harder gluons [20]. Beyond the analytic limits, MIE is also amenable to direct numerical methods, such as in refs. [21–23]. Recently, there has also been a progression in the better understanding of the two regimes and the scales involved [24–26]. In our work, we adopt the latter strategy to explore the impact of the MIE spectrum. The MIE can be resummed accounting for multiple induced emissions. The resulting formalism is the quenching weight [27, 28]. It is easy to generalize for all jet constituents, including coherence effects [29], spectrum shapes, and elastic energy loss. Within this framework, we show the appearance of the spectrum bias.

The paper is organized as follows. In section 2, we define our novel framework to calculate the jet spectrum in heavy-ion collisions. In section 3, we show, for the first time, the properties of the cumulative spectrum, and we apply the quantile procedure. In section 4 we give predictions for measurements, considering cone size dependence for the single-inclusive jet sample produced in dijet and boson+jet events (we focus concretely on Z+jet). We also show how to use cumulative arguments to improve quark-gluon discrimination. Beside,. in appendix D, we use the quenching weight formalism for elastic scatterings and in the hybrid weak/strong-coupling model.

2 Quenching effects in the spectrum

The main observable considered in this work is the single-inclusive spectrum of reconstructed jets in heavy-ion collisions. In this work, we will both consider jets produced in conventional QCD processes, that is dijet events, and jets produced in conjunction with a photon or weak boson, so-called boson-jet events. In the context of high-energy collisions, it is natural to assume a factorization of the partonic hard cross-section from the subsequent medium processes. This can be justified by invoking the large separation of momentum scales involved in jet production; typically the hard scattering $Q_{\rm hard} \sim 10^3\,\text{GeV}$, is much bigger than the jet scale $Q_{\rm jet} \sim p_T R \sim 10^2\,\text{GeV}$, where p_T is the reconstructed transverse momentum of the jet and R is jet cone parameter. These scales are much bigger than the typical medium scale, for instance, the temperature of the medium $T \sim 0.5\,\text{GeV}$. Hence, one can write the medium modification of the vacuum jet spectrum due to energy loss [27],

$$\frac{\mathrm{d}\sigma_R^{\rm med}}{\mathrm{d}p_T}(p_T) = \int_0^\infty \mathrm{d}\varepsilon\, \mathcal{P}_>(\varepsilon)\, \frac{\mathrm{d}\sigma_R^{\rm vac}}{\mathrm{d}p'_T}\bigg|_{p'_T=p_T+\varepsilon} . \tag{2.1}$$

The $\mathcal{P}_>(\varepsilon)$ describes the probability of a vacuum jet to distribute (or lose) energy out of the jet cone. Above, $\mathrm{d}\sigma_R^{\rm vac}/\mathrm{d}p_T$ refers to the partonic cross-section to produce a jet with

JHEP10(2021)038

R in the collinear factorization and can be calculated up to high precision [5, 6]. In this work, instead, we extract the spectrum using a Monte Carlo event generator, see section 4 for further details.

The vacuum spectrum is well approximated by a power-law, $\mathrm{d}\sigma_R^{\mathrm{vac}}/\mathrm{d}p_T \propto 1/p_T^n$, and is steeply falling, i.e. $n \gg 1$. Therefore, we approximate $\mathrm{d}\sigma_R^{\mathrm{vac}}(p_T+\varepsilon)/\mathrm{d}p_T = A\,(p_T+\varepsilon)^{-n} \approx \mathrm{e}^{-n\varepsilon/p_T}\mathrm{d}\sigma_{\mathrm{vac}}(p_T)/\mathrm{d}p_T$, where A is a constant and $n \equiv n(p_T, R)$ is the power index of the spectrum. In the last step, we additionally assumed that $\varepsilon \ll p_T$. The cone-size dependent nuclear modification factor defined as

$$R_{\mathrm{med}}(p_T, R) \equiv \frac{\mathrm{d}\sigma_R^{\mathrm{med}}}{\mathrm{d}p_T} \Big/ \frac{\mathrm{d}\sigma_R^{\mathrm{vac}}}{\mathrm{d}p_T} \,, \tag{2.2}$$

is consequently related to the quenching factor $\mathcal{Q}(\nu \equiv n/p_T)$, which is the Laplace transform of $\mathcal{P}_>(\varepsilon)$, i.e.

$$R_{\mathrm{med}}(p_T, R) \approx \int_0^\infty \mathrm{d}\varepsilon\, \mathcal{P}_>(\varepsilon)\, \mathrm{e}^{-\frac{n\varepsilon}{p_T}} \equiv \mathcal{Q}_>(\nu)\,. \tag{2.3}$$

This approximation is precise within at most a few percent for realistic parameters in a wide kinematic range $10 < p_T < 1000\,\mathrm{GeV}$ that we consider here. Therefore, in what follows, we simply identify $R_{\mathrm{med}} = \mathcal{Q}_>$. For further details on such corrections, see appendix A. In order to go from R_{med} to the experimentally measured R_{AA}, one needs to include a p_T-dependent quark and gluon jet production. Other differences are mostly due to geometry (for a review see ref. [30]), and cold nuclear effects (nPDF), that we include in section 4.

The energy loss distribution $\mathcal{P}_>(\varepsilon)$ — and therefore the quenching factor $\mathcal{Q}_>(\nu)$ — depends on vacuum jet properties (like the jet p_T and the cone size R) and the properties of the medium (e.g., medium length L, and the jet transport coefficient $\hat{q}$). Its normalization condition, $\int_0^\infty \mathrm{d}\varepsilon\, \mathcal{P}_>(\varepsilon) = 1$, translates to $\mathcal{Q}_>(0) = 1$. The assumptions underlying eq. (2.2) are quite robust for a wide range of applications. Therefore, one is flexible in defining the precise nature of the energy loss distribution $\mathcal{P}_>(\varepsilon)$. The introduction of a probability distribution to describe effects of quenching in heavy-ion collisions has a long history in the analysis of single-inclusive hadron [27, 28, 31] and jet spectra [29, 32–34].

We will derive the quenching factor $\mathcal{Q}_>$ of the jet in several steps. First, we consider the induced radiation spectrum of a single color-charge propagating through the medium and how to account for their multiple emissions. We thus arrive at the quenching factor for a single parton, $\mathcal{Q}_>^{(0)}(\nu)$. Next, we consider the effect of jet fragmentation which leads to multiple vacuum-like emissions on short time-scales inside the jets. Partons from these emissions contribute to the quenching of the full jet. This is accounted for by the so-called collimator function that provides a fully resummed quenching factor $\mathcal{Q}_>(\nu)$. The corresponding jet quenching probability distribution can then be found via an inverse Laplace transform, but we will not pursue this further in this work.

In our numerical results in section 4, we will also include elastic energy loss, since it potentially can contribute to the $\sim 10-30\%$ level to the final jet suppression factor. As we said, the formulation above is quite general and allows to separately formulate `a)` the mechanism of quenching, and `b)` the phase space where the jet is affected. In appendix D, therefore, we show how to formulate other energy loss models in terms of quenching weights.

JHEP10(2021)038

2.1 Constructing the quenching weight

We construct the radiative energy loss distribution $\mathcal{P}_>(\varepsilon)$ (or the quenching factor $\mathcal{Q}_>(p_T)$) of a jet starting from a single parton, that we denote $\mathcal{P}_>^{(0)}(\varepsilon)$. For the radiation of a single medium-induced gluon, this energy loss probability is simply given by

$$\mathcal{P}_>^{(0)}(\varepsilon) \approx \left.\frac{\mathrm{d}I_>}{\mathrm{d}\omega}\right|_{\omega=\varepsilon} + \delta(\varepsilon)\left[1 - \int_0^\infty \mathrm{d}\omega\, \frac{\mathrm{d}I_>}{\mathrm{d}\omega}\right], \tag{2.4}$$

where $\mathrm{d}I_>/\mathrm{d}\omega$ is the spectrum of medium-induced gluon radiation spectrum that emerges at angles larger than the jet cone, or

$$\frac{\mathrm{d}I_>}{\mathrm{d}\omega} = \int_{(\omega R)^2}^\infty \mathrm{d}\boldsymbol{k}^2 \frac{\mathrm{d}I}{\mathrm{d}\omega\,\mathrm{d}\boldsymbol{k}^2}\,. \tag{2.5}$$

This is an important difference with respect to the more common use of the quenching weights applied to single-hadron spectra in refs. [27, 28]. The first term in eq. (2.4) describes an emission, while the second term is a virtual correction and provides the normalization. Our starting point in section 2.2 is to discuss the specific details of this induced-emission spectrum. Accounting for multiple such emissions in course of the medium propagation allows us to derive the single parton quenching weight in section 2.3.

Then, having derived how *one parton* contributes to the energy loss of the whole jet, we next turn to the calculation of how *multiple partons* in the jet, resolved by the medium during their fragmentation process, add up to the total quenching effect. This will be described in section 2.4.

2.2 The medium induced gluon spectrum at finite cone

The medium-induced spectrum from multiple scattering in a QCD medium was derived independently by Zakharov [18, 35] and Baier-Dokshitzer-Mueller-Peigne-Schiff (BDMPS) [17, 19, 36], see also refs. [37, 38]. In the limit of soft gluon emission with ω energy and $\boldsymbol{k}$ transverse momentum, this spectrum can be written as

$$\begin{aligned}\omega\frac{\mathrm{d}I}{\mathrm{d}\omega\mathrm{d}^2\boldsymbol{k}} = \frac{\alpha_s C_R}{(2\pi)^2\omega^2} 2\mathrm{Re}\int_0^\infty \mathrm{d}t_1 \int_{t_1}^\infty \mathrm{d}t_2 \int \mathrm{d}^2\boldsymbol{x}\, \mathrm{e}^{-\int_{t_2}^\infty \mathrm{d}s\, v(\boldsymbol{x},s)} \\ \times\, \partial_{\boldsymbol{x}}\cdot\partial_{\boldsymbol{y}}\mathcal{K}(\boldsymbol{x},t_2;\boldsymbol{y},t_1)\big|_{\boldsymbol{y}=0}\,,\end{aligned} \tag{2.6}$$

where $C_R = C_F$ if the radiator is a quark (C_A for gluon), the path integral is

$$\mathcal{K}(t_2,\boldsymbol{x};t_1,\boldsymbol{y}) = \int_{\boldsymbol{r}(t_1)=\boldsymbol{y}}^{\boldsymbol{r}(t_2)=\boldsymbol{x}} \mathcal{D}\boldsymbol{r}\, \exp\left\{\int_{t_1}^{t_2}\mathrm{d}s\, \left[i\frac{\omega}{2}\dot{\boldsymbol{r}}^2 - v(\boldsymbol{r},s)\right]\right\}, \tag{2.7}$$

and

$$v(\boldsymbol{x},t) = N_c \int \frac{\mathrm{d}^2\boldsymbol{q}}{(2\pi)^2}\frac{\mathrm{d}^2\sigma_{\mathrm{el}}}{\mathrm{d}^2\boldsymbol{q}}\left(1-\mathrm{e}^{i\boldsymbol{x}\cdot\boldsymbol{q}}\right), \tag{2.8}$$

is related to the elastic scattering potential in medium. The leading logarithmic behavior of the scattering potential reads for any hard Coulomb tailed elastic potential,

$$v(\boldsymbol{x},t)_{\mathrm{LL}} = \frac{1}{4}\hat{q}_0(t)\boldsymbol{x}^2\log\left(\frac{1}{\boldsymbol{x}^2\mu_*^2}\right) + \mathcal{O}\left(\boldsymbol{x}^4\mu_*^2\right), \tag{2.9}$$

JHEP10(2021)038

where $\hat{q}_0(t)$ is a local transport coefficient and μ_* is related to the infrared scale that regularizes the in-medium scattering potential $\mathrm{d}^2\sigma_{\rm el}/\mathrm{d}^2\boldsymbol{q}$. For example, the Gyulassy-Wang scattering potential [39], describes scattering off a plasma made up of Yukawa-screened quasi-particles, reads

$$\left.\frac{\mathrm{d}^2\sigma_{\rm el}}{\mathrm{d}^2\boldsymbol{q}}\right|_{\rm GW} = \frac{g^4 n(t)}{(\boldsymbol{q}^2+\mu^2)^2}\,, \tag{2.10}$$

where μ is a screening mass, and $n(t)$ is the density of scattering centers in the medium. In this case $\hat{q}_0(t) = 4\pi\alpha_s^2 N_c n(t)$, and $\mu_*^2 = \mu^2 \mathrm{e}^{-1+2\gamma_E}/4$.

The path integral in eq. (2.7) can be solved numerically, see e.g. in refs. [21–23], but analytic solutions are available in limiting cases. Here, two of the most frequent schemes are discussed. In the "harmonic oscillator" (HO) approximation, valid when the logarithm is slowly varying around a constant scale, i.e. $\mu_*^2 \ll 1/\boldsymbol{x}^2 \sim Q^2$, one can absorb the logarithm in the definition of the transport coefficient so that $v(\boldsymbol{x}) \approx \hat{q}\boldsymbol{x}^2/4$. This corresponds to a purely Gaussian momentum broadening in transverse momentum given by $\langle \boldsymbol{k}^2\rangle = \hat{q}t$ during the propagation in the medium. In this case, eq. (2.7) describes a harmonic oscillator with imaginary frequency $\Omega^2 = \hat{q}/(2i\omega)$ in transverse to the propagation, and whose solution is well known. This approximation, however, fails to describe the hard tail of medium emissions. On the other hand, one can also truncate the resummation of medium scatterings at a fixed order N, giving rise to the so-called "opacity expansion" [20, 40]. The $N=1$ approximation amounts to considering a single, incoherent scattering with the medium during the propagation.

A discussion of the regions of validity of these approaches was recently addressed in refs. [24–26, 41]. In the soft scattering regime, the formation time of emissions $t_{\rm f} = 2\omega/\boldsymbol{k}^2$ becomes modified due to Gaussian broadening, i.e. $t_{\rm f} \sim \sqrt{2\omega/\hat{q}}$. For emissions with large formation times $t_{\rm f} \sim L$ correspond to $\omega \sim \hat{q}L^2/2 \equiv \omega_c$. It follows that emissions with $\omega > \omega_c$ cannot be produced by soft collisions and Gaussian broadening, and are dominated by a single, hard scattering with the medium constituents. Similarly, at short formation times of the order of the medium mean free path $t \sim \lambda \ll L$, or $\omega \sim \hat{q}\lambda^2/2 = \omega_{\rm BH} \ll \omega_c$, the spectrum is again dominated by single scattering [40, 41]. Since this latter regime gives a small contribution to energy loss, we will not discuss it further here.

Recently, the contribution of hard emissions has been shown to matter for precision comparisons with high-p_T single-hadron spectra at RHIC and LHC [42]. A systematic procedure to calculate the spectrum for a large range of relevant emission energies $\omega > \omega_{\rm BH}$ was developed in the so-called "improved opacity expansion" (IOE) [24–26]. This framework rewrites the leading-log scattering potential form eq. (2.9) as

$$v_{\rm LL}(\boldsymbol{x},t) = \frac{1}{4}\hat{q}_0\boldsymbol{x}^2\left[\log\left(\frac{Q_c^2}{\mu_*^2}\right) + \log\left(\frac{1}{Q_c^2\boldsymbol{x}^2}\right)\right] = v_{\rm HO}(\boldsymbol{x},t) + \delta v(\boldsymbol{x},t)\,, \tag{2.11}$$

where Q_c is a separation scale of the harmonic potential. In the limit of $Q_c \gg \mu_*$, one can then expand the solution of the path integral in eq. (2.7) around the HO solution with an effective $\hat{q}(Q_c^2)$ and treat hard scatterings with the medium, given by $\delta v(\boldsymbol{x},s)$, as higher-order perturbations. This approach is systematically improvable and, up to next-

JHEP10(2021)038

to-next-to-leading order in this expansion, the effective $\hat{q}$ parameter is given by [26]

$$\hat{q}(Q_c^2) = \hat{q}_0 \ln\left(\frac{Q_c^2}{\mu_*^2}\right)\left[1 + \frac{1.013}{\ln\left(\frac{Q_c^2}{\mu_*^2}\right)} + \frac{0.318}{\ln^2\left(\frac{Q_c^2}{\mu_*^2}\right)}\right], \tag{2.12}$$

where the scale Q_c^2 is itself found through an implicit equation, $Q_c^4 = \hat{q}_0\omega\ln(Q_c^2/\mu_*^2)$. This equation has solution, if $\omega \geq \omega_{\min} \equiv 2e\mu_\star^4/\hat{q}_0$, when $Q_c^2 = \mu_\star^2 \exp[-W_{-1}(2\mu_\star^4/(w\hat{q}_0))/2]$, where $W_i(x)$ is the Lambert function on the i^{th} branch. Since jet quenching is not sensitive to the details of very soft gluon emissions, in our numerical results we freeze the logarithms at 1, i.e. $\ln Q_c^2/\mu_*^2 \geq 1$.

The IOE has so far only been developed for the spectrum integrated over transverse momenta, i.e. $\mathrm{d}I/\mathrm{d}\omega$, and for the momentum broadening of a single particle in the medium [43]. Since the contributions to jet energy loss rely on out-of-cone emissions, cf. eq. (2.5), we instead have to consider the matching of the partially integrated spectrum $\mathrm{d}I_>/\mathrm{d}\omega$ between the multiple-scattering HO and single-scattering $N = 1$ regimes. We use the quenching parameter $\hat{q}$ from eq. (2.12) which correctly connects the $\hat{q}_0$ parameter from the elastic potential with the multiple scattering formalism. We propose a simple interpolation scheme that relies on a single matching scale $\omega_\star$, that is related to the broadening of soft gluons and will be defined below. At small gluon energies, below the matching scale, the spectrum is described by the HO approach with the effective $\hat{q}$. Above the matching scale, the spectrum is given by the $N = 1$ spectrum.

We will treat the medium as a "brick" of constant $\hat{q}_0$ and fixed length L. In the absence of the cone constraint, a natural matching scale is $\omega_\star \sim \omega_c \equiv \hat{q}L^2/2$. For the out-of-cone spectrum, the effect of broadening after emission cannot be neglected. For Gaussian broadening that presents in HO, a particle emitted at initial time accumulates $\langle \boldsymbol{k}^2\rangle \sim \hat{q}L$ after propagating through the medium. This corresponds to an angle $\theta \sim \sqrt{\hat{q}L}/\omega$ in the small-angle approximation. Demanding that this angle is larger than the jet cone R for energy loss, a cut-off in energy arises $\omega < \omega_R \equiv \sqrt{\hat{q}L}/R$, above which the HO spectrum falls rapidly. For more details, see appendix B. It turns out that a relatively smooth matching between the HO and $N = 1$ regimes is achieved by choosing

$$\omega_\star = \min\left(\omega_c, \omega_R\right), \tag{2.13}$$

where ω_c and ω_R are defined with the effective $\hat{q}$ parameter in eq. (2.12). For our final results, see the left and right panels in figure 1. The postulated matching works extremely well, up to some negligible discontinuities in the spectrum. The uncertainty in this matching procedure is small compare to other approximation that we will make in the following sections.

The regime of soft gluon emissions, $\omega < \omega_\star$, is dominated by multiple scattering where we can employ the HO approximation. The spectrum in this approximation is given by

$$\omega\frac{\mathrm{d}I^{\text{HO}}}{\mathrm{d}\omega\,\mathrm{d}^2\boldsymbol{k}} = \frac{\bar{\alpha}}{\pi}\,\mathrm{Im}\Big[\mathcal{R}^{\text{in-in}} + \mathcal{R}^{\text{in-out}}\Big], \tag{2.14}$$

JHEP10(2021)038

where $\bar{\alpha} = \alpha_s C_R/\pi$ and the two factors read [44]

$$\mathcal{R}^{\text{in-in}} = \int_0^L \mathrm{d}t\, \frac{(1+i)\sqrt{\omega\hat{q}}\cot(\Omega t)}{\hat{q}(L-t)-(1+i)\sqrt{\omega\hat{q}}\cot(\Omega t)} \exp\left[-\frac{\boldsymbol{k}^2}{\hat{q}(L-t)-(1+i)\sqrt{\omega\hat{q}}\cot(\Omega t)}\right], \tag{2.15}$$

$$\mathcal{R}^{\text{in-out}} = \int_0^L \mathrm{d}t\, \frac{1}{\cos^2(\Omega t)} \exp\left[-i\frac{\boldsymbol{k}^2}{2\omega\Omega}\tan(\Omega t)\right], \tag{2.16}$$

where $\Omega = (1-i)\sqrt{\hat{q}/\omega}/2$. The integrated spectrum in eq. (2.5) reads then

$$\omega\frac{\mathrm{d}I^{\text{HO}}_{>}}{\mathrm{d}\omega} = \frac{\bar{\alpha}}{\omega}\,\mathrm{Im}\left[\mathcal{R}^{\text{in-in}}_{>} + \mathcal{R}^{\text{in-out}}_{>}\right], \tag{2.17}$$

where now

$$\mathcal{R}^{\text{in-in}}_{>} = \int_0^L \mathrm{d}t\,(1+i)\sqrt{\omega\hat{q}}\cot(\Omega t)\exp\left[-\frac{(\omega R)^2}{\hat{q}(L-t)-(1+i)\sqrt{\omega\hat{q}}\cot(\Omega t)}\right], \tag{2.18}$$

$$\mathcal{R}^{\text{in-out}}_{>} = -\int_0^L \mathrm{d}t\,\frac{(1+i)\sqrt{\hat{q}\omega}}{\cos(\Omega t)\,\sin(\Omega t)}\exp\left[-i\frac{(\omega R)^2}{2\omega\Omega}\tan(\Omega t)\right], \tag{2.19}$$

We find that in the $R\to 0$ limit, eq. (2.17) yields

$$\omega\frac{\mathrm{d}I^{\text{HO}}}{\mathrm{d}\omega} = 2\bar{\alpha}\,\ln\left|\cos(1-i)\sqrt{\frac{\omega_c}{2\omega}}\right|. \tag{2.20}$$

which is the celebrated BDMPS-Z spectrum [18, 19].

As discussed above, at $\omega > \omega_\star$, the HO spectrum has to be corrected with the single hard gluon emission spectrum ($N=1$) [24–26], for which

$$\omega\frac{\mathrm{d}I^{N=1}}{\mathrm{d}\omega\mathrm{d}^2\boldsymbol{k}} = 8\pi\bar{\alpha}N_c\int_0^L \mathrm{d}t\int\frac{\mathrm{d}^2\boldsymbol{q}}{(2\pi)^2}\,\frac{\mathrm{d}^2\sigma_{\text{el}}}{\mathrm{d}^2\boldsymbol{q}}\,\frac{\boldsymbol{k}\cdot\boldsymbol{q}}{\boldsymbol{k}^2(\boldsymbol{k}-\boldsymbol{q})^2}\left[1-\cos\frac{(\boldsymbol{k}-\boldsymbol{q})^2}{2\omega}t\right]. \tag{2.21}$$

Using eq. (2.5) and the Gyulassy-Wang potential from eq. (2.10), we immediately find that the integrated spectrum reads

$$\omega\frac{\mathrm{d}I^{N=1}_{>}}{\mathrm{d}\omega} = \bar{\alpha}\frac{\hat{q}_0 L^2}{\omega}\int_0^\infty \mathrm{d}u\,\frac{u-\sin u}{u^2}\,\frac{1}{[(\zeta+u+y)^2-4\zeta u]^{1/2}}, \tag{2.22}$$

where $\zeta \equiv \omega R^2 L/2$ and $y = \mu^2 L/(2\omega)$. Again, for $R\to 0$, we recover the familiar form of the integrated $N=1$ spectrum [20, 28],

$$\omega\frac{\mathrm{d}I^{N=1}}{\mathrm{d}\omega} = \bar{\alpha}\frac{\hat{q}_0 L^2}{\omega}\int_0^\infty \mathrm{d}u\,\frac{u-\sin u}{u^2}\,\frac{1}{u+y}, \tag{2.23}$$

as expected.

The full spectrum is therefore postulated to be well approximated by the following interpolation,

$$\frac{\mathrm{d}I_{>}}{\mathrm{d}\omega} = \Theta(\omega_\star-\omega)\frac{\mathrm{d}I^{\text{HO}}_{>}}{\mathrm{d}\omega} + \Theta(\omega-\omega_\star)\frac{\mathrm{d}I^{N=1}_{>}}{\mathrm{d}\omega}, \tag{2.24}$$

JHEP10(2021)038

Medium parameter	Value
α_s	0.3
T_0	0.45 GeV
$\hat{q}_0$	0.095 GeV^3
L	4 fm

Table 1. Choice of medium parameters corresponding to 0-10% central PbPb collisions with $\sqrt{s_{NN}} = 5.02\,\text{TeV}$.

where all parameters in the HO spectrum and in the matching scales contain the effective $\hat{q}$, given in eq. (2.12). Note that the spectrum $\mathrm{d}I_>/\mathrm{d}\omega$ depends on the initial energy only in the combination $\omega = x(1-x)E \approx xE$.

At this point, we pause to discuss the choice of medium parameters. We choose the medium coupling to be fixed at $\alpha_s = \alpha_{\text{med}} = 0.3$, corresponding to $g_{\text{med}} \approx 1.94$. The IR cutoff scale is $\mu \equiv \sqrt{2/3} g_{\text{med}} T_0 \approx 0.71\,\text{GeV}$, and $\omega_{\text{min}} \approx 1.5\,\text{GeV}$. The remaining parameters are chosen to reflect the conditions in 0–10% central PbPb events at $\sqrt{s_{NN}} = 5.02\,\text{TeV}$, see table 1. As demonstrated below, this parameter set gives a good description of the experimentally measured jet suppression factor, see section 4. Since our work does not deal with the precise description of experimental data, we have not attempted to fix these parameters from a model of the underlying medium nor fitted them to experimental data, which was done in [34]. This choice finally leads to a matching scale $Q_c^2 = (1-100)\mu_\star^2$ and $\hat{q} = (1-5)\hat{q}_0$, depending on ω, in the IOE-matched spectrum eq. (2.24).

The MIE spectrum is shown in figure 1 for quarks (left) and for gluons (right). The matching points ω_* from eq. (2.13) are shown with bullets and below (above) the spectrum is the HO ($N = 1$) spectrum. The matching works very well capturing the cone size dependence, however it is not perfectly smooth (see $R = 0$). It is good enough to study the integral of this spectrum, presented in the quenching factor. With different colors the cone size dependence points out, it is less probable to lose energy by opening the cone (i.e., recapturing emissions). The difference in the quark and gluon spectrum is the color factor $C_A/C_F = 9/4$, and thus gluons lose more energy. The dotted line in figure 1, represents the energy scale which below secondary branching start to dominate [37], corresponding to $\omega \sim \alpha_s^2 \hat{q} L^2$, see eq. (2.29) (for more details, see the next subsection). Finally, the grey band in figure 1 corresponds to emissions with $\omega \sim \omega_{\text{BH}}$, which are given by the Bethe-Heitler spectrum [40, 41]. In what follows, we will neglect such emissions since these emissions do not contribute significantly to jet energy loss at high-p_T [27, 28].

2.3 Single-parton contribution to out-of-cone energy loss

When the number of medium-induced gluon emissions becomes large, $\int_\omega^\infty \mathrm{d}\omega\, \mathrm{d}I/\mathrm{d}\omega > 1$, one needs to go beyond (2.4) and account for multiple emissions to the energy loss distribution. Assuming independent emissions, we can treat it as a Poisson process. This allows to define a probability distribution of radiating energy ε off a single parton at angles

JHEP10(2021)038

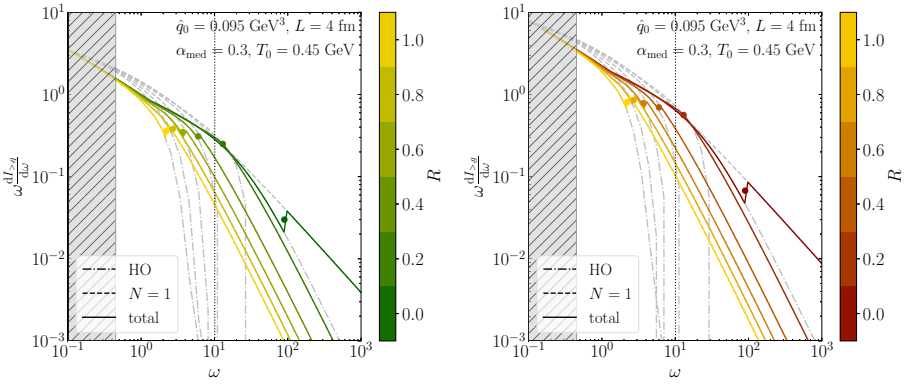


Figure 1. The out of cone emission spectrum, matched at NNLO for quarks (left) and gluons (right) from eq. (2.24) for different cone sizes. The dash-dotted (dashed) grey lines are the HO ($N = 1$) spectra, the bullets show the matching points ω_*. The dotted line shows the region ω_s below in which emitted gluons thermalize during the broadening. The hatched band shows where $\omega < T_0$, considered and neglected as background.

larger than the jet cone. This probability distribution reads

$$\mathcal{P}_{>}^{(0)}(\varepsilon) = \sum_{n=0}^{\infty} \frac{1}{n!} \left[\prod_{j=1}^{n} \int \mathrm{d}\omega_j \frac{\mathrm{d}I_{>}}{\mathrm{d}\omega_j} \right] \delta\left(\varepsilon - \sum_{j=1}^{n} \omega_j \right) \exp\left[- \int \mathrm{d}\omega \frac{\mathrm{d}I_{>}}{\mathrm{d}\omega} \right], \tag{2.25}$$

where the parton radiates n soft gluons with ω_j energies summing up to ε, and there is a Sudakov exponential factor to resum virtual contributions [27]. In eq. (2.25), we only account for the primary emissions off the leading particle and neglect any secondary splittings. Performing the Laplace transform, the quenching factor of a single parton emitting multiple gluons can be calculated using

$$\mathcal{Q}_{>}^{(0)}(\nu) \equiv \int_0^{\infty} \mathrm{d}\varepsilon\, \mathcal{P}_{>}^{(0)}(\varepsilon)\, \mathrm{e}^{-\nu\varepsilon} = \exp\left[- \int_0^{\infty} \mathrm{d}\omega \frac{\mathrm{d}I_{>}}{\mathrm{d}\omega} \left(1 - \mathrm{e}^{-\nu\omega}\right) \right], \tag{2.26}$$

where $\nu = n/p_T$.

It is here worth emphasizing the role of the hard emissions described by the $N = 1$ spectrum. Neglecting for the moment broadening, i.e. setting $R = 0$ in eq. (2.22), the spectrum at large $\omega > \omega_c \gg \mu^2 L/2$ is simply

$$\left. \frac{\mathrm{d}I^{N=1}}{\mathrm{d}\omega} \right|_{\omega > \omega_c} = \frac{\bar{\alpha}\pi}{4} \frac{\hat{q}_0 L^2}{\omega^2}\,. \tag{2.27}$$

The resulting single-parton quenching factor for this regime behaves as

$$\mathcal{Q}_{>}^{(0),N=1}(\nu) = \exp\left[-\frac{\bar{\alpha}\pi}{2} \frac{\hat{q}_0}{\hat{q}} \left(1 - \mathrm{e}^{-\nu\omega_c} + \nu\omega_c \Gamma(0, \nu\omega_c)\right) \right], \tag{2.28}$$

where $\Gamma(s,x) = \int_x^{\infty} \mathrm{d}t\, t^{s-1} \mathrm{e}^{-t}$ is the upper incomplete gamma function. At low p_T, i.e. $p_T \ll n\omega_c$, the quenching becomes at most $\mathcal{Q}^{(0),N=1}(\nu)|_{p_T \ll n\omega_c} \approx 1 - \alpha_s C_R \hat{q}_0/(2\hat{q})$ constant

JHEP10(2021)038

factor. Hence, we conclude that the impact of hard radiation at LHC, where we expect $\hat{q} > \hat{q}_0$, is relatively small. However, at high-p_T, i.e. $p_T \gg n\omega_c$, where the leading behavior is $Q^{(0)}(\nu) \approx Q^{(0),N=1}(\nu)|_{p_T \gg n\omega_c} \approx 1 - \frac{\alpha_s}{4} C_R \hat{q}_0 L^2 \nu (1 - \gamma_E - \ln \nu\omega_c)$, it becomes more important. Also, since hard radiation takes place at small angles, the effect is even smaller at $R > 0$.

Equation (2.26) also includes contributions from small energy gluons, and an important modification should be included to improve their description in the medium. As mentioned above, gluons emitted with energy $\omega_{\rm BH} < \omega < \omega_s$, where

$$\omega_s \equiv \pi \left(\frac{N_c \alpha_s}{\pi} \right)^2 \hat{q}_0 L^2 \,, \tag{2.29}$$

will thermalize quasi-instantaneously in the plasma via multiple branching [37]. Their energy will basically be redistributed randomly over a cone with characteristic opening angle $R_{\rm rec} \sim \pi/2$. Hence, instead of losing energy ω out of the cone, the jet loses $\omega[1 - (R/R_{\rm rec})^2]$, where the second power comes from the area proportionality. This process describes the thermalization of soft jet particles. Moreover, if $R_{\rm rec}(\omega, \eta, \phi)$, where (η, ϕ) describes the jet direction with respect to the reaction plane, one could use it to describe back-propagation of the thermalized energy to the cone, and thus medium response. The dashed line in figure 1 shows the location of $\omega_s \approx 8.5\,\text{GeV}$ for our parameters. Depending on the medium and jet parameters, this scale can be below or above the matching scale $\omega_\star$. In ref. [34], the importance of $R_{\rm rec}$ was studied and small dependence in the result was observed and thus we used $R_{\rm rec} = \pi/2$.

Finally, after neglecting emissions below the Bethe-Heitler energy (we assume that it is given by the plasma temperature $\omega_{\rm BH} = T_0$), our final form for the quenching factor of a single parton inside the jet is therefore

$$\mathcal{Q}^{(0)}_{>}(\nu) = \exp\left[-\int_{T_0}^{\infty} \mathrm{d}\omega \frac{\mathrm{d}I_{>}}{\mathrm{d}\omega} \left(1 - \mathrm{e}^{-\nu\omega\left(1 - \Theta(\omega_s - \omega)\frac{R^2}{R^2_{\rm rec}}\right)} \right) \right] , \tag{2.30}$$

where the cone size dependence is implicit in the integration limits of the out-of-cone spectrum. The single parton quenching factor eq. (2.30) is shown in figure 2 with dashed lines for quark and gluon initiators. By opening the cone, the emitted energy gets gradually recovered, and thus the quenching factor becomes closer to 1. The difference in between quark and gluon initiators is $\mathcal{Q}^{(0)}_{>,g} = (\mathcal{Q}^{(0)}_{>,q})^{C_A/C_F}$, resulting more quenching for gluon. We expect our description to be less valid at smaller energies.

2.4 Quenching the whole jet

Having derived how *one parton* contributes to the energy loss of the whole jet, we next turn to the calculation of how *multiple partons* in the jet, resolved by the medium during their fragmentation process, add up to the total quenching effect.

Due to the large phase space for radiation between the jet scale $\sim p_T R$ and the hadronization scale $\sim \Lambda_{\rm QCD}$, the jet forms through multiple emissions. It can be estimated, from formation time arguments, that many of these emissions occur while the parton(s) are interacting with the surrounding medium [29]. In the limit of complete decoherence,

JHEP10(2021)038

the constituents are expected to lose energy independently, following an incoherent superposition of single-particle quenching factors in eq. (2.30). However, one has account for coherence effects leading to a finite resolution power of the medium. As long as two partons are closer to each other than the medium resolution length, the medium cannot resolve them individually. The two partons are affected coherently (as a whole color charge) by medium interactions and, in particular, by induced energy loss [45]. The relevant time-scale can be estimated by comparing the size of a dipole, that in the small-angle approximation roughly scales as $r_\perp \sim \theta t$ (where θ is the angle of the dipole and t is the propagation time), to the resolution length of the medium, that scales as $\lambda_\perp \sim 1/\langle k_\perp^2 \rangle^{1/2} \sim 1/\sqrt{\hat{q}t}$, where we assume Gaussian transverse-momentum broadening. The two transverse sizes are equal at the decoherence time $t_\mathrm{d} \sim (\hat{q}\theta^2)^{-1/3}$.

This condition can be translated to emission times: emissions with formation times smaller than the medium decoherence time $t_\mathrm{f} < t_\mathrm{d}$ are vacuum-like. In other words, they are generated according to the probability distribution to split in the vacuum. The core constituents should ultimately be resolved while they are still in the medium, i.e. $t_\mathrm{d} < L$. Therefore, jet constituents produced in the phase space delimited by $t_\mathrm{f} \ll t_\mathrm{d} \ll L$, will be resolved by the medium and are affected by quenching. The rest of the phase space stays unaffected.

These two effects (vacuum fragmentation and medium resolution) are captured by the collimator function $\mathcal{C}(p_T, R)$ [29], which is a function of the jet and medium scales. It takes into account the additional energy loss of resolved vacuum-like emissions in the medium (see also ref. [46] for an application to heavy-quark jets). The total quenching of the jet is therefore given as a product of the quenching of the total charge of the jet and the collimator, that is

$$\mathcal{Q}_{>,i}(p_T, R) = \mathcal{Q}^{(0)}_{>,i}\left(\frac{n}{p_T}\right) \mathcal{C}_i(p_T, R)\,, \tag{2.31}$$

where $i = q, g$ indicates the dependence on the color charge. This is what we refer to as the fully resummed quenching factor of a jet. The functions $\mathcal{C}_i$ obey a set of coupled, non-linear evolution equations, see in ref. [29]. Here, we use its linear approximation, where the quark and gluon solutions decouple, resulting in the resummation of all primary emissions off the initiator. This allows to write the solution explicitly as

$$\mathcal{C}_i(p_T, R) = \exp\left[-\int_0^R \frac{\mathrm{d}\theta}{\theta}\int_0^1 \mathrm{d}z\, \frac{\alpha_s(k_\perp)}{\pi} P_{gi}(z)\Theta_\mathrm{res}\left(\mathcal{Q}^{(0)}_{>,g}(n/p_T) - 1\right)\right], \tag{2.32}$$

where α_s is the 1-loop running coupling, the relative transverse momentum is $k_\perp = z(1-z)p_T\theta$, and $P_{gi}(z)$ is the Altarelli-Parisi LO splitting function. The finiteness of the integrals is ensured by the phase space measure $\Theta_\mathrm{res} = \Theta(L - t_\mathrm{d})\Theta(t_\mathrm{d} - t_\mathrm{f})$, with the corresponding times $t_\mathrm{f} = 2z(1-z)p_T/k_\perp^2$ and $t_\mathrm{d} = [12/(\hat{q}_0\theta^2)]^{1/3}$. At large p_T, $p_T \gtrsim \hat{q}_0 L^2$, this implies that the angular integral is directly regulated by $\theta > \theta_c$, where $\theta_c = [12/(\hat{q}_0 L^3)]^{1/2}$. In the opposite case, $p_T \lesssim \hat{q}_0 L^2$, the angular integral is regulated by $\theta \gtrsim (\hat{q}/p_T^3)^{1/4}$. Finally, if $\theta_c > R$ the jet is completely coherent and $\mathcal{C}(p_T, R) = 1$. The color dependence of the collimator is not trivial. Color dependence appears in the splitting function P_{gi} and bare quenching weight $\mathcal{Q}^{(0)}_{>,i}$, through the Casimir factors C_R, and, most complicatedly, in

JHEP10(2021)038

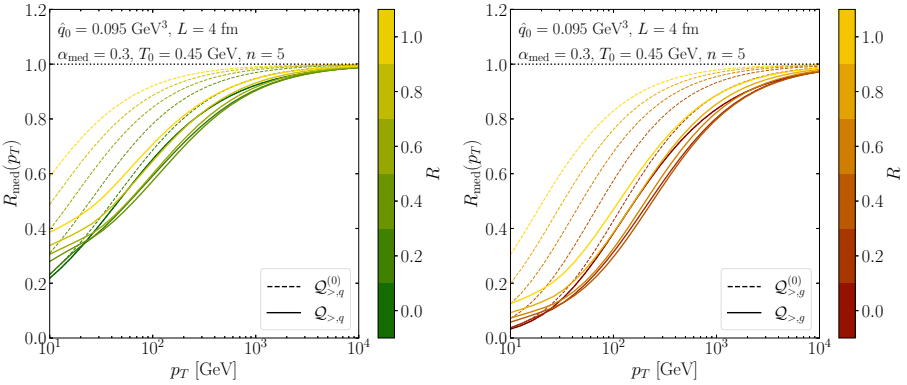


Figure 2. The nuclear modification factor using the matched radiation spectrum for single parton quenching form eq. (2.30) (dashed line), and quenching the whole jet with coherence effects from eq. (2.31) (solid line). Quark jets on the left and gluon jets on the right panel with different cone sizes.

the spectrum power indices $n_i(p_T)$. Therefore, the exact quark/gluon dependence of the collimator is process and nPDF dependent.

This linearized version of the collimator function is analytically calculable, which is a big advantage in comparison to the full, non-linear version. We also tested against the full non-linear solution, which resulted in small, <10%, deviations even for big $R \sim 1$ cones.

The resummed quenching weight $Q_{>,i}(p_T, R)$ is shown in figure 2 with solid lines for quarks (left) and gluons (right) for different cone sizes. It results in more quenching (with the same medium parameters) compared to the single-parton quenching factor because there are more jet constituents that contribute to the total energy loss. The R-dependence is a result of two competing effects: a) the recapture of medium-induced gluons by opening the cone, and b) the opening of phase space for vacuum-like emissions that source additional energy loss. The combination of these effects balances out, leading to a very mild cone size dependence. For a full discussion of the uncertainties related to the choice of medium scales and parameters involved in the quenching, see also ref. [34].

We would like to emphasize the flexibility of the collimator function $\mathcal{C}_i$, which is independent of the particulars of the model of energy loss employed in the previous sections. One could start with other models for single-parton quenching $\mathcal{Q}^{(0)}$ and the resolved phase space $\Theta_{\rm res}$, and then use the collimator to describe the quenching of the multiple resolved sources inside the jet. As a concrete example, we provide an alternative calculation in the context of the hybrid weak- and strong-coupling model [47] in appendix D.

3 Quenching effects in the cumulative spectrum

The jet suppression factor, defined in eq. (2.2), compares the jet spectra in heavy-ion collisions (medium), to that in proton-proton collisions (vacuum) at the same final p_T. In this section, we turn to the discussion of other observables that could be constructed from

JHEP10(2021)038

the inclusive jet spectra measured in these two systems. One alternative is to compare the cumulative of the jet spectra (i.e., integrated above a fixed p_T cut). The cumulative is less affected by the initial shape of the hard spectrum and has better statistical uncertainties. Another approach is to estimate the p_T-shift necessary to match the pp and AA spectra. This is closely related to the typical amount of energy lost by a jet. Both of these procedures are straightforwardly related to the quenching factors entering the jet spectrum and will be discussed in further detail in section 3.3. Now, we turn to an observable that combines the strength of the two examples given above, namely the quantile procedure.

3.1 The quantile ratio

The quantile procedure was introduced in ref. [10] and aims to unfold the average momentum shift $\langle\varepsilon\rangle$ between vacuum and quenched jets. First, let us introduce the tail cumulative of the spectrum,

$$\Sigma(p_T, R) \equiv \int_{p_T}^{\infty} \mathrm{d}p'_T \, \frac{\mathrm{d}\sigma_R}{\mathrm{d}p'_T} \,, \tag{3.1}$$

which is a probability after dividing with the full integral. The quantile procedure compares the medium and the vacuum spectrum at equal probabilities, $\Sigma^{\mathrm{med}}(p_T^{\mathrm{q,med}}, R) = \Sigma^{\mathrm{vac}}(p_T^{\mathrm{q,vac}}, R)$. In heavy-ion collisions, for a fixed $p_T^{\mathrm{q,med}}$, this condition allows identifying the corresponding $p_T^{\mathrm{q,vac}}$. Finally, the quantile momentum ratio is defined as

$$Q_{\mathrm{med}}(p_T^{\mathrm{q,med}}) \equiv \left.\frac{p_T^{\mathrm{q,med}}}{p_T^{\mathrm{q,vac}}}\right|_{\Sigma} . \tag{3.2}$$

Therefore, $p_T^{\mathrm{q,vac}}$ is the momentum of vacuum jets above which vacuum and medium jets have equal probability to be produced.

For a quick estimate, let us assume a steeply falling spectrum with a fixed power $n = \mathrm{const}$, and neglect the R-dependence of the quenching. The tail cumulative cross-sections in vacuum and in medium (see eq. (2.1)–(2.2)) are simply

$$\begin{aligned} \Sigma^{\mathrm{vac}}\left(p_T^{\mathrm{q,vac}}\right)\big|_{n=\mathrm{const}} &= \frac{1}{n-1}\left(p_T^{\mathrm{q,vac}}\right)^{1-n} , \\ \Sigma^{\mathrm{med}}\left(p_T^{\mathrm{q,med}}\right)\Big|_{n=\mathrm{const}} &= \int_{p_T^{\mathrm{med}}}^{\infty} \mathrm{d}p_T \, p_T^{-n} \mathcal{Q}(n/p_T) \,. \end{aligned} \tag{3.3}$$

This results in the quantile momentum ratio

$$Q_{\mathrm{med}}(p_T^{\mathrm{q,med}})\Big|_{n=\mathrm{const}} = p_T^{\mathrm{q,med}} \left[(n-1)\int_{p_T^{\mathrm{q,med}}}^{\infty} \mathrm{d}p_T \, p_T^{-n} \mathcal{Q}(n/p_T)\right]^{\frac{1}{n-1}} . \tag{3.4}$$

To get the feeling for this quantity, it is instructive to consider a few simplified scenarios for the quenching factor $\mathcal{Q}(n/p_T)$. First, for a constant quenching factor $\mathcal{Q}(n/p_T) = \mathcal{Q}_0$, the quantile ratio is a trivial function of the quenching factor $Q_{\mathrm{med}} = \mathcal{Q}_0^{1/(n-1)}$. Next, we will consider the single-parton quenching factor obtained by using the soft limit ($\omega \ll \omega_c$ in eq. (2.20)) of the BDMPS-Z spectrum, which is derived in appendix B. The interplay between the jet cone and the broadening introduces a characteristic energy scale $\omega_R = \sqrt{\hat{q}L}/R$ which defines two regimes that we discuss below:

JHEP10(2021)038

- For $p_T \ll n\omega_R$, it scales parametrically as $\mathcal{Q}_>^{(0)}(p_T) = \exp[-\sqrt{\omega_1 n/p_T}]$ where $\omega_1 = 8\bar{\alpha}^2\pi\omega_c$ is a characteristic energy scale of the medium. In this case the quantile ratio becomes

$$Q_{\text{med}}(p_T) \approx \exp\left[-\sqrt{\omega_1/((n-1)p_T)}\right] = \left[\mathcal{Q}_>^{(0)}(p_T)\right]^{\frac{1}{\sqrt{n(n-1)}}} . \tag{3.5}$$

- In the high-p_T regime, $p_T \gg n\omega_R$, we instead have $\mathcal{Q}_>^{(0)} = \exp[-\omega_2 n/p_T]$ with $\omega_2 = 2\bar{\alpha}\sqrt{2\omega_c\omega_R}$. This p_T dependence is similar to that of medium-induced single hard scattering, cf. eq. (2.28), and elastic drag, cf. eq. (4.6). The quantile reads

$$Q_{\text{med}}(p_T) = \left[\frac{n-1}{y^{n-1}}\,\gamma\,(n-1,y)\right]^{\frac{1}{n-1}} , \tag{3.6}$$

  where $y = n\omega_2/p_T$ is the scaling variable and $\gamma(s,x) = \Gamma(s) - \Gamma(s,x)$ is the lower incomplete gamma function.

Both the quantile ratio Q_{med}, and the jet suppression factor R_{med} depends identically on a dimensionless ratio of a medium scale over the jet transverse momentum. Strikingly, the main difference resides in the n dependence. It turns out that the relation between quantile and quenching factor $Q_{\text{med}} \simeq \mathcal{Q}_>^{1/(n-1)}$ holds approximately also for p_T-dependent quenching factors — at least for the case of fixed n. In particular, given that $\ln \mathcal{Q}_>^{(0)} \propto 1 - n\omega_2/p_T$ at $p_T \gg n\omega_R$, we should expect a reduced sensitivity of the quantile to the details of the initial spectrum at high-p_T, i.e. $\ln Q_{\text{med}} \propto 1 - \omega_2/p_T$.

Here, we have mostly focused on the contribution from the out-of-cone, soft radiation spectrum to quenching. However, both quenching by hard emissions, see eq. (2.28) and discussion below, and elastic energy loss, see appendix D.1, behave in a similar fashion. For a single parton species, we should therefore expect to see a universal behavior, independent of the hard spectrum of the quenching at high-p_T.

We study the onset of the independence of the power-index of the hard spectrum n in figure 3. In these plots, we have computed the single-parton quenching factor $\mathcal{Q}_>^{(0)}$ for a *single parton species*, in this case quarks, using as input the full radiative spectrum from eq. (2.24), which encompass both soft and hard medium-induced emissions (dashed lines). We also plot the resummed quenching factor $\mathcal{Q}_>$, given in eq. (2.31), (solid lines). On the left in figure 3, we plot quenching factor $R_{\text{med}}(p_T)$ for a wide range of constant n values, $3 \leq n \leq 9$. Strikingly, the amount of quenching varies significantly with n as a function of p_T, encapsulating the strong bias effects. The quantile ratio, on the other hand, plotted on the right in figure 3, is remarkably resilient to the details of the hard spectrum. This holds both for the single-parton quenching factor, which was expected based on the discussion above, and the fully resummed $\mathcal{Q}_>$.

This robustness to the details of the hard, partonic spectrum was observed but not derived analytically in the original paper [10]. It is the result of the combined effect of using a momentum ratio, and using the cumulative distribution instead of the spectrum. Note, however, that we have observed scaling for quark and gluon contribution independently

JHEP10(2021)038

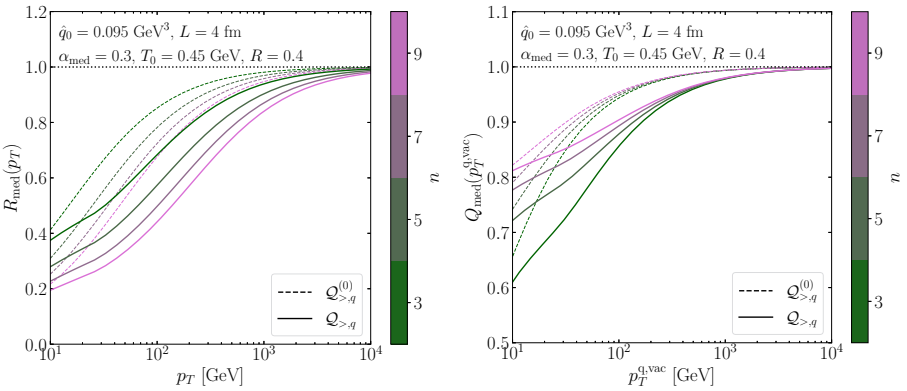


Figure 3. The quenching factor (left) and the quantile ratio (right) for quarks only, using single parton quenching (dashed lines) and quenching the whole jet (solid lines) for different spectrum power.

which only makes sense for pure samples of quark and gluon jets.[1] For realistic situations, e.g. dijet or boson-jet events, one has first to add up these contributions to the total spectrum (cf. eq. (4.2)), before computing the cumulative. In this case, the scaling features of the quantile ratio are not necessarily as transparent. We will discuss these issues in more detail in section 4.

In order to make contact with the main objective of this paper, namely to provide predictions for the quantile in heavy-ion collisions at the LHC, we round off this section by studying the R-dependence of the quantile for pure quark and gluon jets in figure 4 for fixed $n = 5$ and medium parameters given in table 1. As before, the dashed lines correspond to using eq. (2.30) for the quenching factor, which assumes that the whole jet is quenched as a single parton, while the solid lines employ the resummed quenching factor eq. (2.31). The single parton quenching trivially results in less modification and thus smaller momentum shift and quantile ratio for the same medium parameters. Generally, the qualitative features follows the naive expectation $Q_{\rm med} \simeq \mathcal{Q}_>^{1/(n-1)}$, cf. figure 2. The R-dependence is analogous to our previous discussion; less quenching results in a quenching factor closer to 1, and thus a smaller difference between the quantile momenta.

3.2 Statistical advantage of the cumulative spectrum

Using the tail-cumulative distribution has certain advantages in statistical analysis. For a steeply falling distribution $f(p_T) \sim p_T^{-n}$ (like the hadronic or jet spectrum), there are fewer and fewer entries in the higher p_T bins. This results in a rapid increase of the relative statistical uncertainty with p_T such as $\Delta f/f(p_T) \sim p_T^{n/2}(N\Delta p_T)^{-1/2}$, where N is the total number of hits, and Δp_T is the size of a bin. By using $\Sigma_f(p_T) \equiv \int_{p_T}^\infty \mathrm{d}x\, f(x)$ tail-cumulative distribution, the sum of the higher bins results in less uncertainty $\Delta\Sigma_f/\Sigma_f(p_T) \sim p_T^{-1/2}\Delta f/f(p_T)$. This can be further slowed down using $\Delta p_T(p_T) \sim p_T$ logarithmic binning,

[1] The scaling of the "gluon quantile" follows the same trends as for the quarks in figure 3, the only difference being the Casimir scaling of the quenching factors, $\mathcal{Q}_{q,>} = (\mathcal{Q}_{g,>})^{4/9}$.

JHEP10(2021)038

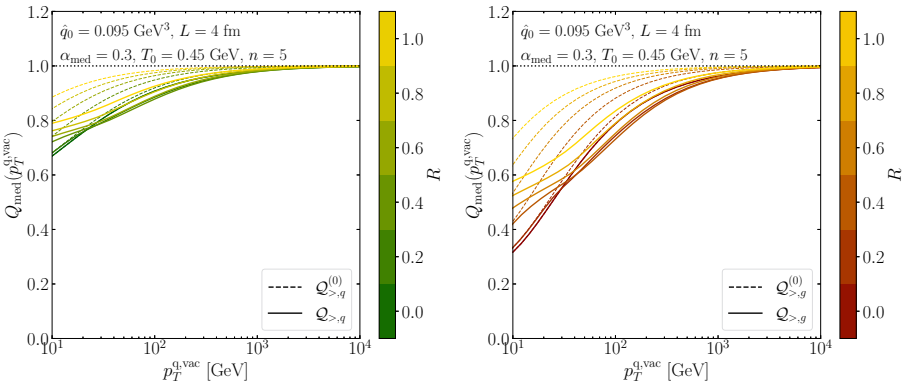


Figure 4. The quantile ratio using the eq. (2.30) for a single parton quenching (dashed line) and quenching the whole jet with coherence effects from eq. (2.31) (solid line). Quark jets on the left and gluon jets on the right panel for different cone sizes.

resulting in $\Delta f/f \sim p_T^{(n-1)/2}$ and $\Delta\Sigma_f/\Sigma_f \sim p_T^{(n-2)/2}$. It is also true that the cumulative is equivalent to the original distribution, therefore for a given set of statistical samples the tail-cumulative distribution could be advantageous.[2]

We would like to note, however, that one also has to consider systematic uncertainties. Some of them cancel in ratio observables, such as the nuclear modification factor $R_{\rm AA}$. This cancellation is less trivial in the cumulative case.

3.3 Other types of observables based on the spectrum

So far we have discussed the nuclear modification factor and the quantile procedure. But other observables related to the jet spectrum could also be defined (see also in refs. [10, 27, 48]). We demonstrate the relation between these observables within the quenching weight formalism and show how these observables are related to the quenching factor $\mathcal{Q}(n/p_T)$ (or $R_{\rm med}$) and the quantile ratio $Q_{\rm med}$ below.

- *Pseudo-quantile* [10]: is a version of the quantile procedure which matches directly the spectrum instead of the cumulative, $\mathrm{d}\sigma^{\rm med}/\mathrm{d}\tilde{p}_T^{\rm q,med} \equiv \mathrm{d}\sigma^{\rm vac}/\mathrm{d}\tilde{p}_T^{\rm q,vac}$. The condition relates the two momenta, i.e. $\tilde{p}_T^{\rm q,vac}(\tilde{p}_T^{\rm q,med})$. We obtain then,

$$\widetilde{Q}_{\rm med}\left(\tilde{p}_T^{\rm q,med}\right) \equiv \left.\frac{\tilde{p}_T^{\rm q,med}}{\tilde{p}_T^{\rm q,vac}}\right|_{\sigma} \approx \left[\mathcal{Q}_{>}\left(\tilde{p}_T^{\rm q,med}\right)\right]^{-\frac{1}{n(\tilde{p}_T^{\rm med})}}. \tag{3.7}$$

  Because of the momentum ratio, this observable has similar n dependence to the quantile ratio, and therefore it is more robust against the initial spectrum. The statistical uncertainty, however, is similar to the $R_{\rm med}$ since bins are not summed. It is equivalent with the momentum shift parameter.

[2] Up to truncation in the domain or co-domain of the distribution.

JHEP10(2021)038

- *Momentum shift* [27, 48]: is another interpretation of the pseudo-quantile, defined by demanding

$$\frac{\mathrm{d}\sigma^{\mathrm{med}}(p_T)}{\mathrm{d}p_T} \equiv \frac{\mathrm{d}\sigma^{\mathrm{vac}}(p_T + S(p_T))}{\mathrm{d}p_T}\,. \tag{3.8}$$

  The trivial connection between the definitions is $p_T^{\mathrm{vac}} = p_T + S(p_T)$, therefore

$$\widetilde{Q}_{\mathrm{med}}(p_T) = 1 + \frac{S(p_T)}{p_T}\,. \tag{3.9}$$

  Therefore, the pseudo-quantile and the momentum shift are equivalent. The spectrum shift can also be expressed with the quenching factor,

$$S(p_T) = p_T \left[\mathcal{Q}_{>}(p_T)^{-\frac{1}{n(p_T)}} - 1 \right], \tag{3.10}$$

  expressing the connection between the pseudo-quantile $\widetilde{Q}_{\mathrm{med}}$ and the shift parameter $S(p_T)$.

  A different definition of the momentum shift parameter was used by the PHENIX collaboration $S_{\mathrm{loss}}(p_T)$ [48]. Their definition, however, assumes $n = \mathrm{const}$ and postulates $S_{\mathrm{loss}}(p_T) = S_0 p_T$. The $S(p_T)$ and $\widetilde{Q}(p_T)$ are more general. We found to be necessary to consider changing power $n(p_T)$ in the spectrum, see section 4 and appendix A.

- *Cumulative-R_{med}* (or pseudo-ratio [10]): is similar to the R_{med} but uses the ratio of the cumulative spectrum instead,

$$\widetilde{R}_{\mathrm{med}}(p_T) \equiv \frac{\Sigma^{\mathrm{med}}(p_T)}{\Sigma^{\mathrm{vac}}(p_T)}\,, \tag{3.11}$$

  where the cumulative of the spectrum is defined in eq. (3.1), and where we have suppressed the R dependence for now. The integral reduces bias effects from the initial spectrum and improves the statistics as we showed in section 3.2. For a p_T-independent quenching weight, as in the single-parton $\mathcal{P}(\varepsilon)$ in the soft BDMPS-Z limit, the cumulative $\widetilde{R}_{\mathrm{med}}$ can be written as

$$\widetilde{R}_{\mathrm{med}} = \int_0^\infty \mathrm{d}\varepsilon\, \mathcal{P}(\varepsilon) \frac{\Sigma^{\mathrm{vac}}(p_T + \varepsilon)}{\Sigma^{\mathrm{vac}}(p_T)}\,. \tag{3.12}$$

  A deviation from this expectation indicates a p_T dependence of the quenching weight. Furthermore, for a spectrum with constant n, where $\Sigma^{\mathrm{vac}}(p_T + \varepsilon)/\Sigma^{\mathrm{vac}}(p_T) = (1 + \varepsilon/p_T)^{1-n}$, we simply get that $\widetilde{R}_{\mathrm{med}} = \mathcal{Q}((n-1)/p_T)$. Finally, a trivial connection between the $\widetilde{R}_{\mathrm{med}}$ and the quantile ratio for constant n is $\widetilde{R}_{\mathrm{med}}(p_T) \approx Q_{\mathrm{med}}(p_T)^{n-1}$.

In conclusion, we demonstrated the relationship between the different observables one can construct from the jet spectrum. It turns out that the quantile procedure is particularly appealing due to the reduced sensitivity to the hard spectrum and because of the improvement of the statistical uncertainties.

JHEP10(2021)038

4 Predictions

In the previous sections, we introduced the quenching weight formalism and showed the way one can construct the quenched jet spectrum and its cumulative. In this section, we extend our formalism to provide more reliable predictions for measurements in heavy-ion collisions.

4.1 Including the realistic spectrum and elastic energy loss

A realistic calculation of quenching effects has to include the partonic cross-section for jet production. This is evaluated at the hard scale of the collision, $Q_{\rm hard} \sim p_T$. The jet spectrum at a given cone size is computed then by including a DGLAP evolution to the scale $Q_{\rm jet} \sim p_T R$ [5, 49, 50]. Alternatively, the partonic cross-section that results in a jet with a given cone can be parameterized using a Monte Carlo event generator in which the partonic cross-section and the parton shower are matched. In our study, we generate events with Pythia8 [51] to fit the p_T-dependence of the spectrum with reconstruction parameter R, for quark-, and gluon-initiated jets. The large-angle DGLAP radiation results in additional R-dependence of the $R_{\rm AA}$, through the recapture of vacuum radiation. The spectrum parameterization, therefore, includes the vacuum radiation recaptured by the cone, resulting in an additional R dependence [52], see appendix C and figure 10, in particular. Currently, we restrict our study to inclusive jets in dijet samples, generated in pp, and 0–10% central PbPb collision at $\sqrt{s_{NN}} = 5.02\,\text{TeV}$, with $p_T = 20 - 1000\,\text{GeV}$ and $|\eta| < 2.8$, similar to the kinematics used by ATLAS [53].[3] The details of the event generation (excluding ISR, MPI and including the effect of nuclear PDFs in Pb, jet selection, and quark/gluon flavor assignment procedure) are described in appendix C. We also describe the proposed functional form, following ref. [32], to fit the spectral indices $n_q(p_T, R)$ and $n_g(p_T, R)$, that automatically parameterizes the (p_T, R) dependence of the quark-gluon fraction.

The nuclear modification factor $R_{\rm AA}$, defined as

$$R_{\rm AA}(p_T) = \frac{{\rm d}N^{\rm AA}(p_T, R)/{\rm d}p_T}{N_{\rm coll}\,{\rm d}\sigma^{\rm pp}(p_T, R)/{\rm d}p_T}\,, \tag{4.1}$$

where $N_{\rm coll}$ gives the number of collisions in the nuclear overlap at a given impact parameter and we identify ${\rm d}N^{\rm AA}/N_{\rm coll} = {\rm d}\sigma^{\rm AA}$. The main difference between this ratio and the previously defined $R_{\rm med}$, defined in eq. (2.2), is the addition of both quark and gluon jets with their respective quenching factors and the nPDFs. As a result, in our framework, we get

$$R_{\rm AA}(p_T, R) = \left[\sum_{i=q,g} \mathcal{Q}_i\left(p_T, R; n_i^{\rm AA_0}(p_T, R)\right) \frac{{\rm d}\sigma_i^{\rm AA_0}}{{\rm d}p_T}(p_T, R)\right] \Bigg/ \sum_{i=q,g} \frac{{\rm d}\sigma_i^{\rm pp}}{{\rm d}p_T}(p_T, R)\ , \tag{4.2}$$

where the extracted spectra are ${\rm d}\sigma_i^{\rm pp/AA_0}(p_T, R)/{\rm d}p_T$, $(i = q, g)$, and we have explicitly written out the dependence of the quenching factor on the spectral index $n_i^{\rm AA_0}(p_T, R)$.

[3] We refer the dijet $R_{\rm AA}$ as single-inclusive because jets contribute independently.

JHEP10(2021)038

Similarly, the cumulative spectrum in vacuum and medium are given by

$$\Sigma^{\mathrm{pp}}(p_T^{\mathrm{q}}, R) = \int_{p_T^{\mathrm{q}}}^{\infty} \mathrm{d}p_T \sum_{i=q,g} \frac{\mathrm{d}\sigma_i^{\mathrm{pp}}}{\mathrm{d}p_T}(p_T, R)\,, \tag{4.3}$$

$$\Sigma^{\mathrm{AA}}(p_T^{\mathrm{q}}, R) = \int_{p_T^{\mathrm{q}}}^{\infty} \mathrm{d}p_T \sum_{j=q,g} \mathcal{Q}_j\left(p_T, R, n_i^{\mathrm{AA}_0}(p_T, R)\right) \frac{\mathrm{d}\sigma_j^{\mathrm{AA}_0}}{\mathrm{d}p_T}(p_T, R)\,. \tag{4.4}$$

The quantile is finally defined as,

$$Q_{\mathrm{AA}} = \frac{p_T^{\mathrm{q,AA}}}{p_T^{\mathrm{q,pp}}}\,, \tag{4.5}$$

where the two momenta are determined from the condition $\Sigma^{\mathrm{pp}}(p_T^{\mathrm{q,pp}}, R) = \Sigma^{\mathrm{AA}}(p_T^{\mathrm{q,AA}}, R)$.

In section 2, we focused our discussion on the quenching effects emerging from medium-induced radiation and broadening. However, for realistic predictions we should also include quenching from elastic scattering. Elastic energy loss is described by the transport coefficient $\hat{e}$, which is related to $\hat{q}$ through Einstein's fluctuation-dissipation relation $\hat{e}_g = \hat{q}/(4T)$ for gluons and $\hat{e}_q = \hat{e}_g C_F/N_c$ for quarks [54, 55]. Here, $T = T_0$ is the local temperature of the plasma. We model the single-particle energy loss distribution simply as $\mathcal{P}(\varepsilon) = \delta(\varepsilon - \hat{e}L)$, where the flavor index is suppressed. Assuming that the energy lost in elastic processes thermalize instantaneously, we also build in the possibility to recover part of this energy through the phenomenological parameter R_{rec}, see eq. (2.30). This finally results in a single-parton quenching factor from elastic energy loss, given by

$$\mathcal{Q}_{\mathrm{el}}^{(0)}(\nu) = \exp\left[-\hat{e}Ln\nu\left(1 - \frac{R^2}{R_{\mathrm{rec}}^2}\right)\right]\,. \tag{4.6}$$

The criteria for resolving the partons in the jet are based on geometry and are therefore assumed to be identical for elastic and radiative processes. Therefore, the complete single-particle quenching factor $\mathcal{Q}_{>}^{(0)}$, appearing in eqs. (2.31)–(2.32), should be replaced by

$$\mathcal{Q}_{>}^{(0)}(p_T) = \mathcal{Q}_{>,\mathrm{rad}}^{(0)}(p_T)\, \mathcal{Q}_{>,\mathrm{el}}^{(0)}(p_T)\,, \tag{4.7}$$

where $\mathcal{Q}_{>,\mathrm{rad}}^{(0)}(p_T)$ is given by eq. (2.30) and $\mathcal{Q}_{>,\mathrm{el}}^{(0)}(p_T)$ is given by eq. (4.6). Including elastic effects has an important effect on the magnitude of the total quenching factor. For further detalils see appendix D.1 and figure 11.

4.2 Numerical results for dijet events

The single-inclusive jet R_{AA}, generated from a sample of dijet events, for a set of cone sizes, $0.2 < R < 1$, is shown in the left panel of figure 5 (solid curves). In the current work, the medium is treated as a static brick with fixed $\hat{q}_0$ and length L, see table 1 for details that is generally a good approximation even for expanding media, see ref. [56]. There is a notable change of the curves at high p_T due to the inclusion of nPDFs (see also in figure 10 in appendix C for only the nPDF effects). The overall R dependence is very modest and will be discussed in more detail shortly.

JHEP10(2021)038

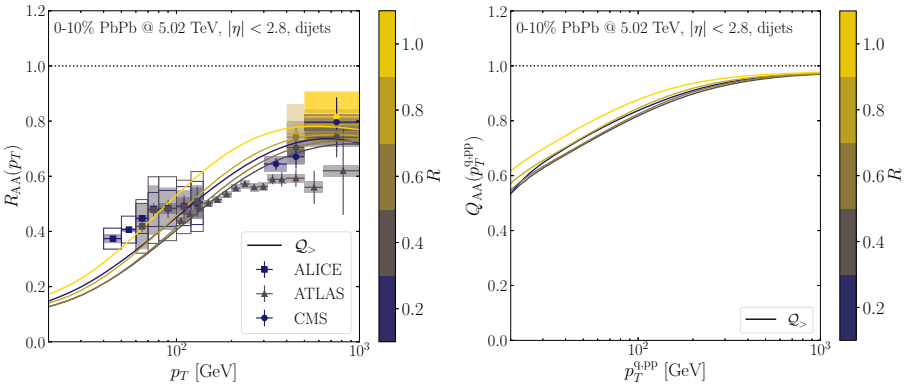


Figure 5. The $R_{\rm AA}$ (left panel) and $Q_{\rm AA}$ (right panel) from eqs. (4.2) and (4.5), respectively, for single-inclusive jets in dijets events in 0–10% PbPb collision at $\sqrt{s_{NN}}$ =5.02 TeV. The parameters are chosen to reproduce the $R = 0.4$ ATLAS [53] measurement. Data from ALICE [57] and CMS [58] are also shown.

The parameters of the calculation are tuned to the measured inclusive jet data from ATLAS [53] at $p_T \simeq 100\,\text{GeV}$ and $R = 0.4$ with $|\eta| < 2.8$, cf. table 1, resulting in good agreement between data and theory for the whole p_T range. The measured inclusive jet $R_{\rm AA}$ from ALICE [57] for $R = 0.2$ and $R = 0.4$ are also shown in figure 5, where the rapidity range for the jet selection $|\eta| < 0.5$ is slightly different. The recent CMS [58] results are also shown in figure 5 for various R, where the rapidity is $|\eta| < 2$. We would like to note, there is a disagreement between the ATLAS and CMS data that was not pointed out in the CMS' latter publication. We would like to also note that the magnitude of $R_{\rm AA}(p_T)$ in any BDMPS-Z type of calculation is mostly sensitive to the combination $\omega_s \sim \alpha_{\rm med}^2 \hat{q}_0 L^2$, as observed in refs. [27, 59]. The slope of the $R_{\rm AA}(p_T)$ is quite robust to changes in the parameters.[4] Given our simplified modeling of the medium, we do not attempt to reproduce the centrality dependence of the jet $R_{\rm AA}$ at high-p_T which will be left to future work, see also in ref. [34].

Having constrained the medium parameters with $R_{\rm AA}$, we now turn to the predictions for the quantile momentum ratio which is showed in the right panel of figure 5. There is no drastic change due to the nPDFs in comparison to the $R_{\rm AA}$ at high-p_T, demonstrating the robustness of the quantile procedure against modifications in the partonic cross-section. The shape of the curves and even the R dependence is very similar to the $R_{\rm AA}$ and is well captured by the approximate relation $Q_{\rm AA} \sim R_{\rm AA}^{1/(n-1)}$. To date, there are no experimental measurements of the quantile ratio.

Our formalism accounts for the cone size dependence of jet quenching through competing effects related to the early vacuum shower and medium-induced elastic and radiative processes. On the one hand, by opening the cone, one captures more of the particles that

[4] By including event-by-event fluctuations in the jet position and path length, the slope becomes flatter in ref. [34].

JHEP10(2021)038

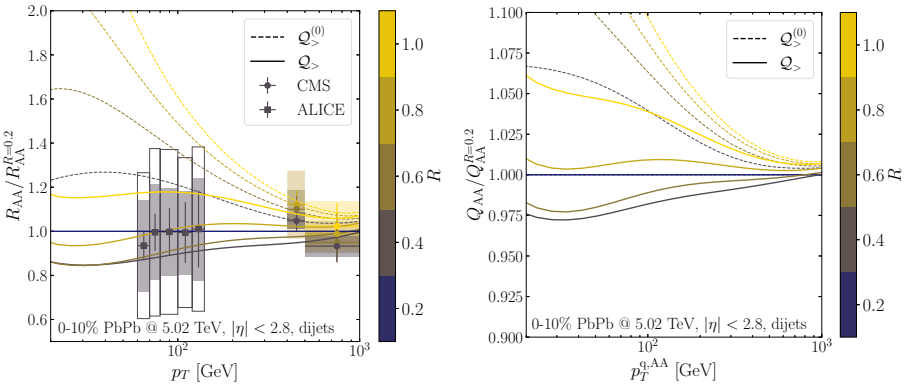


Figure 6. Same plots as in figure 5, but taking the ratio of different cone sizes to enhance the differences between single and multi-parton quenching. On the x-axis of the right panel we used $p_T^{q,AA}$ instead of $p_T^{q,pp}$ from previous.

are affected by medium interactions and thus recover the lost energy. This is manifested as a suppression of the spectrum $\mathrm{d}I_>/\mathrm{d}\omega$ at $\omega > \sqrt{\hat{q}L}/R$ in figure 1, meaning it is less probable to lose energy. For $R \sim R_{\mathrm{rec}}$ all quasi-thermalized modes, both in the radiative spectrum ($\omega < \omega_s$) and due to elastic collisions, are recaptured within the jet cone. On the other hand, opening the reconstructed jet cone results in more phase space for vacuum fragmentation at an early stage of the evolution. This leads to a higher multiplicity of vacuum-like emissions and, therefore, more sources for energy loss.

The overall effect is a relative cancellation of the R dependence (see figure 5). figure 6 shows the R_{AA} (left) and quantile ratio (right) at a given R divided by the $R = 0.2$ results. As an illustration, here the dashed curves are the results obtained by using the single-parton quenching factors $Q_>^{(0)}$, which amounts of treating the whole jet as a completely coherent single parton that is not resolved by medium interactions. Their R-dependence reflects directly the effect of recapturing energy at large angles without sourcing more energy loss though vacuum fragmentation. The full curves are for the full quenching factors which leads to a more complicated R dependence where, several effects contribute, such as vacuum fragmentation and its recapture, color coherence effects and the thermalization of the medium-induced emissions. This ratio was measured by ALICE [57] (for $|\eta| < 0.5$), and CMS [58] (for $|\eta| < 2$), showing great agreement with our model. Reference [58] compares many theory prediction and we can say, currently, our model has the best agreement.

The main differences between the "bare", and "resummed" quenching factors is at lower $p_T \sim 50 - 100\,\mathrm{GeV}$. As we mentioned, in our model the R_{AA} is mostly sensitive to the $\omega_s \sim \alpha_{\mathrm{med}}^2 \hat{q}_0 L^2$ combination of the parameters. A precise measurement on the R-dependence would help to constraint more parameters. The right side of figure 6 shows the R-dependence of the quantile ratio. It is much less sensitive to the jet cone.

JHEP10(2021)038

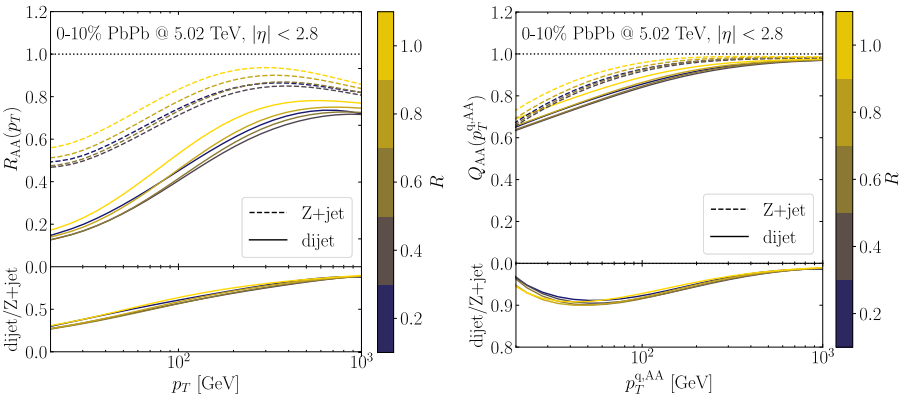


Figure 7. The $R_{\rm AA}$ and $Q_{\rm AA}$ from eqs. (4.2) and (4.5) for single-inclusive jets from dijet and Z+jet events in 0–10% PbPb collision at 5.02 TeV, keeping the same parameters as in figure 5. On the x-axis of the right panel we used $p_T^{\rm q,AA}$ instead of $p_T^{\rm q,pp}$ like in previous plots.

4.3 Differences between dijets and Z+jets

At first glance, there should be no difference in the quenching mechanism in dijet and boson+jet events.[5] However, their $R_{\rm AA}$ are not expected to be the same. In this subsection we explain why, and what are the consequences for quark and gluon classification. The Z+jet process is used as an illustration. We also show how can one use the cumulative distribution for quark-gluon discrimination to perform better statistics.

4.3.1 Difference in the $R_{\rm AA}$

In recent years, a lot of effort has been put to measure and understand the boson+jet processes both in pp and AA collisions. Firstly, it is a favorable process for quark-, and gluon-jet discrimination, because it provides a natural definition of the initial jet flavor. Secondly, measuring the boson momentum, one gets a label on the initial momentum of the recoiling jet. This is especially advantageous in heavy-ion collisions, where the quenching of bosons is suppressed,[6] and one, therefore, gains knowledge about the jet before final-state interactions with the medium.

Here, we focus on Z+jet process, but the arguments are valid for other boson+jet processes. We generated the Z+jet events with Pythia and identified jets using the same cuts as in the dijet study. Further details on the event generation and parameterization are summarized in appendix C. We include the corresponding spectrum using eq. (4.2) and the result is shown in figure 7. Dijet and boson+jet processes have different spectra and thus their bias is different on the energy loss distribution. As discussed before, this appears through the spectrum power $n(p_T)$ in the quenching factor. For dijets $n \approx 6$,

[5]The so-called "surface bias" is in our context built in due to the bias from the initial steeply falling spectrum.

[6]The electro-weak bosons can still interact with the background for example through electromagnetism.

JHEP10(2021)038

for Z+jets $n \approx 4$, therefore based on the approximate $R_{\rm AA} \sim \exp(-2\sqrt{\pi\bar{\alpha}^2\omega_c n/p_T})$, the bigger the power, the stronger is the quenching. To stress this, the dijet-, and Z+jet-$R_{\rm AA}$ ratio is also shown. The ratio of two exponential remains to be exponential that looks linear on the semi-log scale. Moreover, dijets are gluon dominated, while Z+jets are quark dominated, see in figure 10 in appendix C. Therefore we expect Z+jets to have smaller quenching, which is in agreement with figure 7. The p_T dependence of the power is also different, resulting in a different slope of $R_{\rm AA}$ as high-p_T. Note also that the Z+jet quark contribution is relatively constant in p_T in contrast to the increasing quark contribution in dijets. Since the quenching roughly scales with the color charge (gluons have more quenching), at higher p_T, Z+jets has stronger quenching, which is in agreement with the smaller slope in the plot. Although the Z+jet and dijet spectra are different, the relative R dependence is similar.

The quantile ratio is also shown in the right panel of figure 7. The difference is smaller between dijets and Z+jets than for the nuclear modification factor, pointing out the reduced sensitivity to the details of the initial spectrum (see also figure 10). Using the cumulative spectrum, and the ratio of momenta, one gets much less sensitive to the initial shape of the spectrum. The degree of scaling the initial spectrum of the quantile is nevertheless not as ideal as for the results in figure 3. This can be traced to the fact that the single inclusive jets in neither dijet nor Z+jet events are pure samples of quark-, or gluon-initiated jets. The different admixture of parton species, as well as the different level of quenching of the two both, contribute to delaying the onset of scaling effects up to higher transverse momenta. One also has to point out that figure 3 was obtained by assuming n =const, while for realistic calculations n_q and n_g are complicated functions of p_T and differ significantly for dijet and Z+jet events, see figure 10.

All in all, our results are also qualitatively similar to the ones observed in ref. [10]. Similarly to dijets, the R-dependence is very similar to the Z+jet $R_{\rm AA}$ and is to a great degree captured by the approximate relation $Q_{\rm AA} \sim R_{\rm AA}^{1/(n-1)}$.

4.3.2 Improving quark-gluon discrimination

As mentioned above, jets recoiling from a boson (γ or Z/W) is the preferred process for quark-gluon (q/g) discrimination, because the hard scattering naturally defines the initial flavor of the jet at leading order in perturbation theory. Most q/g discriminators apply cut on jet substructure observables, e.g., jet mass or soft-drop multiplicity, to classify jets, see in refs. [1, 60]. While the best performance is achieved by machine learning algorithms, these nonetheless rely on training sets resulting in model dependence. In pp, however, the main description of jets is predominantly perturbative or it can be made perturbative with grooming. Model dependence, therefore, is not as crucial as for heavy-ions, where the description is not yet unique and where non-perturbative effects are more prominent.

The recently introduced *topic modeling* [15, 61, 62] is a data-driven method that is largely model independent, and its outstanding performance was demonstrated for event generator samples both in pp and AA. For this reason, it is also applicable to heavy-ions. There are some caveats of the classifier; (i) it works only with certain observables for which quarks and gluons are *mutually irreducible* (usually counting-type observables [63, 64]), (ii)

JHEP10(2021)038

it is heavily limited by statistical uncertainty, (iii) the performance is limited by the cuts on the phase space, and (iv) it works on statistical samples.

Topic modeling aims to un-mix the sample probability distributions (e.g., dijet and boson+jet samples) to a common basis (quark and gluon distribution). Consider, for example,

$$\begin{aligned} p^{\mathrm{Zj}}(x) &= f_q^{\mathrm{Zj}} p_q(x) + (1 - f_q^{\mathrm{Zj}}) p_g(x)\,, \\ p^{\mathrm{2j}}(x) &= f_q^{\mathrm{2j}} p_q(x) + (1 - f_q^{\mathrm{2j}}) p_g(x)\,, \end{aligned} \tag{4.8}$$

where $p(x)$ is a probability density of some observable for which eq. (4.8) is true (mutual reducibility), and $f_q^{Zj/2j}$ and $f_g^{Zj/2j}$ are the weight factors. To un-mix, one uses the fact that phase-space of x exists, where either the dijet or boson+jet (and thus the quark or gluon) dominates the distribution. This usually happens on the domain border of the observable (e.g., small/big jet multiplicity) [63, 64]. With this, called anchor-bin, one can statistically decouple the basis using $p^{\mathrm{Zj}}(x) - \kappa p^{\mathrm{2j}}(x) \geq 0$,

$$\begin{aligned} p_q(x) &= \frac{p^{\mathrm{Zj}}(x) - \kappa(\mathrm{Zj}|\mathrm{2j}) p^{\mathrm{2j}}(x)}{1 - \kappa(\mathrm{Zj}|\mathrm{2j})}\,, \\ p_g(x) &= \frac{p^{\mathrm{2j}}(x) - \kappa(\mathrm{2j}|\mathrm{Zj}) p^{\mathrm{Zj}}(x)}{1 - \kappa(\mathrm{2j}|\mathrm{Zj})}\,, \end{aligned} \tag{4.9}$$

where the reducibility factor is

$$\kappa(i|j) \equiv \inf_x \frac{p^i(x)}{p^j(x)}\,, \tag{4.10}$$

where $i, j = \mathrm{Zj}, \mathrm{2j}$.

Equation (4.9) is true if eq. (4.8) is possible, and both $\inf_x p_q(x)/p_g(x) = 0$ and $\inf_x p_g(x)/p_q(x) = 0$. However, for real data it has a finite minimum. The minimum is typically on the edge of the x distribution, and thus the extraction of κ is limited by the statistical uncertainty of this corner bin. Because of the linearity, one could integrate both sides of eq. (4.8), and rewrite eq. (4.9) using the cumulative distribution of $p(x)$,

$$\tilde{\kappa}(\mathrm{Zj}|\mathrm{2j}) \equiv \inf_x \frac{\int_x^\infty \mathrm{d}x\, p^{\mathrm{Zj}}(x)}{\int_x^\infty \mathrm{d}x\, p^{\mathrm{2j}}(x)}\,, \tag{4.11}$$

$$\tilde{\kappa}(\mathrm{2j}|\mathrm{Zj}) \equiv \inf_x \frac{\int_0^x \mathrm{d}x\, p^{\mathrm{2j}}(x)}{\int_0^x \mathrm{d}x\, p^{\mathrm{Zj}}(x)}\,. \tag{4.12}$$

Our cumulative method improves the statistical uncertainty by definition (see section 3.2) that can be trivially tested with arbitrary combined distributions.[7] We would like to note the cumulative in this subsection refers to the $p(x)$ distribution, and has nothing to do with the cumulative of the jet spectrum.

Unfortunately, the jet spectrum is not mutually irreducible. One can see this from eq. (4.2), where the quenching factor depends on the quark/gluon spectrum through their n indices. However, the medium modified quark-gluon ratio of the jet spectrum is important for any quark-gluon discriminator, and thus we provide it in figure 8, for both dijets

[7] Ref. [62] mentions the possibility to use the cumulative instead of the binned histogram.

JHEP10(2021)038

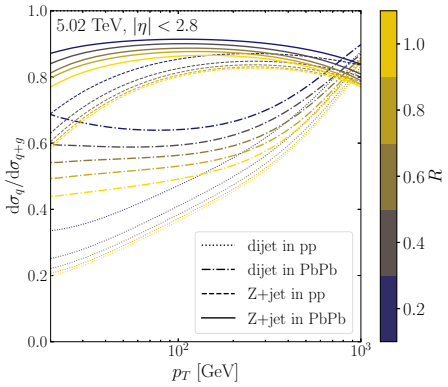


Figure 8. The quark and gluon ratio before and after quenching for dijets and Z+jets in pp and PbPb.

and Z+jets. We observe that at high-p_T the ratio barely changes, however, at lower p_T quarks start to dominate. This is the result of the stronger quenching of gluons, effectively suppressing them in the samples. For future quark-gluon classification, this suppression effect indicates that the Z+jet quark-gluon ratio will be less different from the dijet ratio, making the separation, unfortunately, harder, in line with what was observed in ref. [65].

5 Conclusion

In heavy-ion collisions, the steeply falling jet spectrum, convolved with the probability for quenching, biases the measured jet observables. In this paper, we studied the origin of this bias and its presence in recently suggested observables, e.g., the quantile ratio and its comparison of single-inclusive jet spectra in dijet and boson+jet events in heavy-ion collisions.

We developed a novel analytic framework based on quenching factors to construct the jet spectrum in heavy-ion collisions. Starting from single parton energy loss, we showed the necessity to include the full medium-induced radiation spectrum, including both multiple-soft (described by the BDMPS-Z spectrum) and single-hard (included in the GLV spectrum) scattering regimes. By keeping track of the jet cone size, the energy is only lost if the emissions propagate out of the cone. We also account for the quenching of multiple jet partons resolved by the medium interactions through the collimator function, see eq. (2.31). Therefore, opening the cone, vacuum-like jet fragmentation sources more partons to quench, resulting in a relative cancellation of the cone-size dependence. It is also important to use realistic jet spectrum for predictions by including the p_T dependent spectrum power $n(p_T)$, p_T-dependent quark-gluon jet ratio, and nPDFs. We revealed the importance of elastic energy loss, and also included it in our quenching framework for jets.

Different observables have been introduced to study the jet spectrum in heavy-ion collisions, e.g., spectrum shift, cumulative-R_{AA}, quantile or pseudo-quantile procedure. We

JHEP10(2021)038

showed their relation to the nuclear modification factor, to each other, and elucidated their dependence on the jet momentum p_T, the spectrum power n, and the jet cone size R. The cumulative-based observables reduce statistical uncertainty, and the momentum-ratio-based ones change the spectrum power dependence. The quantile momentum ratio maximizes both of these advantages, and we provided predictions for its cone-size dependence for the first time. Our approximate formula for the quantile momentum ratio $Q_{\rm AA} \sim R_{\rm AA}^{1/(n-1)}$ captures the rough properties of the observable: it is similar to $R_{\rm AA}$, with reduced spectrum power index n.

Finally, we demonstrated the bias effect by comparing quenched dijet and Z+jet spectra. For a pure sample of quark or gluon jets, the quenching factors for single-partons scale like

$$-\ln \mathcal{Q}_{>}^{(0)} \approx \begin{cases} \bar{\alpha}\sqrt{\dfrac{\hat{q}L^2 n}{p_T}} & \text{for } p_T \ll n\omega_R \\ \bar{\alpha}\dfrac{\sqrt{\omega_R \hat{q}L^2}\, n}{p_T} & \text{for } p_T \gg n\omega_R \end{cases}, \tag{5.1}$$

where $\omega_R = \sqrt{\hat{q}L}/R$. This qualitative behavior is also numerically confirmed for the resummed quenching factor $\mathcal{Q}_{>}$. As a direct manifestation of the bias effect, i.e. the n-dependence in eq. (5.1), the jet spectrum in Z+jet events, which is less steep, results in a smaller nuclear modification factor $R_{\rm AA}$ than for dijet events, see figure 3 (left). In other words, the steeper the spectrum the stronger the effect of quenching.

The quantile ratio $Q_{\rm AA}$ is much more resilient to details of the hard spectrum, both the actual values of n and of the relative admixture of quarks and gluons. For pure samples of quark-, and gluon-jets, the behavior in eq. (5.1) predicts an almost ideal scaling at $p_T \gg n\omega_R$, which was largely confirmed in figure 3 for n =const. For realistic samples of jets in dijet and Z+jet events, however, the universal behavior of the quantile ratio can only be expected to be approximate, see figure 7 (right).

The dijet and Z+jet events are also useful for quark and gluon discrimination. Following our cumulative spectrum experiences, we improved quark-, and gluon-jet discrimination based on topics modeling in general. However, due to the n-dependence of the quenching factors, quarks and gluons in the jet spectrum are not mutually irreducible. We will return to the challenging of quark/gluon discrimination in heavy-ion collisions in future work.

Many of the assumptions underlying the concrete realization behind our numerical results in section 2 are already implemented in varying degrees in various phenomenological Monte Carlo models. Our framework, therefore, provides theoretical tools to organize the effects of quark/gluon contributions, jet fragmentation, and finally, the details of medium interactions. In the BDMPS-Z framework, these are, to a first approximation, all controlled by a single transport coefficient $\hat{q}$.

Acknowledgments

We thank T. S. Biró, J. Brewer, P. Caucal, J. Isaksen, G. Milhano, D. Pablos, J. Thaler and D. Perepelitsa for helpful discussions. The work is supported by a Starting Grant

JHEP10(2021)038

from Trond Mohn Foundation (BFS2018REK01) and the University of Bergen. A.T. is also supported by the MCnetITN3 H2020 Marie Curie Initial Training Network, contract 722104, and wishes to thank the Institut de Physique Theorique (IPhT) and Gregory Soyez for the hospitality.

A Corrections to the Laplace transformation and to the running power

Our objective in this paper, is to compute the ratio of medium to vacuum inclusive jet spectra, which can be written as (see eqs. (2.1)–(2.2))

$$R_{\rm med}(p_T) = \int_0^\infty \mathrm{d}\varepsilon\, \mathcal{P}(\varepsilon) \left(1 + \frac{\varepsilon}{p_T}\right)^{-n} , \tag{A.1}$$

where we omit the R dependence for now, and assume that $n = \text{const}$. In eq. (2.3), we took the $(1 + \varepsilon/p_T)^{-n} \approx \exp(-n\varepsilon/p_T)$ approximation. We can, however, easily include corrections to this by noting that

$$\begin{aligned} R_{\rm med}(p_T) &= \int_0^\infty \mathrm{d}\varepsilon\, \mathcal{P}(\varepsilon) \left[1 + \frac{(\nu\varepsilon)^2}{2n} - \frac{(\nu\varepsilon)^3}{3n^2} + \mathcal{O}((\nu\varepsilon)^4)\right] \mathrm{e}^{-\nu\varepsilon} \\ &= \left[1 + \frac{\nu^2}{2n}\frac{\partial^2}{\partial\nu^2} + \frac{\nu^3}{3n^2}\frac{\partial^3}{\partial\nu^3} + \mathcal{O}(\nu^4\partial_\nu^4)\right] \mathcal{Q}(\nu)\,, \end{aligned} \tag{A.2}$$

where $\mathcal{Q}(\nu) \equiv \int_0^\infty \mathrm{d}\varepsilon\, \mathcal{P}(\varepsilon)\mathrm{e}^{-\nu\varepsilon}$ is the Laplace transform of the energy loss distribution and $\nu = n/p_T$. In this appendix, we investigate the impact of these higher-order corrections for a concrete example that can be solved analytically, namely the energy loss distribution obtained in the strictly soft limit of the BDMPS-Z spectrum. It is given by

$$\mathcal{P}(\varepsilon) = \sqrt{\frac{\omega_s}{\varepsilon^3}}\mathrm{e}^{-\frac{\pi\omega_s}{\varepsilon}} , \tag{A.3}$$

which only depends on the energy scale ω_s and is properly normalized. In this case, its Laplace transform is $\mathcal{Q} = \exp(-2\sqrt{\pi\omega_s\nu})$. We can, in fact find any of the terms in eq. (A.2) by noticing that

$$\mathcal{I}_m \equiv \frac{\partial^m}{\partial\nu^m}\mathcal{Q}(\nu) = (-1)^m 2 \left(\frac{\nu}{\pi}\right)^{\frac{1-2m}{4}} \omega_s^{\frac{1+2m}{4}} K_{m-\frac{1}{2}}(2\sqrt{\pi\omega_s\nu})\,, \tag{A.4}$$

where $K_m(x)$ is the modified Bessel function of the second kind and $\mathcal{I}_0 = \mathcal{Q}(\nu)$. We can therefore write

$$R_{\rm med}(p_T) = \sum_{m=0}^{\infty} c_m \mathcal{I}_m\,, \tag{A.5}$$

where, $c_0 = 1$, $c_1 = 0$, $c_2 = \nu^2/(2n)$, $c_3 = \nu^3/(3n^2)$ and so forth, by assuming $R_{\rm med}(p_T)$ to be analytic function. On the left of figure 9, we study the corrections by comparing to the exact value from eq. (A.1). The parameters we use are $\omega_s = 5\,\text{GeV}$ and $\omega_s = 10\,\text{GeV}$ and $n = 5$. For the realistic choice of $\omega_s \lesssim 5\,\text{GeV}$, the leading behaviour is already of the order of $\mathcal{O}(10^{-2})$, even at low $p_T \sim 100\,\text{GeV}$.

JHEP10(2021)038

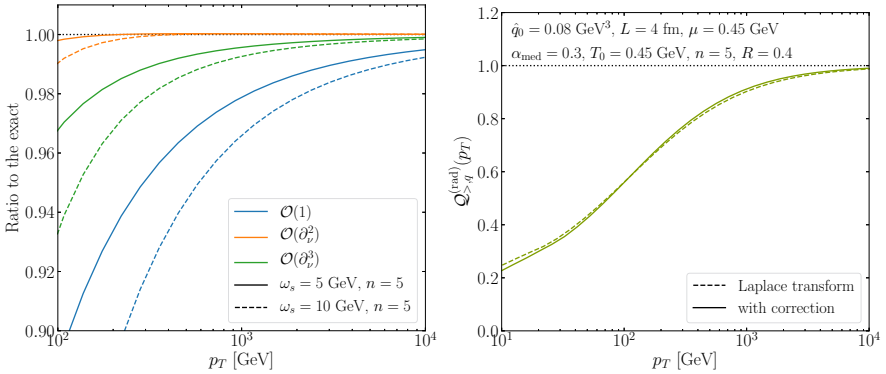


Figure 9. Ratio of the truncated expansion of $R_{\rm med}$ from eq. (A.2) to the exact value for the soft BDMPS-Z model from eq. (A.1) (left), and the same with the full quenching weight from eq. (2.31).

Finally, the effect of the corrections from eq. (A.2) for the realistic quenching weight employed in the main body of the paper is showcased in figure 9 (right) up to the leading correction $\mathcal{O}(\partial_\nu^2)$. The sign of the correction depends on the concavity/convexity of the quenching weight (note the second derivative in p_T in eq. (A.2)). For our choice of parameters, the correction is tiny $< \mathcal{O}(10^{-3})$, and can safely be neglected.

The conclusions from above also hold if we consider a p_T dependent power $n(p_T)$ of the hard spectrum. One can trivially replace $n \mapsto n(p_T)$ in the quenching weight capturing most of the effects of the running power. The correction to this simple replacement is

$$R_{\rm med}(p_T) \approx \left[\left(1 - \frac{1}{p_T}\frac{\partial n}{\partial p_T}\frac{\partial^2}{\partial \nu^2}\right)\left(1 + \ln(p_T)\frac{\partial n}{\partial p_T}\frac{\partial}{\partial \nu}\right)\right]\mathcal{Q}(\nu)\,, \tag{A.6}$$

resulting in negligible $\ll \mathcal{O}(10^{-3})$ corrections similar to the corrections to the Laplace transformation.

B p_T-scaling of the out-of-cone emission quenching factor

Consider the BDMPS-Z spectrum in the soft limit. In the short formation-time approximation the emission and subsequent broadening of a soft gluon factorizes, and we can write

$$\omega\frac{\mathrm{d}I}{\mathrm{d}\omega\mathrm{d}k_\perp^2\,\mathrm{d}t} = \bar{\alpha}\sqrt{\frac{\hat{q}}{\omega}}\mathcal{P}(k_\perp, L-t)\,, \tag{B.1}$$

where $\mathcal{P}(k_\perp, L-t)$ is the probability for a particle emitted at time t to acquire transverse momentum $k_\perp = \omega\theta$ through elastic scattering up to the end of the medium L. In the Gaussian approximation, it reads

$$\mathcal{P}(k_\perp, L-t) = \frac{4\pi}{\hat{q}(L-t)}\mathrm{e}^{-\frac{k_\perp^2}{\hat{q}(L-t)}}\,. \tag{B.2}$$

JHEP10(2021)038

For this simplified ansatz, the out-of-cone spectrum reads

$$\frac{\mathrm{d}I_{>}}{\mathrm{d}\omega} = \bar{\alpha}\sqrt{\frac{\hat{q}L^2}{\omega}}\mathcal{B}\left(\frac{\omega^2}{\omega_R^2}\right), \tag{B.3}$$

where $\omega_R = \sqrt{\hat{q}L}/R$ and the broadening factor $\mathcal{B}$ is given by

$$\mathcal{B}(y) = \frac{1}{L}\int_0^L \mathrm{d}t \int_{(\omega R^2)}^{\infty} \mathrm{d}k_\perp^2\, \mathcal{P}(k_\perp, L-t) = \mathrm{e}^{-y} - y\Gamma(0,y)\,, \tag{B.4}$$

where $y = \omega^2/\omega_R^2$. Therefore, eq. (B.3) falls rapidly for $\omega > \omega_R$, and thus we used ω_R in the matching definition in eq. (2.13). The single-parton quenching factor reads then

$$\begin{aligned}\ln \mathcal{Q}_{>}^{(0)}(\nu) &= -\bar{\alpha}\sqrt{2\omega_c\nu}\int_0^{\infty}\mathrm{d}x\,\frac{1}{x^{3/2}}\left(1-\mathrm{e}^{-x}\right)\mathcal{B}\left(\frac{x^2}{x_R^2}\right), \\ &\approx -\bar{\alpha}\sqrt{2\omega_c\nu}\int_0^{x_R}\mathrm{d}x\,\frac{1}{x^{3/2}}\left(1-\mathrm{e}^{-x}\right)\end{aligned} \tag{B.5}$$

where we changed variables to $x = \omega\nu$ and $x_R \equiv \omega_R\nu$. We will solve this integral in two limiting cases, namely $x_R \ll 1$ and $x_R \gg 1$. In the former case, which corresponds to the high-p_T regime where $p_T \gg \omega_R n$, we can expand the terms in the bracket and find

$$\ln \mathcal{Q}_{>}^{(0)}(\nu)\Big|_{x_R\ll 1} = -2\bar{\alpha}\sqrt{2\omega_c\omega_R}\nu\,. \tag{B.6}$$

In the opposite limit, for $p_T \ll \omega_R n$, we can extend the upper integration limit to infinity, to obtain

$$\ln \mathcal{Q}_{>}^{(0)}(\nu)\Big|_{x_R\gg 1} = -2\bar{\alpha}\sqrt{2\pi\omega_c\nu}\,, \tag{B.7}$$

which is independent of the jet cone.

C Generating and parametrizing the jet spectrum

As we mentioned in section 2, the partonic cross-section to produce a jet with a given p_T and R is perturbatively calculable in the collinear factorization up to high precision [5, 6]. In our work, instead, we extract the spectrum using the Pythia8.235 event generator [51]. To generate dijet events we used default settings and tunes with `HardQCD:All` both in pp and in 0–10% PbPb collision at 5.02 TeV. This results in LO $2 \to 2$ matrix elements. The nPDF was EPS09LO which has a relatively important effect on the R_{AA}, see in figure 10. The ISR, MPI, and hadronization were turned off to focus on final state radiation only. We reconstructed jets using anti-k_t algorithm with FastJet3 [66] for $R = 0-1$, $p_{T,\mathrm{jet}} = 10 - 1000\,\mathrm{GeV}$ and $|\eta_{\mathrm{jet}}| < 2.8$, similar to the kinematic cuts of ATLAS.[8] To label the flavors of the jets, we compared them to the outgoing partons from the hard scattering, and we kept the closest in angle if it was less than 2R. We only associated one jet (the hardest) with an initiator, and thus we only kept the two hardest associated jets.

[8] At this rapidity selection and jet cones, the ISR and MPI could contribute to jet production, that we address in a future study.

JHEP10(2021)038

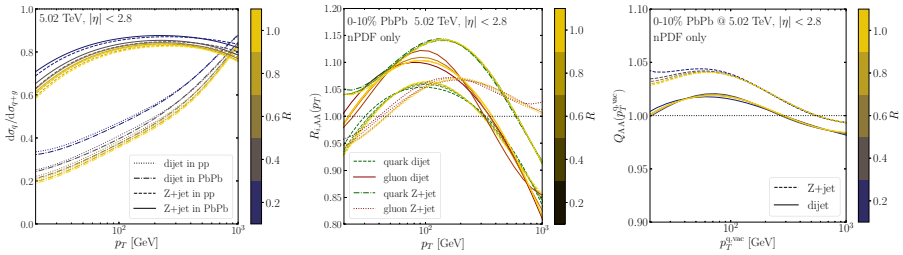


Figure 10. *Left*: the parametrized quark jet contribution in pp and in PbPb using the nPDF in Pythia. *Middle*: the quenching factors for quarks and gluons resulting from nPDFs (no quenching). *Right:* the quantile ratio resulting from nPDFs alone.

This selection detail becomes important for small R jets, where more than one jet can be reconstructed for one initiator. We kept those events in which there no jet passing the criteria, which are important in the proper $R_{\rm AA}$ ratio (before quenching). We parametrized the spectrum following [32],

$$\frac{\mathrm{d}\sigma_i^{\rm pp/AA}}{\mathrm{d}p_T}(p_T,R) = c_0\left(\frac{p_T}{p_0}\right)^{-n_i^{\rm pp/AA}(p_T,R)}\,, \tag{C.1}$$

$$n_i^{\rm pp/AA}(p_T,R) = \sum_{n=1} c_n \log^n\left(\frac{p_T}{p_0}\right)\,, \tag{C.2}$$

where i is the flavor of the initiator of the jet and $\{p_0, c_n\}$ are $(p_T, R, i, \mathrm{pp/AA})$ dependent fitting parameters. We kept terms up to $n = 3$, achieving $< 3\%$ relative deviation. The resulted parametrization is showed in figure 10. On the left, there is the quark contribution, which increases with p_T. The cone size dependence shows, gluons are emitted at larger angles even in the vacuum. On the middle the $R_{\rm AA}$ is shown, resulted by the nPDF (no quenching on the plot). The inclusive jet spectrum would be similar to the dijets keeping not only the two hardest jets, however, the flavor assignment would be less trivial especially for smaller cone sizes, therefore we preferred to use the dijet samples.[9] On the right, the quantile ratio is shown resulted only from the nPDF effects. One can see that the ratio is close to 1, in contrast to the quenching observed in figure 5. The p_T dependence is also different compared to quenching. At high $p_T \gtrsim 300\,\mathrm{GeV}$, the nPDF suppresses the spectrum (because of the EMC effect, see the $R_{\rm AA}$ in the middle), resulting $Q_{\rm AA} < 1$, while at $p_T \approx 20 - 300\,\mathrm{GeV}$, the nPDF enhances the spectrum (due to anti-shadowing), resulting $Q_{\rm AA} > 1$, see [68] for more details.

For the Z+jet samples, we followed the logic from previous. We used however the `WeakBosonAndParton:qg2gmZq` and `WeakBosonAndParton:qqbar2gmZg` processes and we forced the Z-boson to decay invisibly. In this case, we kept only the hardest recoiling jet if it fulfilled the same criteria as before. In the Z+jet case, the spectrum power is smaller in comparison to dijets $n_{\rm Zj} \approx 4 < n_{\rm 2j} \approx 6$. The quark and gluon ratio is different, in comparison to dijets, but the cone size dependence is similar, see on the left in figure 10.

[9] For a recent development on jet flavor definition, see ref. [67].

JHEP10(2021)038

In both cases, opening the cone, one captures more gluons meaning, they are radiated at larger angles. The $R_{\rm AA}$ resulted from the nPDF is also shown in figure 10, relatively similar for both dijet and Z+jet processes. The quark/gluon content is however slightly different. On the right, the quantile ratio is also shown to be similar to dijets.

D Other models

As we mentioned in section 2 the quenching weight and collimator formalism are independent of the energy loss description. To illustrate this we show how to use this formalism to describe quenching of elastic scattering and within the strong coupling approximation.

D.1 Momentum shift and elastic energy loss

The simplest example for the jet energy loss probability $\mathcal{P}(\varepsilon)$ is to take a momentum shift of vacuum jets $\mathcal{P}(\varepsilon) = \delta(\varepsilon - S(p_T))$, resulting equivalent definition to the old fashioned momentum shift parameter $\mathrm{d}\sigma_{\rm med}(p_T)/\mathrm{d}p_T \equiv \mathrm{d}\sigma_{\rm vac}(p_T + S(p_T))/\mathrm{d}p_T$ [27, 48]. By expanding the steeply falling spectrum in $S(p_T) \ll p_T$,

$$\frac{\mathrm{d}\sigma_{\rm med}}{\mathrm{d}p_T}(p_T) = \sum_n^\infty \langle \varepsilon^n \rangle \frac{\mathrm{d}^n}{\mathrm{d}p_T^n}\left(\frac{\mathrm{d}\sigma_{\rm vac}}{\mathrm{d}p_T}\right) \approx \frac{\mathrm{d}\sigma_{\rm vac}}{\mathrm{d}p_T}(p_T + \langle\varepsilon\rangle)\,, \tag{D.1}$$

where we used $\langle \varepsilon^i \rangle \approx \langle \varepsilon \rangle^i$. The shift parameter is roughly the mean energy loss

$$S(p_T) \approx \langle \varepsilon \rangle = \int \mathrm{d}\varepsilon\, \varepsilon\, \mathcal{P}(\varepsilon) \equiv \Delta E\,. \tag{D.2}$$

The quenching factor in the simple power-law case is

$$R_{\rm med}(p_T) = p_T^{n(p_T) - n(p_T + S(p_T))} \left[1 + \frac{S(p_T)}{p_T}\right]^{-n(p_T + S(p_T))}\,, \tag{D.3}$$

which is well approximated by $R_{\rm med}(p_T) \approx [1 + S(p_T)/p_T]^{-n(p_T)}$ (see in appendix A). The quantile ratio is straightforward by using eq. (3.2) with $\mathcal{Q} = [1 + S(p_T)/p_T]^{-n(p_T)}$. In the constant power approximation it is

$$Q_{\rm med}(p_T) \approx p_T \left[(n-1)\int_{p_T}^\infty \mathrm{d}p\, (p_T + S(p_T))^{-n}\right]^{-\frac{1}{1-n}}\,. \tag{D.4}$$

The elastic scattering is approximated by a constant momentum shift $\mathcal{P}_{{\rm el},i}(\varepsilon) = \delta(\varepsilon - \hat{e}_i L)$, where $\hat{e} = -\mathrm{d}\langle E\rangle/\mathrm{d}t \approx C_i \hat{q}_0/(4 N_c T_0)$, and $T_0 \approx 0.45\,\mathrm{GeV}$ [55, 69]. This can be translated to an additional quenching weight compared to the form eq. (2.30),

$$\mathcal{Q}^{(0)}_{>,{\rm el}}(p_T, R) = \exp\left[-\frac{\hat{e} L\, n}{p_T}\left(1 - \frac{R^2}{R^2_{\rm rec}}\right)\right], \tag{D.5}$$

where we included some energy recapture through the second term with $R_{\rm rec} = \pi/2$. Therefore in total $\mathcal{Q}^{(0)}_{>,{\rm tot}} = \mathcal{Q}^{(0)}_{>,{\rm rad}}\mathcal{Q}^{(0)}_{>,{\rm el}}$. This factor runs slower with p_T than the BDMPS-Z, similar to the "$N = 1$" (or GLV) spectrum. We can also include the elastic energy loss of each jet constituent, by using the collimator function form eq. (2.32), $\mathcal{Q}_{>,{\rm tot}} = \mathcal{Q}_{>,{\rm rad}}\mathcal{Q}_{>,{\rm el}}$. Figure 11 shows eq. (D.5) with dashed lines and with the collimator with full lines, indicating the importance of elastic scattering in the overall quenching. Therefore we included this effect in section 4.

JHEP10(2021)038

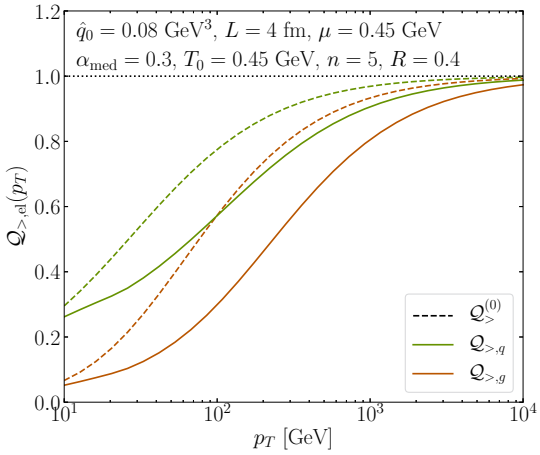


Figure 11. The quenching factor of elastic scatterings from eq. (D.5) and with the collimator from eq. (2.32).

D.2 Hybrid weak-, strong-coupling

In the strong coupling approximation, the energy loss is assumed to be described by a semi-classical string falling inside a black hole horizon [70]. This model does not use $\mathrm{d}I/\mathrm{d}\omega$ because there are no emitted gluons, and thus the energy loss is directly connected with $\mathcal{P}^{(0)}(\varepsilon)$. The average lost energy of a single parton traversing through an L sized, T temperature strongly coupled medium is

$$\frac{\Delta E}{E} = 1 - \frac{2}{\pi}\left[\frac{L}{x_s}\sqrt{1-\left(\frac{L}{x_s}\right)^2} + \cos^{-1}\left(\frac{L}{x_s}\right)\right], \tag{D.6}$$

where E is the initial energy, $x_s = E^{1/3}/(2\kappa_{\rm sc}T^{4/3})$ is the stopping length and $\kappa_{\rm sc} = 1.05\, g^{1/3}N_c^{1/6}$. Using the definition of momentum shift from eq. (D.2), the single parton quenching is estimated by

$$\mathcal{Q}^{(0)}(p_T) = \left[2 - \frac{2}{\pi}\left(\frac{L}{x_s}\sqrt{1-\left(\frac{L}{x_s}\right)^2} + \cos^{-1}\left(\frac{L}{x_s}\right)\right)\right]^{-n}, \tag{D.7}$$

where the p_T dependence is presented in x_s. For high p_T, $\mathcal{Q}^{(0)}(p_T) \approx 1 - 32n\kappa_{\rm sc}^3 L^3T^4/(3\pi p_T)$ the same p_T dependence obtained from GLV (see eq. (2.28) and below). We did not include here the broadening in and out of the cone due to the lack of particles, however, one could include the linearized hydro response through $R_{\rm rec}(\eta,\phi)$. We include multi parton quenching and thus the R-dependence by using the collimator from eq. (2.32). We used eq. (D.7) as $\mathcal{Q}^{(0)}(p_T)$, and for $\Theta_{\rm res}$ we used the resolution condition used in the hybrid model [71], $t_{\rm f} < t_{\rm d} = (\theta\mu)^{-1} < L$, where $\mu = \pi T/2$ is the IR screening scale used in the hybrid model (the Debye mass would be gT). eq. (D.7) is shown in figure 12

JHEP10(2021)038

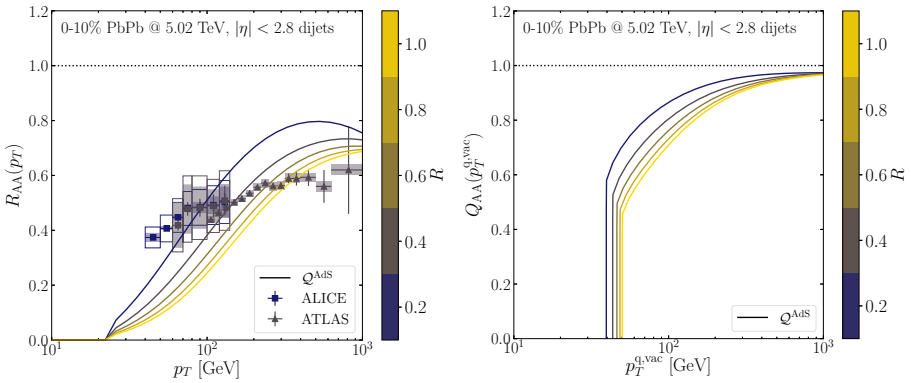


Figure 12. The quenching and the quantile ratio in the strong coupling assumption with the proper quark and gluon mixture using eq. (D.7) and including the collimator to consider multi-parton effects from eq. (2.32).

with the Pythia spectrum, resulting similar quenching to the GLV assumption. The curves in figure 12 similar to the results in [72] without the medium response, where the sharp cutoff is where the medium fully absorbs the jet (the stopping length). The R-dependence is simplified in our case, because we neglected the medium response. We also estimate the quantile ratio using eq. (D.3). The parameters are $T = 0.27\,\text{GeV}$, $L = 4\,\text{fm}$ and $\kappa_{sc} = 0.4$.

References

[1] A.J. Larkoski, I. Moult and B. Nachman, *Jet substructure at the Large Hadron Collider: a review of recent advances in theory and machine learning*, *Phys. Rept.* **841** (2020) 1 [arXiv:1709.04464] [INSPIRE].

[2] S. Marzani, G. Soyez, and M. Spannowsky, *Looking inside jets: an introduction to jet substructure and boosted-object phenomenology*, Springer, Germany (2019).

[3] M. Dasgupta, F.A. Dreyer, K. Hamilton, P.F. Monni, G.P. Salam and G. Soyez, *Parton showers beyond leading logarithmic accuracy*, *Phys. Rev. Lett.* **125** (2020) 052002 [arXiv:2002.11114] [INSPIRE].

[4] R. Kogler et al., *Jet substructure at the Large Hadron Collider: experimental review*, *Rev. Mod. Phys.* **91** (2019) 045003 [arXiv:1803.06991] [INSPIRE].

[5] M. Dasgupta, F. Dreyer, G.P. Salam and G. Soyez, *Small-radius jets to all orders in QCD*, *JHEP* **04** (2015) 039 [arXiv:1411.5182] [INSPIRE].

[6] M. Dasgupta, F.A. Dreyer, G.P. Salam and G. Soyez, *Inclusive jet spectrum for small-radius jets*, *JHEP* **06** (2016) 057 [arXiv:1602.01110] [INSPIRE].

JHEP10(2021)038

[7] K. Rajagopal, A.V. Sadofyev and W. van der Schee, *Evolution of the jet opening angle distribution in holographic plasma*, *Phys. Rev. Lett.* **116** (2016) 211603 [arXiv:1602.04187] [INSPIRE].

[8] J. Casalderrey-Solana, Z. Hulcher, G. Milhano, D. Pablos and K. Rajagopal, *Simultaneous description of hadron and jet suppression in heavy-ion collisions*, *Phys. Rev. C* **99** (2019) 051901 [arXiv:1808.07386] [INSPIRE].

[9] Y.-L. Du, D. Pablos and K. Tywoniuk, *Deep learning jet modifications in heavy-ion collisions*, *JHEP* **21** (2020) 206 [arXiv:2012.07797] [INSPIRE].

[10] J. Brewer, J.G. Milhano and J. Thaler, *Sorting out quenched jets*, *Phys. Rev. Lett.* **122** (2019) 222301 [arXiv:1812.05111] [INSPIRE].

[11] CMS collaboration, *Studies of jet quenching using isolated-photon+jet correlations in PbPb and pp collisions at* $\sqrt{s_{NN}} = 2.76$ *TeV*, *Phys. Lett. B* **718** (2013) 773 [arXiv:1205.0206] [INSPIRE].

[12] CMS collaboration, *Study of jet quenching with* $Z +$ *jet correlations in Pb-Pb and pp collisions at* $\sqrt{s_{NN}} = 5.02$ *TeV*, *Phys. Rev. Lett.* **119** (2017) 082301 [arXiv:1702.01060] [INSPIRE].

[13] ATLAS collaboration, *Measurement of photon-jet transverse momentum correlations in* 5.02 *TeV Pb + Pb and pp collisions with ATLAS*, *Phys. Lett. B* **789** (2019) 167 [arXiv:1809.07280] [INSPIRE].

[14] ATLAS collaboration, *Comparison of fragmentation functions for jets dominated by light quarks and gluons from pp and Pb+Pb collisions in ATLAS*, *Phys. Rev. Lett.* **123** (2019) 042001 [arXiv:1902.10007] [INSPIRE].

[15] J. Brewer, J. Thaler and A.P. Turner, *Data-driven quark and gluon jet modification in heavy-ion collisions*, *Phys. Rev. C* **103** (2021) L021901 [arXiv:2008.08596] [INSPIRE].

[16] Y. He, L.-G. Pang and X.-N. Wang, *Bayesian extraction of jet energy loss distributions in heavy-ion collisions*, *Phys. Rev. Lett.* **122** (2019) 252302 [arXiv:1808.05310] [INSPIRE].

[17] R. Baier, Y.L. Dokshitzer, A.H. Mueller, S. Peigne and D. Schiff, *Radiative energy loss and* p_T *broadening of high-energy partons in nuclei*, *Nucl. Phys. B* **484** (1997) 265 [hep-ph/9608322] [INSPIRE].

[18] B.G. Zakharov, *Fully quantum treatment of the Landau-Pomeranchuk-Migdal effect in QED and QCD*, *JETP Lett.* **63** (1996) 952 [hep-ph/9607440] [INSPIRE].

[19] R. Baier, Y.L. Dokshitzer, A.H. Mueller and D. Schiff, *Medium induced radiative energy loss: equivalence between the BDMPS and Zakharov formalisms*, *Nucl. Phys. B* **531** (1998) 403 [hep-ph/9804212] [INSPIRE].

[20] M. Gyulassy, P. Levai and I. Vitev, *Reaction operator approach to nonAbelian energy loss*, *Nucl. Phys. B* **594** (2001) 371 [nucl-th/0006010] [INSPIRE].

[21] S. Caron-Huot and C. Gale, *Finite-size effects on the radiative energy loss of a fast parton in hot and dense strongly interacting matter*, *Phys. Rev. C* **82** (2010) 064902 [arXiv:1006.2379] [INSPIRE].

[22] X. Feal and R. Vazquez, *Intensity of gluon bremsstrahlung in a finite plasma*, *Phys. Rev. D* **98** (2018) 074029 [arXiv:1811.01591] [INSPIRE].

JHEP10(2021)038

[23] C. Andres, L. Apolinário and F. Dominguez, *Medium-induced gluon radiation with full resummation of multiple scatterings for realistic parton-medium interactions*, *JHEP* **07** (2020) 114 [arXiv:2002.01517] [INSPIRE].

[24] Y. Mehtar-Tani, *Gluon bremsstrahlung in finite media beyond multiple soft scattering approximation*, *JHEP* **07** (2019) 057 [arXiv:1903.00506] [INSPIRE].

[25] Y. Mehtar-Tani and K. Tywoniuk, *Improved opacity expansion for medium-induced parton splitting*, *JHEP* **06** (2020) 187 [arXiv:1910.02032] [INSPIRE].

[26] J. Barata and Y. Mehtar-Tani, *Improved opacity expansion at NNLO for medium induced gluon radiation*, *JHEP* **10** (2020) 176 [arXiv:2004.02323] [INSPIRE].

[27] R. Baier, Y.L. Dokshitzer, A.H. Mueller and D. Schiff, *Quenching of hadron spectra in media*, *JHEP* **09** (2001) 033 [hep-ph/0106347] [INSPIRE].

[28] C.A. Salgado and U.A. Wiedemann, *Calculating quenching weights*, *Phys. Rev. D* **68** (2003) 014008 [hep-ph/0302184] [INSPIRE].

[29] Y. Mehtar-Tani and K. Tywoniuk, *Sudakov suppression of jets in QCD media*, *Phys. Rev. D* **98** (2018) 051501 [arXiv:1707.07361] [INSPIRE].

[30] M.L. Miller, K. Reygers, S.J. Sanders and P. Steinberg, *Glauber modeling in high energy nuclear collisions*, *Ann. Rev. Nucl. Part. Sci.* **57** (2007) 205 [nucl-ex/0701025] [INSPIRE].

[31] F. Arleo, *Quenching of hadron spectra in heavy ion collisions at the LHC*, *Phys. Rev. Lett.* **119** (2017) 062302 [arXiv:1703.10852] [INSPIRE].

[32] M. Spousta and B. Cole, *Interpreting single jet measurements in Pb + Pb collisions at the LHC*, *Eur. Phys. J. C* **76** (2016) 50 [arXiv:1504.05169] [INSPIRE].

[33] J.-W. Qiu, F. Ringer, N. Sato and P. Zurita, *Factorization of jet cross sections in heavy-ion collisions*, *Phys. Rev. Lett.* **122** (2019) 252301 [arXiv:1903.01993] [INSPIRE].

[34] Y. Mehtar-Tani, D. Pablos and K. Tywoniuk, *Cone size dependence of jet suppression in heavy-ion collisions*, arXiv:2101.01742 [INSPIRE].

[35] B.G. Zakharov, *Radiative energy loss of high-energy quarks in finite size nuclear matter and quark-gluon plasma*, *JETP Lett.* **65** (1997) 615 [hep-ph/9704255] [INSPIRE].

[36] R. Baier, Y.L. Dokshitzer, A.H. Mueller, S. Peigne and D. Schiff, *Radiative energy loss of high-energy quarks and gluons in a finite volume quark-gluon plasma*, *Nucl. Phys. B* **483** (1997) 291 [hep-ph/9607355] [INSPIRE].

[37] J.-P. Blaizot, F. Dominguez, E. Iancu and Y. Mehtar-Tani, *Medium-induced gluon branching*, *JHEP* **01** (2013) 143 [arXiv:1209.4585] [INSPIRE].

[38] L. Apolinário, N. Armesto, J.G. Milhano and C.A. Salgado, *Medium-induced gluon radiation and colour decoherence beyond the soft approximation*, *JHEP* **02** (2015) 119 [arXiv:1407.0599] [INSPIRE].

[39] M. Gyulassy and X.-n. Wang, *Multiple collisions and induced gluon Bremsstrahlung in QCD*, *Nucl. Phys. B* **420** (1994) 583 [nucl-th/9306003] [INSPIRE].

[40] U.A. Wiedemann, *Gluon radiation off hard quarks in a nuclear environment: opacity expansion*, *Nucl. Phys. B* **588** (2000) 303 [hep-ph/0005129] [INSPIRE].

[41] C. Andres, F. Dominguez and M. Gonzalez Martinez, *From soft to hard radiation: the role of multiple scatterings in medium-induced gluon emissions*, *JHEP* **03** (2021) 102 [arXiv:2011.06522] [INSPIRE].

JHEP10(2021)038

[42] X. Feal, C.A. Salgado and R.A. Vazquez, *Jet quenching test of the QCD matter created at RHIC and the LHC needs opacity-resummed medium induced radiation*, *Phys. Lett. B* **816** (2021) 136251 [arXiv:1911.01309] [INSPIRE].

[43] J. Barata, Y. Mehtar-Tani, A. Soto-Ontoso and K. Tywoniuk, *Revisiting transverse momentum broadening in dense QCD media*, *Phys. Rev. D* **104** (2021) 054047 [arXiv:2009.13667] [INSPIRE].

[44] Y. Mehtar-Tani, C.A. Salgado and K. Tywoniuk, *The radiation pattern of a QCD antenna in a dense medium*, *JHEP* **10** (2012) 197 [arXiv:1205.5739] [INSPIRE].

[45] Y. Mehtar-Tani and K. Tywoniuk, *Radiative energy loss of neighboring subjets*, *Nucl. Phys. A* **979** (2018) 165 [arXiv:1706.06047] [INSPIRE].

[46] B. Blok and K. Tywoniuk, *Higher-order corrections to heavy-quark jet quenching*, *Eur. Phys. J. C* **79** (2019) 560 [arXiv:1901.07864] [INSPIRE].

[47] J. Casalderrey-Solana, D.C. Gulhan, J.G. Milhano, D. Pablos and K. Rajagopal, *A hybrid strong/weak coupling approach to jet quenching*, *JHEP* **10** (2014) 019 [*Erratum ibid.* **09** (2015) 175] [arXiv:1405.3864] [INSPIRE].

[48] PHENIX collaboration, *A detailed study of high-p_T neutral pion suppression and azimuthal anisotropy in Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV*, *Phys. Rev. C* **76** (2007) 034904 [nucl-ex/0611007] [INSPIRE].

[49] Z.-B. Kang, F. Ringer and I. Vitev, *The semi-inclusive jet function in SCET and small radius resummation for inclusive jet production*, *JHEP* **10** (2016) 125 [arXiv:1606.06732] [INSPIRE].

[50] L. Dai, C. Kim and A.K. Leibovich, *Fragmentation of a jet with small radius*, *Phys. Rev. D* **94** (2016) 114023 [arXiv:1606.07411] [INSPIRE].

[51] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159 [arXiv:1410.3012] [INSPIRE].

[52] M. Dasgupta, L. Magnea and G.P. Salam, *Non-perturbative QCD effects in jets at hadron colliders*, *JHEP* **02** (2008) 055 [arXiv:0712.3014] [INSPIRE].

[53] ATLAS collaboration, *Measurement of the nuclear modification factor for inclusive jets in Pb+Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV with the ATLAS detector*, *Phys. Lett. B* **790** (2019) 108 [arXiv:1805.05635] [INSPIRE].

[54] G.D. Moore and D. Teaney, *How much do heavy quarks thermalize in a heavy ion collision?*, *Phys. Rev. C* **71** (2005) 064904 [hep-ph/0412346] [INSPIRE].

[55] Y. Tachibana, N.-B. Chang and G.-Y. Qin, *Full jet in quark-gluon plasma with hydrodynamic medium response*, *Phys. Rev. C* **95** (2017) 044909 [arXiv:1701.07951] [INSPIRE].

[56] P. Caucal, E. Iancu and G. Soyez, *Jet radiation in a longitudinally expanding medium*, *JHEP* **04** (2021) 209 [arXiv:2012.01457] [INSPIRE].

[57] ALICE collaboration, *Measurements of inclusive jet spectra in pp and central Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV*, *Phys. Rev. C* **101** (2020) 034911 [arXiv:1909.09718] [INSPIRE].

[58] CMS collaboration, *First measurement of large area jet transverse momentum spectra in heavy-ion collisions*, *JHEP* **05** (2021) 284 [arXiv:2102.13080] [INSPIRE].

JHEP10(2021)038

[59] P. Caucal, E. Iancu and G. Soyez, *Deciphering the z_g distribution in ultrarelativistic heavy ion collisions*, *JHEP* **10** (2019) 273 [arXiv:1907.04866] [INSPIRE].

[60] A.J. Larkoski and E.M. Metodiev, *A theory of quark vs. gluon discrimination*, *JHEP* **10** (2019) 014 [arXiv:1906.01639] [INSPIRE].

[61] E.M. Metodiev and J. Thaler, *Jet topics: disentangling quarks and gluons at colliders*, *Phys. Rev. Lett.* **120** (2018) 241602 [arXiv:1802.00008] [INSPIRE].

[62] P.T. Komiske, E.M. Metodiev and J. Thaler, *An operational definition of quark and gluon jets*, *JHEP* **11** (2018) 059 [arXiv:1809.01140] [INSPIRE].

[63] A.J. Larkoski, I. Moult and D. Neill, *Power counting to better jet observables*, *JHEP* **12** (2014) 009 [arXiv:1409.6298] [INSPIRE].

[64] C. Frye, A.J. Larkoski, J. Thaler and K. Zhou, *Casimir meets Poisson: improved quark/gluon discrimination with counting observables*, *JHEP* **09** (2017) 083 [arXiv:1704.06266] [INSPIRE].

[65] Y.-T. Chien and R. Kunnawalkam Elayavalli, *Probing heavy ion collisions using quark and gluon jet substructure*, arXiv:1803.03589 [INSPIRE].

[66] M. Cacciari, G.P. Salam and G. Soyez, *FastJet user manual*, *Eur. Phys. J. C* **72** (2012) 1896 [arXiv:1111.6097] [INSPIRE].

[67] J. Baron, D. Reichelt, S. Schumann, N. Schwanemann and V. Theeuwes, *Soft-drop grooming for hadronic event shapes*, *JHEP* **07** (2021) 142 [arXiv:2012.09574] [INSPIRE].

[68] N. Armesto, *Nuclear shadowing*, *J. Phys. G* **32** (2006) R367 [hep-ph/0604108] [INSPIRE].

[69] G.-Y. Qin and X.-N. Wang, *Jet quenching in high-energy heavy-ion collisions*, *Int. J. Mod. Phys. E* **24** (2015) 1530014 [arXiv:1511.00790] [INSPIRE].

[70] P.M. Chesler and K. Rajagopal, *Jet quenching in strongly coupled plasma*, *Phys. Rev. D* **90** (2014) 025033 [arXiv:1402.6756] [INSPIRE].

[71] Z. Hulcher, D. Pablos and K. Rajagopal, *Resolution effects in the hybrid strong/weak coupling model*, *JHEP* **03** (2018) 010 [arXiv:1707.05245] [INSPIRE].

[72] D. Pablos, *Jet suppression from a small to intermediate to large radius*, *Phys. Rev. Lett.* **124** (2020) 052301 [arXiv:1907.12301] [INSPIRE].

JHEP10(2021)038

Article IV.

5.4 Dynamically groomed jet radius in heavy-ion collisions

Paul Caucal, Alba Soto-Ontoso, and Adam Takacs

Phys.Rev D, **105**, 114046 (2022)

Paper about jet substructure in heavy-ion collisions. Keywords: jet substructure, jet energy-loss, out-of-cone radiation, dynamical grooming, event generators.

Dynamically groomed jet radius in heavy-ion collisions

Paul Caucal,[1,*] Alba Soto-Ontoso,[2,†] and Adam Takacs[3,‡]

[1]*Physics Department, Brookhaven National Laboratory, Upton, New York 11973, USA*
[2]*Université Paris-Saclay, CNRS, CEA, Institut de physique théorique, 91191, Gif-sur-Yvette, France*
[3]*Department of Physics and Technology, University of Bergen, Bergen 5020, Norway*

(Received 2 December 2021; accepted 25 May 2022; published 28 June 2022)

We explore the ability of a recently proposed jet substructure technique, dynamical grooming, to pin down the properties of the quark-gluon plasma formed in ultrarelativistic heavy-ion collisions. In particular, we compute, both analytically and via Monte Carlo simulations, the opening angle θ_g of the *hardest* splitting in the jet as defined by dynamical grooming. Our calculation, grounded in perturbative QCD (pQCD), accounts for the factorization in time between vacuum-like and medium-induced processes in the double logarithmic approximation. We observe that the dominant scale in the θ_g distribution is the decoherence angle θ_c which characterizes the resolution power of the medium to propagating color probes. This feature also persists in strong coupling models for jet quenching. We further propose for potential experimental measurements a suitable combination of the dynamical grooming condition and the jet radius that leads to a pQCD-dominated observable with a very small sensitivity ($\lesssim 10\%$) to medium response.

DOI: 10.1103/PhysRevD.105.114046

I. JET SUBSTRUCTURE IN HEAVY-ION COLLISIONS

The use of jet substructure techniques in heavy-ion collisions is ramping up; see Refs. [1,2] and references therein. From a theoretical viewpoint, there are certain advantages when considering observables defined in terms of one or a few jet constituents with respect to global ones such as fragmentation functions [3–5] or jet shapes [6–9]. In particular, jet substructure observables can be engineered to enhance the sensitivity to certain regions of the radiation phase space where perturbative QCD effects dominate, thus enabling first principles calculations. Experimentally, fully corrected jet substructure measurements are now available in heavy-ion collisions both at RHIC and LHC energies [10,11]. They are highly complementary to the rich data set recorded in pp collisions both for low [12,13] and high-p_t jets (e.g., Refs. [14,15]).

Up to now, the jet substructure program in the heavy-ion community has strongly focused on SoftDrop (SD) observables [16,17]. They are defined in terms of the kinematics of the first branching in an angular-ordered splitting tree whose momentum sharing[1] z satisfies the so-called SoftDrop condition, $z > z_{\rm cut}\theta^{\beta}$, where θ is the relative angle of the branching and $(z_{\rm cut}, \beta)$ are free parameters. One of such observables is the distribution of z values that pass the SD cut. In vacuum, the z_g distribution for $\beta = 0$ is known to scale as the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi splitting function [18–20] to lowest order in perturbative QCD (pQCD), i.e., $\mathrm{d}\sigma/\mathrm{d}z_g \sim 1/z_g$ [21,22]. In the medium, several ingredients are expected to play a role. On the one hand, assuming that the interaction between the high energetic jet and the medium is dominated by multiple, soft scatterings, an enhancement of low-z_g splittings is expected due to the $\propto z^{-3/2}$ scaling of the medium-induced radiative spectrum [23,24]. On the other hand, incoherent energy loss leads to more asymmetric splittings being suppressed with respect to the vacuum baseline [25]. These two competing effects are, in general, hard to disentangle and their relative magnitude will depend on the jet p_t together with the parameters of the grooming condition; see Ref. [26]. The first z_g measurement in heavy-ion collisions by CMS showed a steeper z_g distribution with respect to the vacuum baseline [27]. The theoretical interpretation of this softening remains unclear given that the data has been quantitatively reproduced by models whose in-medium dynamics are disparate; e.g., Refs. [28,29] related the enhancement of soft particles to

*pcaucal@bnl.gov
†alba.soto@ipht.fr
‡adam.takacs@uib.no

[1]For a given splitting with prongs i and j, the momentum sharing is defined as $z = \min(E_i + E_j)/(E_i + E_j)$, where E represents the energy.

the medium-modified splitting functions while Ref. [30] proposed a medium-response based description of the data. In addition, no obvious modification of the z_g distribution has been observed at RHIC energies [12]. The simultaneous description of both data sets has been provided in Ref. [31] where it was argued that the energy dependence of the z_g distribution was dominated by the density of the quark-gluon plasma (QGP) together with coherent energy loss. Nevertheless, a back of the envelope calculation shows that the k_t of the splitting probed by the RHIC measurement can be as large as $k_t \propto z p_t R$, that is $k_t = 1$ GeV for $p_t = 25$ GeV, $R = 0.4$ and $z = 0.1$. Thus, nonperturbative dynamics are expected to play a role and a purely pQCD approach might not be well suited. Along these lines, the impact of the fluctuating thermal background on the z_g distribution, among other observables, has been recently assessed in Refs. [1,32]. The authors showed that mistagged splittings can induce a non-negligible contribution [$\mathcal{O}(10\%)$] that mimics a jet quenching signal. This fact lead ALICE to increase the value of $z_{\rm cut}$ from the standard value in pp, i.e., $z_{\rm cut} = 0.1$, to 0.2 in their recent publication [11]. Also, in this recent measurement, the z_g is integrated over all possible angles of the splitting. These two combined ingredients, i.e., the enhanced $z_{\rm cut}$ and the integration over the angles, result in an unmodified z_g distribution.

Another SoftDrop observable that has been studied in the context of jet quenching is the opening angle of the SD splitting, θ_g. The physics motivation in this case is related to the intrinsic medium angular scale θ_c that divides the radiation phase space into resolved and unresolved splittings [33–35]. In a nutshell, splittings with $\theta > \theta_c$ lose more energy than those with $\theta < \theta_c$. Then, the steeply falling nature of the jet p_t spectrum leads to a filtering effect such that only quasicollinear splittings pass the selection cut and thus a narrowing of the θ_g distribution when compared to pp is to be expected [36]. There is a competing effect that leads to a broadening of the θ_g distribution, namely transverse momentum diffusion of each of the resolved branches when $\theta_g > \theta_c$ [37]. Recent measurements by ALICE [11] indicate an overall narrowing of the θ_g distribution with respect to pp. The physics mechanism driving this observation is far from being settled given that models with [38] and without a color (de)coherence mechanism [37,39,40] are able to semi-quantitatively describe the data.

Overall, out of the theoretical curves presented in Ref. [11], the models that correctly reproduce both z_g and θ_g data are (i) JetMed [26,38], where the coherence angle is built in, (ii) the Hybrid [39] with a fully incoherent energy loss picture and (iii) the JetScape [40] result using MATTER + LBT [41,42], a model completely agnostic to θ_c. Given the lack of consensus in the theoretical interpretation of the SoftDrop measurements, a natural question is whether instead of merely adopting jet substructure techniques that were designed by the pp community, one should develop specific tools best suited to in-medium jet physics. As we have already mentioned, the necessity to double the value of $z_{\rm cut}$ to mitigate the impact of the underlying event highlights the specificities of heavy-ion collisions. Another example along this direction of thought is Ref. [43] where a jet clustering algorithm that uses as metric the formation time of the splitting was explored. Regarding groomers, Ref. [44] proposed the dynamical grooming procedure which relies on identifying the hardest splitting in the QCD shower as defined by the maximal value of the so-called "hardness" variable

$$\kappa^{(a)} = \frac{1}{p_{t,\rm jet}} z(1-z) p_t \left(\frac{\theta}{R}\right)^a, \tag{1}$$

where $a > 0$ is a continuous free parameter and $(p_{t,\rm jet}, R)$ is the transverse momentum and cone size of the jet. The most natural values of a from a heavy-ion perspective are $a = 1, 2$. Indeed, when setting $a = 1$ the splitting with the largest transverse momentum, k_t, is selected. This choice is interesting from the point of view of probing the quasiparticle nature of the QGP. In short, rare, hard scatterings between the propagating color probe and the medium lead to a $\propto 1/k_t^4$ scaling of the k_t distribution's tail. Thus, an enhancement in the distribution at large k_t could serve as the smoking gun for Rutherford-like scatterings [45,46]. Further, selecting $a = 2$ corresponds to the splitting with the shortest formation time. These splittings will likely be resolved by the medium and therefore larger in-medium modifications with respect to the vacuum dynamics are expected. One could then subdivide a sample of jets into short and large formation time splittings and compare the size of the modifications with respect to the pp result [43].

So far, the dynamical grooming technique has only been applied to pp physics both theoretically and experimentally. In particular, Ref. [47] presented a thorough examination of the analytic structure of dynamically groomed observables that led to a quantitative description of the ALICE preliminary data [48,49]. Equipped with a solid understanding of the vacuum benchmark, we extend the theoretical calculation to in-medium jet physics. In this paper, we focus on the angle of the splitting tagged by dynamical grooming, while the relative k_t will be presented in a separate publication [50]. The goal of this paper is to showcase the main physics ingredients that enter into the theoretical calculation of θ_g and ease the interpretation of Monte Carlo results. In particular, we demonstrate that this observable is highly sensitive to the quark-gluon plasma resolution angle.

The analytic calculation is presented in Sec. II. After a brief reminder of the vacuum calculation, we move on to the in-medium theoretical analysis in Sec. II B. We build up our toy in-medium shower incrementally such that the

impact of each ingredient in the θ_g distribution can be clearly disentangled. The final theoretical curves can be found in Sec. II C where we make use of the Kolmogorov-Smirnov metric to quantify the discriminating power of the observable. Then, we compare our analytic estimates to a pQCD-based Monte Carlo in Sec. III A. We show that the effects that we observe are qualitatively robust even in a strong coupling description of the medium by using the Hybrid model. Finally, in Sec. III C, we evaluate the impact of the medium response using state-of-the-art jet quenching Monte Carlo generators and present a systematic study of the best setup to enhance the impact of pQCD physics on this observable experimentally. We conclude and outline some additional ideas in Sec. IV. The numerical routines used in this publication can be found in Ref. [51].

II. THEORETICAL SETUP

We begin by formulating the dynamical grooming technique in its most general terms, i.e., independently of whether the emission takes place in vacuum or in the medium. Here, we provide the main formulas and refer the reader to Refs. [44,47] for a more detailed discussion on their derivation. Our main assumption is that we work in the soft and collinear limit such that z, $\theta \ll 1$ and we can neglect momentum degradation along the jet primary branch. Then, Eq. (1) reduces to

$$\kappa^{(a)} = z\left(\frac{\theta}{R}\right)^a. \tag{2}$$

Next, we take the $\kappa \ll 1$ limit in order for resummation techniques to apply [47]. The probability distribution for a splitting to be the hardest in a QCD jet can be written as

$$\frac{\mathrm{d}^2\mathcal{P}(z,\theta|a)}{\mathrm{d}z\mathrm{d}\theta} = \frac{\mathrm{d}^2\tilde{P}(z,\theta)}{\mathrm{d}z\mathrm{d}\theta}\Delta(\kappa|a), \tag{3}$$

where $\mathrm{d}^2\tilde{P}(z,\theta)$ is a branching kernel that represents the probability of a splitting with (z,θ) to take place along the jet fragmentation and $\Delta(\kappa|a)$ is a Sudakov form factor that is the probability of no emission with hardness larger than $\kappa^{(a)}$. These two functions are related by

$$\ln\Delta(\kappa|a) = -\int_0^1 \mathrm{d}z' \int_0^R \mathrm{d}\theta' \frac{\mathrm{d}^2\tilde{P}(z',\theta')}{\mathrm{d}z'\mathrm{d}\theta'} \times \Theta(z'(\theta'/R)^a - \kappa^{(a)}). \tag{4}$$

Although left implicit, note that the branching kernel and, thereby, the Sudakov form factor appearing in Eqs. (3) and (4) depend on the color representation of the jet-initiating parton. In addition, to guarantee the collinear safety of the Sudakov form factor, we require $a > 0$.

In this work, we are interested in the angular distribution of the splitting tagged by dynamical grooming. It is obtained directly from Eq. (3) by marginalizing over z. That is,

$$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\theta_g}\bigg|_a = \int_0^1 \mathrm{d}z \frac{\mathrm{d}^2\mathcal{P}(z,\theta|a)}{\mathrm{d}z\mathrm{d}\theta}\delta(\theta-\theta_g). \tag{5}$$

Note that this differential distribution is self-normalized by definition.

The purpose of the next sections is to compute Eq. (5) for vacuum and in-medium jets.

A. Vacuum recap

In the double-logarithmic approximation (DLA) on which we rely throughout this paper[2] it is sufficient to consider the branching kernel in the soft ($z \ll 1$) and collinear ($\theta \ll 1$) limit,

$$\frac{\mathrm{d}^2\tilde{P}^{\mathrm{vac}}(z,\theta)}{\mathrm{d}z\mathrm{d}\theta} = \frac{2\alpha_s C_R}{\pi}\frac{1}{z\theta}, \tag{6}$$

where C_R is the Casimir factor of the representation of the leading parton. At this level of accuracy, the strong coupling constant is fixed to the hardest transverse momentum scale of the problem $Q = p_{t,\mathrm{jet}}R$, namely $\alpha_s \equiv \alpha_s(Q)$. The remaining integrations in Eq. (5) can be carried out analytically and yield

$$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\theta_g} = \frac{1}{\theta_g}\sqrt{\bar{\alpha}\pi a}\left[\mathrm{erf}\left(\sqrt{\bar{\alpha}a}\ln\left(\frac{\theta_g}{R}\right)\right)+1\right], \tag{7}$$

with $\bar{\alpha} \equiv \alpha_s C_R/\pi$. This distribution is shown in Fig. 1 as a function of the grooming parameter a. We observe that the lower the value of a, the more collinear the "hardest" (or tagged) splitting is. In fact, this is confirmed analytically after taking the first derivative of Eq. (7), and we obtain the maximum of the distribution to be

$$\ln\left(\frac{1}{\theta_{\max}}\right) = \frac{1}{2a\bar{\alpha}} + \mathcal{O}(1). \tag{8}$$

To conclude this vacuum recap, we would like to emphasize that the regions of phase space that the θ_g observable explores are heavily correlated with the choice of a. In particular, setting $a \geq 1$ leads to the observable being sensitive to wide angle dynamics. This observation will play an important role in the next section, where we extend the calculation to account for in-medium jet evolution.

B. In-medium calculation

After more than two decades of active theoretical work, our understanding of how to describe the fragmentation process of a highly energetic parton in the medium has

[2]See Ref. [47] for a higher-order computation.

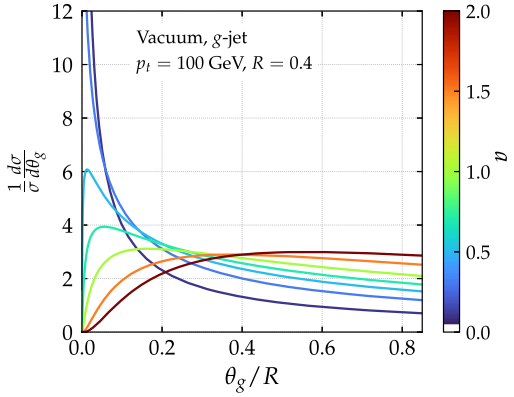


FIG. 1. θ_g distribution in vacuum for various values of the dynamical grooming parameter a. The peak position of the distribution scales as $1/a$.

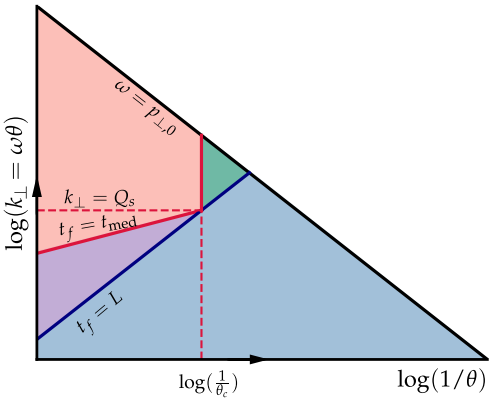


FIG. 2. Schematic Lund plane representation of the relevant regions of phase space for an in-medium jet. The physical meaning of the different lines is explained in the main text.

significantly improved. One of the building blocks of an in-medium parton shower is the medium-induced radiative spectrum. Recently, it has been computed beyond its two asymptotic limits: (i) rare, hard [52] and (ii) multiple, soft scatterings [23,24] between the jet and the medium constituents. A few numerical approaches can be found in Refs. [53–56], while a novel expansion scheme has been proposed at the analytic level [57] and applied to some global observables in Refs. [58,59]. Significant advances have been made on the evolution of this gluonic cascade [60–63] and its cross-talk with the vacuum evolution [38]. In addition, semianalytic approaches [58,64] are now incorporating realistic collision geometries and local medium properties from the hydrodynamic evolution of the medium.

In this work, we make three major simplifications to facilitate analytic manipulations. First, we restrict ourselves to the double logarithmic limit of pQCD. In this limit, there is a factorization in time between vacuum-like emissions (enhanced by large soft and collinear logarithms) and medium-induced emissions [38]. Further, we treat the medium as a static brick of length L. This allows us to neglect the time dependence of the quenching parameter, i.e., $\hat{q}(t) \equiv \hat{q}\Theta(L-t)$. In the first part of this section, L is taken as a constant, while by the end of it we evaluate the impact of its fluctuations by using a probability distribution that mimics the collision geometry. Last, we describe the jet-medium interaction in the multiple soft, scattering approximation. That is, we only account for inelastic collisions with low-momentum exchanges between the propagating parton and the medium constituents. These interactions lead to medium-induced emissions together with a Gaussian diffusion in transverse space characterized by the momentum scale $Q_s^2 \equiv \hat{q}L$.

Within this simplified scenario, the phase space for the first branching can be sketched in a Lund-plane representation [65] as the one provided in Fig. 2, where we use the transverse momentum $k_\perp$ of the emission and its opening angle θ as coordinates. We can approximate the quantum-mechanical formation time of the emission as $t_f \simeq 2/(k_\perp\theta) = 2/(\omega\theta^2)$. Let us discuss the different regions in Fig. 2.

(i) *Blue region:* The most obvious constraint on the radiative phase space is generated by the finite length of the medium; emissions with $t_f > L$ are created outside of it and thus their fragmentation process develops as in vacuum.

(ii) *Red and purple regions:* On the other hand, branchings with $t_f < L$ can be classified into two categories: vacuum-like (VLEs) and medium-induced (MIEs). In this case, the relevant scale arises by considering that any emission inside the medium has a minimum transverse momentum set by the one acquired via multiple soft collisions during its formation time:

$$k_\perp^2 \geq k_{\perp,\rm med}^2 \equiv \hat{q}t_f. \tag{9}$$

Emissions which saturate this constraint, $k_\perp = k_{\perp,\rm med}$, are medium-induced, while emissions with $k_\perp^2 \gg k_{\perp,\rm med}$ are vacuum-like.[3] In terms of formation time, the latter condition becomes $t_f \ll t_f^{\rm med}$ with the formation time of a medium-induced emission given by

$$t_f^{\rm med} = \sqrt{2\omega/\hat{q}}. \tag{10}$$

[3]Large-$k_\perp$ emissions can also be triggered by single hard collisions with a medium scattering center, but we neglect this kind of contribution in this study.

Therefore, vacuum-like emissions have much shorter formation times than medium-induced ones. Consequently, vacuum-like emissions are vetoed in the $t_f^{\rm med} < t_f \ll L$ region, i.e., the purple area in Fig. 2 [38]. Besides $t_f^{\rm med}$, there is yet one more scale that plays a prominent role in this paper, i.e., the decoherence angle θ_c given by (see e.g., Refs. [25,34])

$$\theta_c = \frac{2}{\sqrt{\hat{q}L^3}}. \tag{11}$$

As we have already mentioned, this angular scale separates resolved from unresolved emissions. The purple region contains branchings with $\theta > \theta_c$. As such, the two prongs act as independent emitters after the splitting.

(3) *Green region:* In this area, splittings are typically vacuum-like, but are never resolved by the medium given that $\theta < \theta_c$. These emissions lose energy as a single color charge.

In what follows, we present analytic estimates for the dynamically groomed θ_g distributions of splittings generated in the regions of phase space that we have discussed above. We would like to remark that it is not the aim of this paper to provide precise analytic predictions, but rather to illustrate the main physics ingredients that enter into the theoretical calculation of θ_g in order to facilitate an interpretation of the Monte Carlo results that will be shown by the end of this manuscript.

1. Vacuum-like emissions

Formally, the only leading-logarithmic effect of the medium on the dynamically groomed distributions is caused by the veto constraint on vacuum-like emissions in the presence of a dense medium (see purple region in Fig. 2). As we have described above, emissions whose formation time satisfies $t_f^{\rm med} < t_f < L$ and whose angle is $\theta > \theta_c$ are vetoed, and therefore, Eq. (3) is amended accordingly:

$$\frac{{\rm d}^2\mathcal{P}^{\rm vle}}{{\rm d}z{\rm d}\theta} = \frac{{\rm d}^2\tilde{P}^{\rm vac}(z,\theta)}{{\rm d}z{\rm d}\theta}\Theta_{\notin{\rm veto}}(z,\theta)\Delta_{\notin{\rm veto}}(\kappa|a), \tag{12}$$

where ${\rm d}^2\tilde{P}^{\rm vac}(z,\theta)$ is given by Eq. (6). In this case, the Sudakov form factor reads

$$\ln\Delta_{\notin{\rm veto}}(\kappa|a) = -\int_0^1 {\rm d}z' \int_0^R {\rm d}\theta' \frac{{\rm d}^2\tilde{P}^{\rm vac}(z',\theta')}{{\rm d}z'{\rm d}\theta'} \times \Theta_{\notin{\rm veto}}(z',\theta')\Theta(z'(\theta'/R)^a - \kappa), \tag{13}$$

with

$$\begin{aligned}\Theta_{\notin{\rm veto}}(z,\theta) &= 1-\Theta(\theta-\theta_c)\Theta(t_f - t_f^{\rm med})\Theta(L-t_f)\\ &= 1-\Theta(\theta-\theta_c)\Theta(2\hat{q}-z^3p_t^3\theta^4)\\ &\quad\times\Theta(zp_t\theta^2L-2).\end{aligned} \tag{14}$$

Note that these medium boundaries are known at double logarithmic accuracy only, meaning that the numerical prefactors (such as the factors of 2 in the veto constraint) are not under control and have been chosen in this way for convenience. Consequently, one can perfectly neglect single logarithmic terms such as hard collinear or running coupling corrections in Eq. (12) since our calculation cannot be more accurate than double-log due to medium-related uncertainties in the phase space for vacuum-like emissions.

The calculation of Eq. (13) is provided in Appendix A. The integral over z' is done analytically, while the remaining integral over θ' is performed numerically to avoid the difficulties related to the complicated shape of the integration domain. In Fig. 3, we present the impact of the veto constraint on the θ_g distribution. The kinematic parameters are chosen to resemble an ALICE-like setup[4]: $\alpha_s = 0.2, R = 0.4, p_t = 100$ GeV/c, $L = 4$ fm and $\hat{q} = 1.5$ GeV^2/fm. The medium parameters are tuned such that our final theoretical result agrees with the nuclear modification, R_{AA}, in the ALICE jet selection window [66]. In addition, the jet p_t always refers to the final transverse momentum, i.e., after quenching, although when energy loss is absent this value coincides with the p_t of the initiator. We observe how the presence of the veto region leads to a relative narrowing of the distribution for $a = 1, 2$. This is expected given that the veto region mainly prohibits large angle emissions and thus, collinear radiation is enhanced. Due to the self-normalization of the observable, this leads to a depletion of wide angle splittings. In the case of $a = 0.1$, the effect is negligible since it tags narrow splittings by construction. Overall, the effects are only sizable for $\theta_g/R \ll 1$. Our main interest in this paper is to design an observable that enhances the sensitivity to the critical angle θ_c and that is the reason why we choose an angular observable such as the θ_g distribution. If instead one would like to maximize the impact of the veto region, it would be more convenient to explore observables with large values of a, like the groomed mass ($m_g^2 \sim z\theta^2$), such that the tagging condition is parallel to the $t_{\rm med}$ line in Fig. 2.

2. Medium-induced emissions

The dynamically tagged splitting can also be a medium-induced emission. The differential probability for these type of emissions, within a multiple scattering description

[4]Note that the factorized picture described in Fig. 2 is best suited for large-p_t jets and, therefore, the ALICE-like kinematics is not optimal.

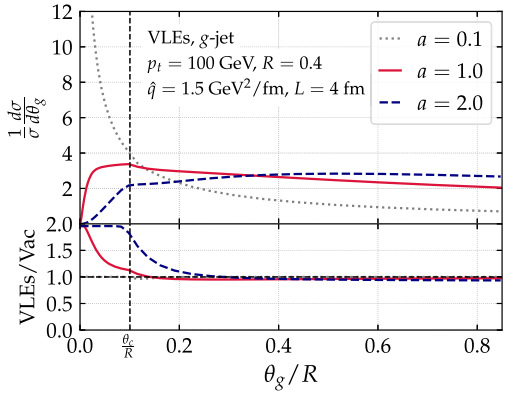


FIG. 3. θ_g distribution for gluon jets in medium including only vacuum-like emissions in the phase space of Fig. 2, for various values of the dynamical grooming parameter a. The lower panel shows the ratio to the pure vacuum expectation presented in Fig. 1. The veto region forbids some of the wide angle emissions and thus enhances emissions with small θ_g.

of the parton-medium interaction, is given by a convolution of the BDMPS-Z energy spectrum with the angular distribution $\mathcal{B}(z,\theta)$ produced via transverse momentum broadening, i.e.,

$$\frac{\mathrm{d}^2\tilde{\mathcal{P}}^{\mathrm{mie}}}{\mathrm{d}z\mathrm{d}\theta} \underset{\substack{k_\perp \ll Q_s \\ \omega \ll \omega_c}}{\approx} \bar{\alpha}_{s,\mathrm{med}} \sqrt{\frac{2\omega_c}{z^3 p_t}} \Theta(\omega_c - z p_t)\mathcal{B}(z,\theta), \quad (15)$$

where $\omega_c = \hat{q}L^2/2$ is the maximum energy that an emission can acquire as it corresponds to $t_f^{\mathrm{med}} = L$. In principle, the emission's energy is also bounded from below by the Bethe-Heitler frequency $\omega_{\mathrm{BH}} \propto \mu^4/\hat{q}$, where μ is an infrared regulator of the order of the Debye mass ($\sim$1 GeV). That said, these soft emissions are suppressed by the Sudakov form factor in DyG (Dynamically Groomed) observables and, consequently, this infrared physics is irrelevant. In this formula, the strong coupling constant $\alpha_{s,\mathrm{med}}$ should be evaluated at the typical transverse momentum scale of a medium-induced emission, $k_{\perp,\mathrm{med}}$. However, following the vacuum calculation, we shall consider $\alpha_{s,\mathrm{med}}$ as a constant parameter to be fixed, just like the other free parameters, by comparing our analytic model to the jet R_{AA} observable in a given $p_{t,\mathrm{jet}}$ window. Note that $\alpha_{s,\mathrm{med}}$ can be distinct from its vacuum counterpart α_s.

The factorization of the exact spectrum into the product of the time-averaged broadening distribution and the energy spectrum is only valid in the soft $\omega \ll \omega_c$ and collinear limit $k_\perp^2 \ll Q_s^2 = \hat{q}L$, i.e., for short-formation time emissions compared to the medium size. Such emissions can happen anywhere along the jet path length. Therefore, $\mathcal{B}(z,\theta)$ describes the transverse diffusion of the emission and is given by the average over the emission time $t \in [0, L]$ of a Gaussian distribution in $k_\perp \simeq \omega\theta$ with variance $\hat{q}(L-t)$, i.e.,

$$\begin{aligned}\mathcal{B}(z,\theta) &= \frac{1}{L}\int_0^L \mathrm{d}t \frac{2\omega^2\theta}{\hat{q}(L-t)} e^{-\frac{\omega^2\theta^2}{\hat{q}(L-t)}} \\ &= 2\theta\frac{z^2 p_t^2}{Q_s^2}\Gamma\left(0, \frac{z^2 p_t^2 \theta^2}{Q_s^2}\right) \end{aligned} \quad (16)$$

where $\Gamma(a,z) = \int_z^\infty \mathrm{d}t t^{a-1}\mathrm{e}^{-t}$ is the incomplete Gamma function. This distribution peaks at the transverse momentum scale Q_s.

Although not realistic from a physics point of view, let us consider a jet evolving via primary medium-induced emissions *only* (without VLEs), distributed according to Eq. (15). Then, the probability distribution for a medium-induced splitting to be the hardest in the shower is given by

$$\frac{\mathrm{d}^2\mathcal{P}^{\mathrm{mie}}}{\mathrm{d}z\mathrm{d}\theta} = \frac{\mathrm{d}^2\tilde{\mathcal{P}}^{\mathrm{mie}}(z,\theta)}{\mathrm{d}z\mathrm{d}\theta}\Delta^{\mathrm{mie}}(\kappa|a) \quad (17)$$

with the in-medium Sudakov form factor related to the medium-induced branching kernel as in Eq. (4). A straightforward calculation gives (for $\kappa < \omega_c/p_T$)

$$\begin{aligned}\ln\Delta^{\mathrm{mie}}(\kappa|a) &= -\int_0^1 \mathrm{d}z' \int_0^R \mathrm{d}\theta' \frac{\mathrm{d}^2 P^{\mathrm{mie}}}{\mathrm{d}z'\mathrm{d}\theta'}\Theta(z'(\theta'/R)^a - \kappa) \\ &= -\bar{\alpha}_{s,\mathrm{med}}\sqrt{\frac{2\omega_c}{p_T}}\int_\kappa^{\omega_c/p_T}\frac{\mathrm{d}z'}{z'^{3/2}} \\ &\quad\times\Bigg[z'^2\chi\Gamma(0,z'^2\chi) \\ &\quad-\left(\frac{\kappa}{z'}\right)^{2/a} z'^2\chi\Gamma\left(0,\left(\frac{\kappa}{z'}\right)^{2/a} z'^2\chi\right) \\ &\quad-\exp(-z'^2\chi) + \exp\left(-\left(\frac{\kappa}{z'}\right)^{2/a} z'^2\chi\right)\Bigg], \end{aligned} \quad (18)$$

with $\chi = Q^2/Q_s^2$ and $Q^2 = p_t^2 R^2$. Note that with Eq. (17), the normalization of $\mathrm{d}^2\mathcal{P}^{\mathrm{mie}}$ is not guaranteed. Indeed, when taking the limit $\kappa \to 0$ in Eq. (18) the Sudakov does not vanish, as it is the case for vacuum emissions, but rather tends to a constant. This difference arises from the absence of a collinear singularity in the medium-induced case. Therefore, in order to maintain the probabilistic interpretation of Eq. (17), one needs to divide by $1 - \Delta(0)$.

In Fig. 4, we represent the θ_g distribution computed with medium-induced emissions only. We observe how the small-θ_g behavior is strongly modified with respect to its vacuum counterpart (see Fig. 3) in the $a \to 0$ limit. This behavior is related to the absence of collinear singularities

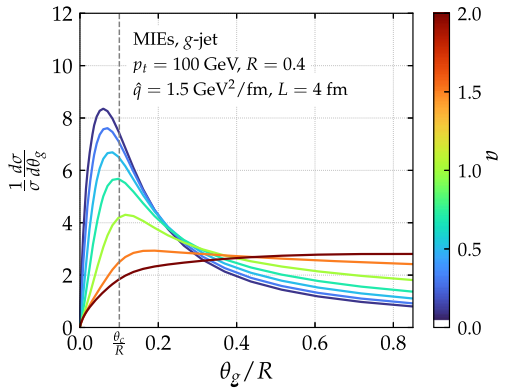


FIG. 4. θ_g distribution for gluon jets using only medium-induced emissions from Eq. (17). The small-θ_g behavior is dominated by the broadening term.

in the medium-induced branching kernel due to the broadening term.

Combining vacuum-like and medium-induced emissions. At this stage, we can construct the probability distribution for a splitting to be the hardest accounting for both vacuum-like and medium-induced emissions. Our formula is grounded on the factorization in time between the two types of processes that holds within the DLA. This factorization states that in-medium vacuum-like emissions occur first, in an angular-ordered way, followed by time-ordered medium-induced emissions [38]. Further, we impose the following pair of approximations:

(1) Transverse momentum broadening after the emission process is neglected for the in-medium vacuum-like splittings. This would shift the final value of θ_g by a typical angle of order Q_s/ω which is indeed negligible in the in-medium region above the line $k_\perp = Q_s$ in Fig. 2. Below this line, this approximation is less justified, and the effect of transverse momentum broadening in jet substructure observables deserves further studies.

(2) Only relatively hard, primary, medium-induced emissions remain inside the jet cone. In principle, these emissions trigger medium-induced cascades that rapidly develop a turbulent behavior leading to the multiplication of soft gluons with energy below the multiple branching scale $\omega_{\rm br} \sim \alpha_{s,\rm med}^2 \omega_c$ [67–69]. Our approximation is valid if these gluons (with $\omega \lesssim \omega_{\rm br}$) are deviated outside the jet cone. Since the typical angle of a gluon in the multiple branching regime is $\theta_{\rm br} \sim (\hat{q}/(\alpha_{s,\rm med}^2\omega^3))^{1/4}$ [70,71], the condition on the jet radius is $R \lesssim \theta_c/\alpha_{s,\rm med}^2$. For our choice of medium parameters, one gets $\theta_c/\alpha_{s,\rm med}^2 \sim 0.7$ which is indeed larger than the cone sizes studied here.

Under these approximations, the probability distribution for a splitting to be the hardest in the full shower can be written as

$$\mathrm{d}^2\mathcal{P}^{\rm med}(z,\theta|a) = (\mathrm{d}^2\tilde{P}^{\rm vac}(z,\theta)\Theta_{\notin\rm veto} + \mathrm{d}^2\tilde{P}^{\rm mie}(z,\theta)) \times \Delta_{\notin\rm veto}(\kappa|a)\Delta^{\rm mie}(\kappa|a). \quad (19)$$

The interpretation of the previous formula is quite transparent from a physical point of view. The tagged splitting can be either a vacuum or a medium-induced emission and, for both cases, one has to ensure that emissions of any type with a $\kappa' > \kappa$ are vetoed. We would like to remark that Eq. (19) can only be taken, at best, as a proxy for a realistic in-medium shower. The Lund plane density of the branching kernels in Eq. (19), i.e., $\mathrm{d}^2\tilde{P}^{\rm vac}(z,\theta)\Theta_{\notin\rm veto} + \mathrm{d}^2\tilde{P}^{\rm mie}(z,\theta)$, can be found in the left panel of Fig. 5. In this representation, the vacuum branching kernel is completely uniform except for the fact that it does not populate the veto region. In contrast, due to momentum broadening, medium-induced emissions have a typical transverse momentum of $k_\perp \sim Q_s$ in the multiple, soft scattering approximation and thus the enhancement is observed in Fig. 5 around this scale.

The difference between the toy shower and the vacuum result is shown in the right panel of Fig. 5. An interesting point to notice is that the average value of the tagged θ_g increases when including the medium-induced component. Indeed, at small θ_g, the MIEs cause the distributions to go to zero faster because of the absence of a collinear singularity in their emission kernel, leading to a depletion when compared to VLEs. The transition angle at which we observe an enhancement depends on the value of a: it will be below ($a \ll 1$) or above [$a \sim \mathcal{O}(1)$] the critical resolution angle θ_c. This is reflected in the lower panel of the plot where we clearly observe that the ratio between the toy shower and the vacuum result goes below 1 at an angle whose value increases with increasing values of a.

3. Energy loss

Up to now, we have ignored one of the main distinctive features of in-medium jet propagation, that is, jet energy loss. We have shown in the previous section that the angular distribution of medium-induced emissions is broader than the corresponding vacuum one. Therefore, a "vacuum" jet with a given cone R and transverse momentum p_{t0}, will lose energy due to MIEs radiated out of the cone, i.e., with $\theta > R$. The main effect of the large angle energy loss consists in a redistribution of jets with given values of z_g and θ_g due to differential energy loss. In other terms, the energy loss by a given jet triggered by a parton with initial transverse momentum p_{t0} depends on the z_g and θ_g values of the jet after evolution. As the hard spectrum tends to bias towards jets losing less energy than average, this

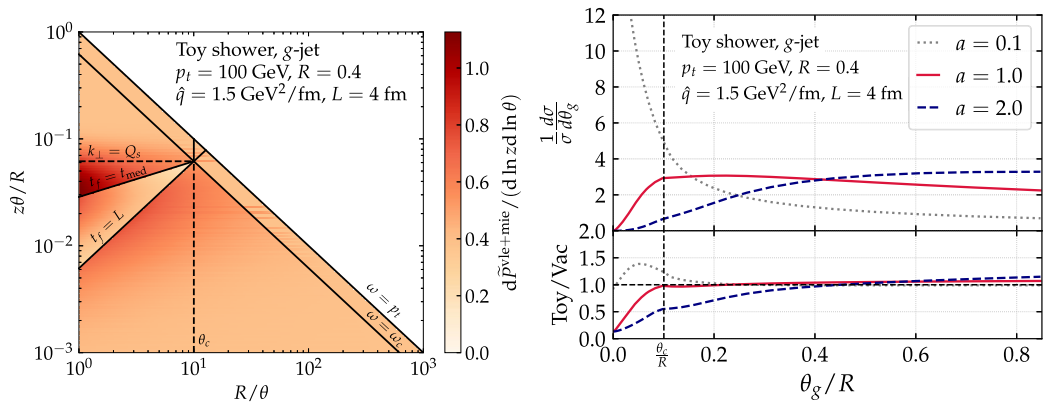


FIG. 5. Left: phase-space density of the branching kernel in the toy shower given by Eq. (19). It is interesting to notice that medium-induced emissions refill the veto region. Right: θ_g distribution for gluon jets using both vacuum-like and medium-induced emissions for various values of the DyG parameter a. The combined effect of VLE and MIE favors wider emissions.

differential energy loss redistributes the amount of jets tagged by (z_g, θ_g).

We now write a general formula that encompasses this idea. We call $\mathcal{E}_{i,p_{t0},R}(\varepsilon|z_g,\theta_g)$ the conditional probability for an i-initiated jet to radiate energy ε out of the jet cone R, *knowing that* the jet has a dynamically groomed hard branch with kinematic (z_g, θ_g). Then, the θ_g distribution for jets having a final transverse momentum p_t is

$$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\theta_g}\bigg|_{p_t} = \mathcal{N}_{\rm med}^{-1}\int \mathrm{d}\varepsilon \sum_{i\in\{q,g\}} \frac{\mathrm{d}\sigma_i}{\mathrm{d}(p_t+\varepsilon)} \times \int \mathrm{d}z_g \mathcal{P}_i^{\rm med}(z_g,\theta_g)\mathcal{E}_{i,p_t,R}(\varepsilon|z_g,\theta_g), \quad (20)$$

where $\mathrm{d}\sigma_i$ is the cross section for producing a jet with flavor i whose extraction is discussed in Appendix B. Further, $\mathcal{N}_{\rm med}$ is a normalization factor given by

$$\mathcal{N}_{\rm med}(p_t) = \sum_{i\in\{q,g\}}\int \mathrm{d}\varepsilon \frac{\mathrm{d}\sigma_i}{\mathrm{d}(p_t+\varepsilon)}\mathcal{E}_{i,p_t,R}(\varepsilon), \quad (21)$$

since using the law of total probability,

$$\mathcal{E}_{i,p_t,R}(\varepsilon) \equiv \int \mathrm{d}\theta_g \mathrm{d}z_g \mathcal{E}_{i,p_t,R}(\varepsilon|z_g,\theta_g)\mathcal{P}_i^{\rm med}(z,\theta_g), \quad (22)$$

where $\mathcal{E}_{i,p_{t0},R}(\varepsilon)$ is the probability for an i jet to lose an energy ε without any knowledge of its substructure. Notice that we use $\mathcal{P}^{\rm med}$ in Eq. (20), i.e., we quench not only vacuum-like emissions, but also intrajet medium-induced ones. Physically speaking, $\mathcal{N}_{\rm med}$ corresponds to the jet cross section.

Given the steeply falling nature of the jet spectrum, i.e., $\mathrm{d}\sigma/\mathrm{d}p_t \sim p_t^{-n}$ with $n \gg 1$, one can write $\mathrm{d}\sigma/\mathrm{d}(p_t+\varepsilon) \approx \mathrm{d}\sigma/\mathrm{d}p_t \exp(-\frac{n\varepsilon}{p_t})$ such that Eq. (20) becomes

$$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\theta_g}\bigg|_{p_t} = \mathcal{N}_{\rm med}^{-1}\sum_{i\in\{q,g\}}\frac{\mathrm{d}\sigma_i}{\mathrm{d}p_t}\int \mathrm{d}z_g \mathcal{P}_i^{\rm med}(z_g,\theta_g) \times \int \mathrm{d}\varepsilon \mathcal{E}_{i,p_t,R}(\varepsilon|z_g,\theta_g)e^{-\frac{n\varepsilon}{p_t}}. \quad (23)$$

The last line of the previous equation is the Laplace transform of the conditional energy loss probability.

Next, we need to specify the energy loss probability distribution $\mathcal{E}_{i,p_t,R}(\varepsilon|z_g,\theta_g)$. In the double logarithmic approximation, the jet is dominated by the hardest emission—the one tagged by dynamical grooming—and is thus made of two subjets. Neglecting the intrajet multiplicity of the subjets, their energy loss probability can be approximated by that of a single parton[5] with flavor i, out of a cone with opening R, denoted by $P_{i,R}^{(1)}(\varepsilon)$. In terms of $P^{(1)}$, the energy loss probability of this two-prong system can be written as

$$\mathcal{E}_{i,p_t,R}(\varepsilon|z_g,\theta_g) = (1-\Theta_{\rm res}(z_g,\theta_g))P_{i,R}^{(1)}(\varepsilon) + \Theta_{\rm res}(z_g,\theta_g) \times \int_0^\infty \mathrm{d}\varepsilon_1 \int_0^\infty \mathrm{d}\varepsilon_2 P_{i,R}^{(1)}(\varepsilon_1)P_{g,R}^{(1)}(\varepsilon_2)\delta(\varepsilon-\varepsilon_1-\varepsilon_2), \quad (24)$$

where the resolution condition reads

[5]We do not take into account the fact that the opening angles of the two subjets are different from R and depend on θ_g.

$$\Theta_{\rm res}(z_g,\theta_g)=\Theta(\theta_g-\theta_c)\Theta(z_g\theta_g p_t-k_{\perp,\rm med}), \quad (25)$$

and it selects splittings in the red region of Fig. 2. Then, Eq. (24) simply states that if the two-prong system is resolved by the medium, the jet energy loss is the sum of the energy losses of each subjet. On the other hand, if the two prongs are not resolved, the full jet loses energy as a single subjet with the color charge of its initiator. Plugging Eq. (24) into Eq. (23) yields

$$\left.\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\theta_g}\right|_{p_t}=\mathcal{N}_{\rm med}^{-1}\sum_{i\in\{q,g\}}\frac{\mathrm{d}\sigma_i^h}{\mathrm{d}p_t}\int \mathrm{d}z_g\mathcal{P}_i^{\rm med}(z_g,\theta_g) \times[(1-\Theta_{\rm res})Q_i(p_t,R)+\Theta_{\rm res}Q_g(p_t,R)Q_i(p_t,R)], \quad (26)$$

where we have defined

$$Q_i(p_t,R)\equiv\int_0^\infty \mathrm{d}\varepsilon P_{i,R}^{(1)}(\varepsilon)\exp\left(-\frac{n\varepsilon}{p_t}\right). \quad (27)$$

The last step is to find an approximation for the function $P_{i,R}^{(1)}(\varepsilon)$ or equivalently, the quenching weight $\mathcal{Q}_i(p_t,R)$. Neglecting the intrajet activity of the subjet, $P_{i,R}^{(1)}(\varepsilon)$ can be approximated by the energy loss probability distribution of a single parton of flavor i out of a cone of size R. Evaluating the Laplace transform, we arrive at the well-known expression for the quenching weight [72,73]:

$$Q_i(p_t,R)=\exp\left[\int_R^\infty \mathrm{d}\theta\int_0^1 \mathrm{d}z\frac{\mathrm{d}^2\tilde{P}^{\rm mie}}{\mathrm{d}\theta\mathrm{d}z}(\mathrm{e}^{-\frac{n\omega}{p_t}}-1)\right]. \quad (28)$$

At this point, an important remark is in order. We have argued that $\mathrm{d}^2\tilde{P}^{\rm mie}$ accurately describes the intrajet medium-induced activity. Then, at first glance, it might seem contradictory to use this very same branching kernel to estimate the number of gluons that are deviated outside the jet cone. The physical reason behind this apparent contradiction was presented in Refs. [67,74] (see Ref. [69] for a review). It is related to the turbulent behavior of the medium-induced cascade that efficiently degrades the initial energy into very soft quanta. This turbulent cascade has a fixed point which is identical to the BDMPS-Z spectrum that gives the z dependence of $\mathrm{d}^2\tilde{P}^{\rm mie}$ and explains, *a posteriori*, why Eq. (28) is a good estimation.

As argued previously, the typical angle of soft gluons in the multiple branching regime is $\theta_{\rm br}(\omega)$. Therefore, the criterion for a medium-induced gluon to be deviated out of the jet cone is $\theta_{\rm br}(\omega)>R$ or $Q_s/\omega>R$. The latter condition corresponds to the case of a relatively hard emission with $\omega>\omega_{\rm br}$. For the values of R we consider, the second condition overwhelms the first one, so that we can safely approximate the angular dependence of $\mathrm{d}^2\tilde{P}^{\rm mie}$ by $\delta(\theta-Q_s/\omega)$. Then, the quenching weight reads

$$\ln\mathcal{Q}_i(p_t,R)=\frac{2\alpha_{s,\rm med}C_i}{\pi}\sqrt{\frac{2\omega_c}{\omega_{\rm max}}}(1-\sqrt{\pi\nu\omega_{\rm max}} \times\mathrm{Erf}(\sqrt{\nu\omega_{\rm max}})-e^{-\nu\omega_{\rm max}}), \quad (29)$$

with $\omega_{\rm max}=\min(Q_s/R,\omega_c)$ and $\nu=n/p_t$. Note that recent works have gone beyond the single parton energy loss picture for global observables by resuming the effects of the fluctuating substructure on the total energy loss [58,59,75]. We will extend that formalism to jet substructure observables in a separate publication [50].

We have checked that this quenching weight gives reasonable values for the R_{AA} ratio of jet cross sections with our choice of medium parameters. As alluded to above, the jet cross section in Pb-Pb is given by $\mathcal{N}_{\rm med}$, which can be obtained either from Eqs. (21)–(22) or from Eq. (26) thanks to the self-normalization of the θ_g distribution. In both cases, we observe a mild a dependence of this jet cross section, as a consequence of the main underlying approximation of Eq. (24), namely the fact that we neglect the vacuum-like multiplicity of the resolved or unresolved subjets. This uncertainty in the jet cross section is harmless for the shape of the θ_g distributions, which are self-normalized by construction.

In Fig. 6, the quenched θ_g distributions are displayed. The most remarkable feature of these distributions is the keen transition at $\theta_g=\theta_c$. This arises due to several reasons that we proceed to analyze. To start with, our energy loss model, i.e., Eq. (24), contains a sharp distinction between resolved and unresolved splittings that translates into branchings with $\theta\geq\theta_c$ losing more energy than those with $\theta\leq\theta_c$. Then, the steeply falling nature of the spectrum drastically reduces the possibilities of these wide angle branchings that lost a substantial amount of energy to end up in the selected p_t window. That is, the

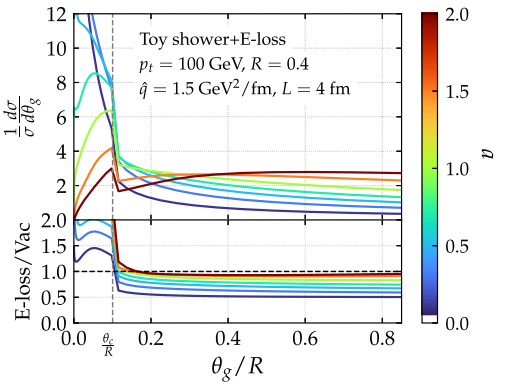


FIG. 6. θ_g distribution obtained with Eq. (26). Including differential energy loss results in a sharp transition in the distributions at θ_c.

least quenched jets, i.e., those splittings with $\theta \leq \theta_c$ (the green region in Fig. 2), dominate the θ_g distribution and therefore a narrowing is expected. This is a well-known effect typically referred to as "selection bias" or "filtering effect"; see e.g., Refs. [59,76–78] for a possible way out. This feature also explains the a dependence of the ratio in the bottom panel. In fact, we have shown in Eq. (8) that the maximum of the θ_g distribution for vacuum splittings is inversely proportional to a. This estimate is still correct when including medium-induced emissions. Therefore, when $a \geq 1$ the probability of tagging a wide splitting is larger than selecting a narrow one. However, once energy loss is included, those very few narrow splittings will lose significantly less energy than their wide angle counterparts. Consequently, their probability is enhanced with respect to vacuum and the ratio goes above one. On the other hand, when $a \leq 1$ small angle splittings are typically selected. Following the same reasoning, the very few large angle splittings will be even more suppressed than in vacuum due to incoherent energy loss.

Finally, we would like to comment on the effect of the quark/gluon fraction resulting from the sum over flavors in Eq. (26). In the vacuum, we expect quark-initiated jets to have a narrower θ_g distribution than gluon-initiated jets [see e.g., Eq. (8)]. Therefore, the θ_g distribution is also sensitive to the different quark-gluon fraction of the hard spectrum in Pb-Pb collisions compared to pp, but as shown in Appendix B, the overall effect is very mild. That said, once large angle jet energy loss is included, since quark jets lose less energy than gluon jets, we expect a filtering effect towards quark-initiated jets, leading to an even narrower θ_g distribution. This effect is accounted for in our analytic calculation and in Fig. 6. To disentangle these two filtering effects, i.e., (i) towards coherent, "unresolved" jets and (ii) quark-initiated jets, an interesting possibility is to measure the θ_g distribution in $\mathrm{Z}/\gamma + \mathrm{jet}$ events [59,76,79].

4. Path-length fluctuations

So far, we assumed that the medium is a homogeneous brick of fixed length. In this section we discuss how to extend our toy theoretical model to account for the fact that in a realistic heavy-ion collision the hard scattering that produces the jet can take place anywhere inside the geometric overlap area between the two colliding nuclei.

Note that our homogenous brick model also ignores the rapid expansion of the medium and the fluctuations of $\hat{q}$ along different path lengths. Regarding the expansion of the medium, previous studies showed that a simple Bjorken-like expansion of the medium is well captured by rescaling the jet quenching parameter $\hat{q}$ of a homogeneous brick [80,81], $\hat{q} \equiv \|\hat{q}(t)\|_{1/2}$, where $\|f(t)\|_{1/2}$ stands for the $1/2$-norm of the function $f(t)$ with compact support. This scaling is a consequence of the local nature of the medium-induced emissions in the multiple soft scattering regime $\omega \ll \omega_c$.[6] Since we do not consider medium-induced emissions harder than ω_c, we invoke this scaling to extend our results to the Bjorken expansion case. Beyond the purely longitudinal expansion scenario, a simple rescaling of $\hat{q}$ does not capture the medium dynamics. Overall, the impact of a more realistic medium description will be studied numerically in Sec. III.

We point out that this scaling applies for the medium-induced emission process for which we can invoke the argument of locality. It is not the case for θ_c, since this angular scale comes from the decoherence of a color singlet dipole traveling through the medium over a distance L. However, θ_c also obeys an approximate scaling law that relates static and expanding medium. This scaling law is tantamount to replacing $\hat{q} \to \hat{q}(L)$ in the definition of θ_c [81]. For a Bjorken expansion, this rescaling of $\hat{q}$ differs by a factor of 2 with respect to the rescaling of $\hat{q}$ that describes medium-induced emissions. Since in the DLA, we do not control overall prefactors, we decided not to study this alternative scaling for the VLE phase space in our qualitative analysis.

To capture the fluctuation in the path length of the jet for central collisions, we propose the following model: (i) the interaction region is approximated by a circle of radius $R = 4$ fm around the center of the collision, (ii) random (x, y) coordinates of hard scatterings are sampled uniformly in the interaction region, (iii) each creation point is connected with a hard-scattering leading-order matrix element from Pythia8 [83] (Monash13 tune [84]) in proton-proton collisions, assigning the 4-momenta of the outgoing legs and (iv) the path lengths are determined by the intersection of the path with the edge of the interaction region. The distribution of the resulting path lengths is shown in the left panel of Fig. 7, centered around 4 fm; however, $\langle L \rangle = 3.75$ fm due to the asymmetry of the distribution. Even though this model is overly simplistic, it is sufficient to qualitatively understand the effects of the path length fluctuations on the θ_g distribution. More precise phenomenology would require accounting for the nuclear thickness function and the precise shape of the interaction region across various centrality classes similar to Ref. [58].

The θ_g distribution obtained with the medium-induced branching kernel [see Eq. (17)] and a fluctuating path length is presented in the right panel of Fig. 7. The ratio to

[6]This scaling is therefore distinct from the one discovered in Ref. [82] that works for processes dominated by the most energetic medium-induced emissions ($\omega \sim \omega_c$). As shown in Ref. [81], it is also violated by VLEs via a change of the phase space boundaries that we neglect in this study.

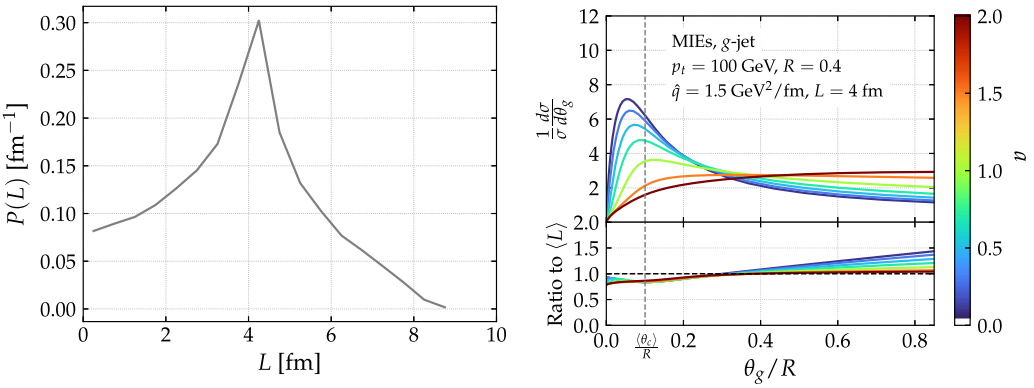


FIG. 7. Left: probability distribution of the fluctuating jet path length in a simplified scenario described in the main text. Right: θ_g distribution for various values of the dynamical grooming parameter a when including medium-induced emissions only and path-length fluctuations. The fluctuations smoothen the sharp transition around θ_c. The asymmetric nature of the jet path length distribution results in a nonflat ratio.

the average L result is displayed in the bottom panel. The enhancement of large θ_g values is rooted in the asymmetric nature of the path-length distribution; see left panel of Fig. 7. More concretely, shorter than average path lengths are more probable. This automatically translates into a distribution of θ_c values that tend to be larger than average due to the $\theta_c \propto L^{-3/2}$ scaling. Consequently, the θ_g distribution gets broader when path length fluctuations are included.

C. Final theoretical results

Finally, we present our theoretical curves including all the ingredients discussed in the previous paragraphs in Fig. 8. Compared to Fig. 6, we observe that the main effect of introducing the path length fluctuations is to smoothen the transition (peak) around the critical angle θ_c. Consequently, the peak of the medium modified θ_g distribution is shifted towards slightly smaller values of the opening angle; however the peak still persists. It is easy to observe by eye that there are choices of the dynamical grooming parameter a which enhance the relative difference between the medium and the vacuum θ_g distributions. For $a \sim 2$, we do not see a significant deviation, whereas values of a close to 1 give a pronounced peak around the mean value of θ_c which is not present in the vacuum distribution. Therefore, we expect that measuring the dynamically groomed jet radius with $a \sim 1$ will provide clear evidence of the existence of a characteristic (de) coherence angle.

In order to gauge the sensitivity of the θ_g observable to medium physics in a more quantitative way, we choose the Kolmogorov-Smirnov metric as a measure of the differences between the vacuum and in-medium distributions. The Kolmogorov-Smirnov (KS) distance,[7] $\mathcal{D}$, is defined as

$$\mathcal{D} = \max_{0 \le \theta_g \le R} |\Sigma^{\text{vac}}(\theta_g) - \Sigma^{\text{med}}(\theta_g)|, \tag{30}$$

where Σ denotes the cumulative distribution

$$\Sigma(\theta_g) = \int_0^{\theta_g} \mathrm{d}\theta' \frac{1}{\sigma} \frac{\mathrm{d}\sigma}{\mathrm{d}\theta'}. \tag{31}$$

That is, the KS metric corresponds to the maximal distance between the cumulated spectra. The larger $\mathcal{D}$ is, the more distinct the two distributions are and, consequently, the larger the discriminating power of θ_g is. We choose to use this more involved metric instead of the usual ratio because of the strong differences in shape between the medium and vacuum distributions. In Fig. 9, we display the value of the Kolmogorov-Smirnov distance resulting from our analytic calculation of the θ_g distribution for several values of the grooming parameter a. We observe a nonmonotonic behavior with a and the largest distance corresponds to $a = 0.3$. The other interesting feature of this plot, which we shall also observe in Monte Carlo simulations, is the reduction of the Kolmogorov-Smirnov distance once path length fluctuations are included, as a consequence of a smoother transition between coherent and incoherent subjet energy loss.

To summarize and conclude this analytic section, we emphasize that our pQCD-motivated theoretical model,

[7]The authors would like to thank Marta Verweij for suggesting this metric in a different context.

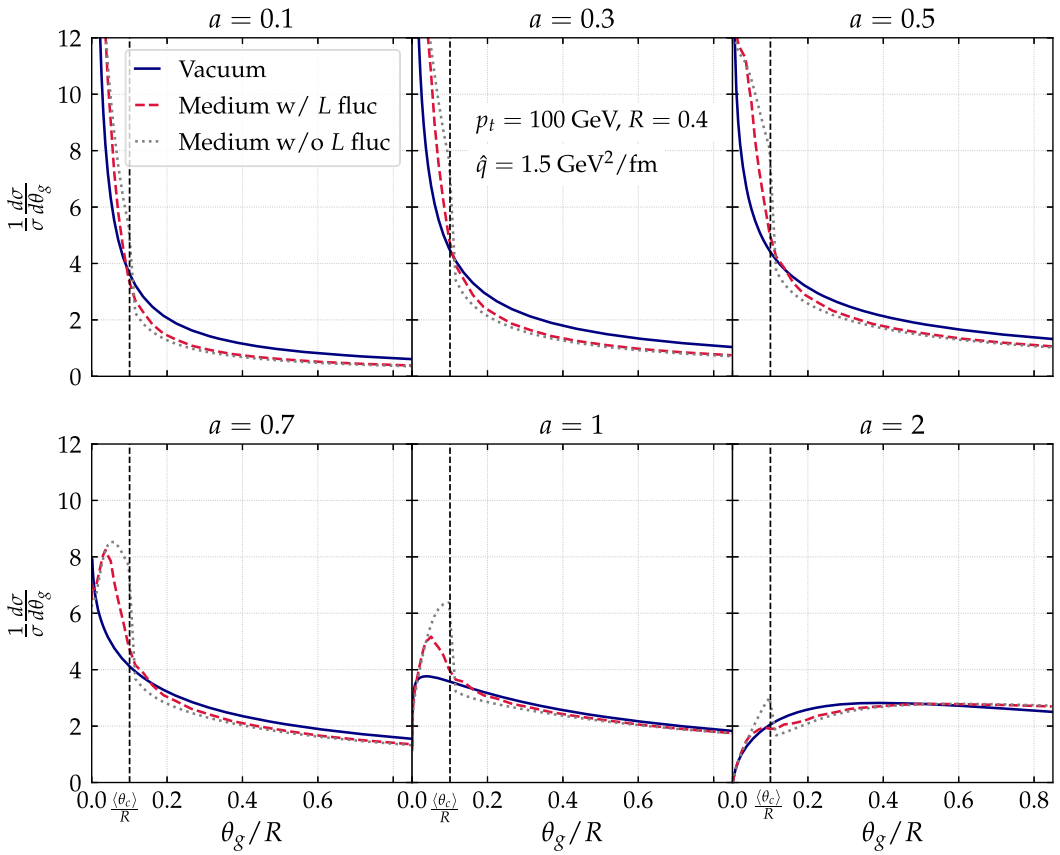


FIG. 8. θ_g distribution for the toy model given by Eq. (26) for various values of the DyG parameter a in vacuum (solid, blue) and in the medium with (dashed, red) and without (dotted, gray) path length fluctuations.

relies on the factorization between VLEs and MIEs in the DLA and the multiple soft scattering approximation. We predict a significant modification of the θ_g distribution around the critical angle θ_c, as a consequence of multiple imprints of different physical mechanisms. The most dominant effect is the filtering towards fully coherent jets. The second important result is that, by pursuing an analytic approach, we are able to provide theoretical guidance on the optimal values of the DyG parameter a that maximize the discrimination power of this observable in order to probe the physics of color (de)coherence experimentally.

III. MONTE CARLO SIMULATIONS

In this section, we numerically explore the θ_g distribution with three state-of-the-art jet quenching Monte Carlo codes: JetMed [38], the Hybrid model [85] and Jewel [86,87]. In all cases, we generate dijet events at $\sqrt{s} = 5.02$ TeV in Pb + Pb collisions.[8] For each event, particles are clustered on an event-by-event basis into anti-k_t jets [88] with $R = 0.4$ and reclustered with the Cambridge/Aachen [89] algorithm to obtain an angular-ordered clustering sequence. The analysis is performed on jets with transverse momenta $75 < p_t < 100$ GeV and rapidities $|y| < 1$.

A. JetMed

To begin with, we present results for the Monte Carlo framework that is closest in spirit to the semianalytic model presented in the previous section. The Monte Carlo JetMed

[8]Actually, we only generated the JetMed events by ourselves. We have obtained the Jewel samples from https://jetquenchingtools.github.io, while the Hybrid events have been kindly provided by Daniel Pablos.

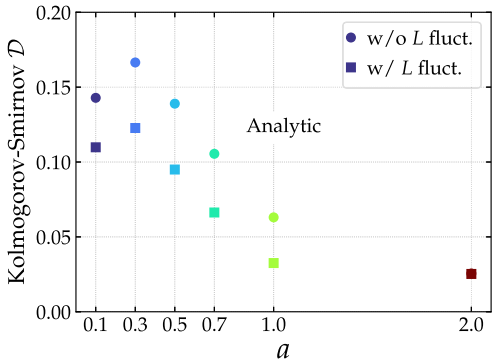


FIG. 9. Kolmogorov-Smirnov distance defined in Eq. (30) as a function of the dynamical grooming parameter a for the theoretical results with (squares) and without (circles) jet path length fluctuations. The bigger the KS value, the easier it is to discriminate between vacuum and medium distributions.

is a parton shower based on the factorization between vacuum-like emissions and medium-induced emissions that holds in the double-logarithmic approximation for the former and multiple soft scattering approximation for the latter. The main differences to be expected between the analytic approach and the numerical results concern (i) the inclusion of part of the single logarithmic corrections to the vacuum-like shower, through the running of the QCD coupling and the hard collinear emissions, (ii) the proper resummation of the medium-induced emissions with formation time $t_f^{\rm med} \ll L$ in the multiple branching regime, (iii) the relaxation of the quenching weight approximation since the jet energy loss is provided for free in a parton shower approach and (iv) accounting for the transverse momentum broadening after emission that leads to a shift in the final θ_g value of the tagged subjet. Notice that we have extended the original code to include jet path length fluctuations using the same model of the geometry as in the analytics.

The resulting θ_g distributions for JetMed are displayed in Fig. 10. Let us start the discussion with the vacuum curves. A clear quantitative discrepancy at small θ_g exists between them and the analytic ones presented in Fig. 1. The main source of this difference is the choice of fixed coupling in the analytic result that moves the vacuum peak towards smaller θ_g. Including running coupling corrections in the analytic calculation is relatively simple; however we opt to not do it since the logic of this paper is to present results at double logarithmic accuracy. Regarding the medium, we observe the same trends as in the analytic calculation: for $a \leq 1$ they are strongly peaked at the average critical angle θ_c, the relevant angular scale in the problem. The peak around θ_c becomes broader when including jet path length fluctuations, similar to what is observed in the analytical results in Fig. 8. An important observation is that the medium curves are overall shifted towards larger angles, due to transverse momentum broadening, while they are still peaked around $\langle \theta_c \rangle$ for $a < 1$.

The Kolmogorov-Smirnov distance between the vacuum and medium distributions is shown in the left panel of Fig. 11. Through this metric we confirm the analytic observation of a nonmonotonic behavior of the KS distance with a. The higher-order corrections included in JetMed slightly shift the value of a at which the maximal KS distance is achieved, i.e., $a = 0.3$ at the DLA and $a = 0.5$ in JetMed. In addition, path-length fluctuations also reduce the discriminating power of this observable as was the case in the analytic calculation.

In order to enhance the sensitivity to jet quenching effects a rather small value of a, i.e., $a \leq 1$, must be chosen. However, one should keep in mind that the lower a is, the larger the nonperturbative corrections to the vacuum distribution are [47]. Hence, there is a trade-off between maximizing the KS distance and minimizing the impact of hadronization corrections. From Fig. 11, we conclude that values of a between ∼0.5 and ∼0.7 fulfill these requirements.

We would like to remark that this optimization exercise is relatively simple in the case of dynamical grooming given that it has a single free parameter and, as noted in Ref. [44], the θ_g distribution is invariant under the $a \to 1/a$ transformation. This last point immediately reduces the range of a values to scan. Obviously, one could also calculate $\mathcal{D}$ in the two-dimensional parameter space spanned by the SoftDrop condition. However, not only the increased dimensionality but also the possible degeneracy between pairs of $(z_{\rm cut}, \beta)$ complicate the analysis.

B. Strong vs weak coupling approach

Throughout this manuscript, we have considered a very specific model of the in-medium shower founded on two basic pillars: a weak coupling description of the medium and the multiple soft scattering approximation. Then, a natural question to ask is: do the observed features of the θ_g distribution arise only in this model or are these features general enough so that any model with some notion of angular dependent energy loss leaves the same footprints in the θ_g distribution? To address this point we make use of the Hybrid model. It is beyond the scope of this paper to provide a thorough description of all the ingredients assembled in this code, but we would like to highlight some of the differences in the phase space of emissions with respect to the discussion surrounding Fig. 2.

In the strong coupling description of the medium, the existence of a finite resolution length was first considered in Ref. [90]. However, this property does not appear naturally as it is the case in the weak-coupling scenario,

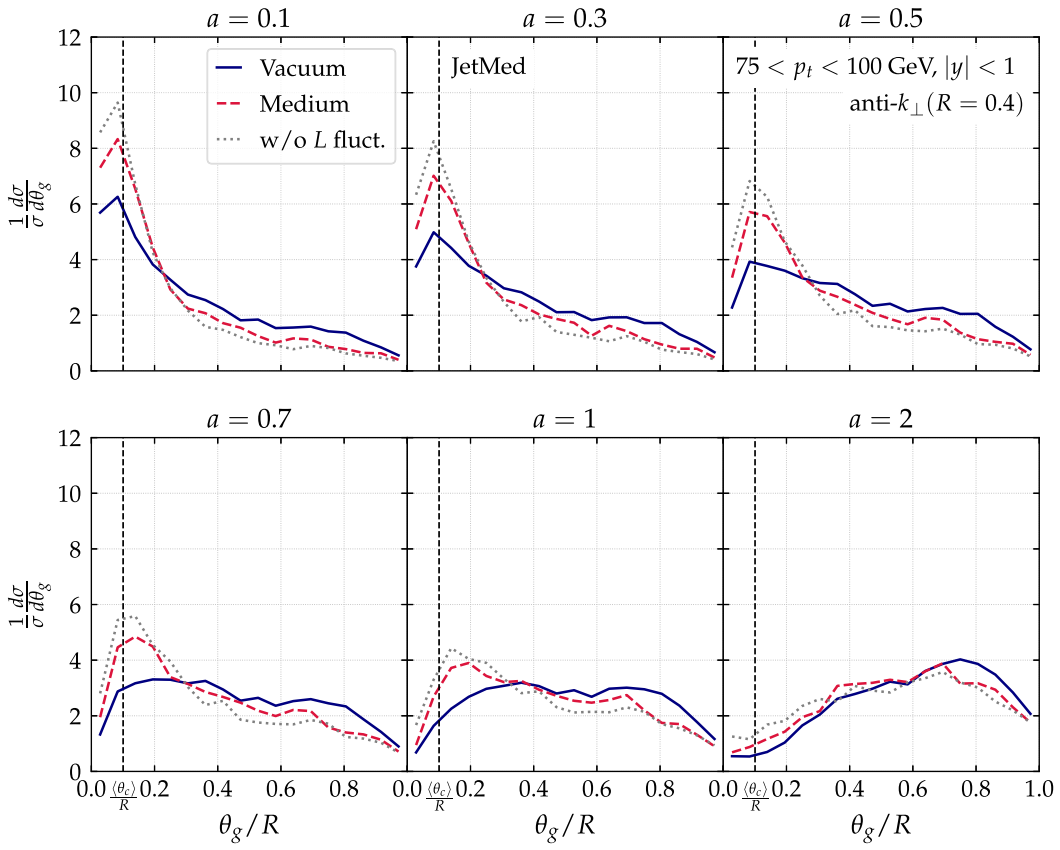


FIG. 10. θ_g distribution for various values of the DyG parameter a with JetMed in vacuum (solid, blue) and in the medium with (dashed, red) and without (dotted, gray) jet path length fluctuations. The medium distribution exhibits a peak around the coherence angle θ_c.

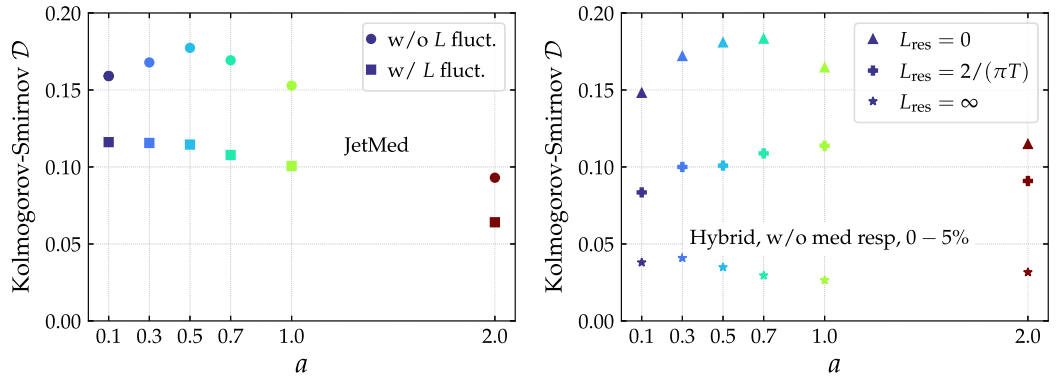


FIG. 11. Kolmogorov-Smirnov distance defined in Eq. (30) as a function of the dynamical grooming parameter a with the JetMed parton shower (left) and the Hybrid model (right). The weak and strong coupling descriptions of the jet-medium interaction are in qualitative agreement.

but rather has to be introduced by hand as a free parameter called $L_{\rm res}$ in the Hybrid model. From a physics point of view, $L_{\rm res}$ corresponds to the minimal distance between two color charges such that they interact with the medium independently. This resolution length is inversely proportional to the local temperature of the medium, the only scale in the problem, and the coefficient of proportionality cannot be computed from first principles but only estimated. In this work we explore three different values of $L_{\rm res}$. The fully incoherent case is considered by setting $L_{\rm res} = 0$. The opposite scenario in which the jet is treated as a single object corresponds to $L_{\rm res} = \infty$. Finally, we take an intermediate value of $L_{\rm res} = 2/(\pi T)$, with T being the local temperature of the plasma. In spite of sharing the same concept of "resolution of colored prongs by the plasma", we would like to emphasize that $L_{\rm res}$ and θ_c are intrinsically different, but lead to qualitatively similar bias effects. Notably, there is no critical angular scale in the Hybrid model, and therefore there is no angular selection bias as in the weak coupling picture discussed so far. The effect of a finite $L_{\rm res}$ is simply to increase the amount of energy lost by jets due to an increase of resolved sources which undergo strong coupling energy loss. This induces a selection bias in a similar way to JetMed, but for a different reason. Another important difference with respect to JetMed is the fact that no medium-induced branching kernel exists in the Hybrid model, i.e., the splitting probability is the same as in vacuum. On top of that, we consider hadronized samples and switch off the medium response for the purpose of this section.

We show the dynamically groomed jet radius distributions in Fig. 12 for the aforementioned values of the screening length $L_{\rm res}$. There are two cases for which no angular scale is present in the energy loss mechanism: $L_{\rm res} = 0$ and $L_{\rm res} = \infty$. Clearly, the $L_{\rm res} = \infty$ results

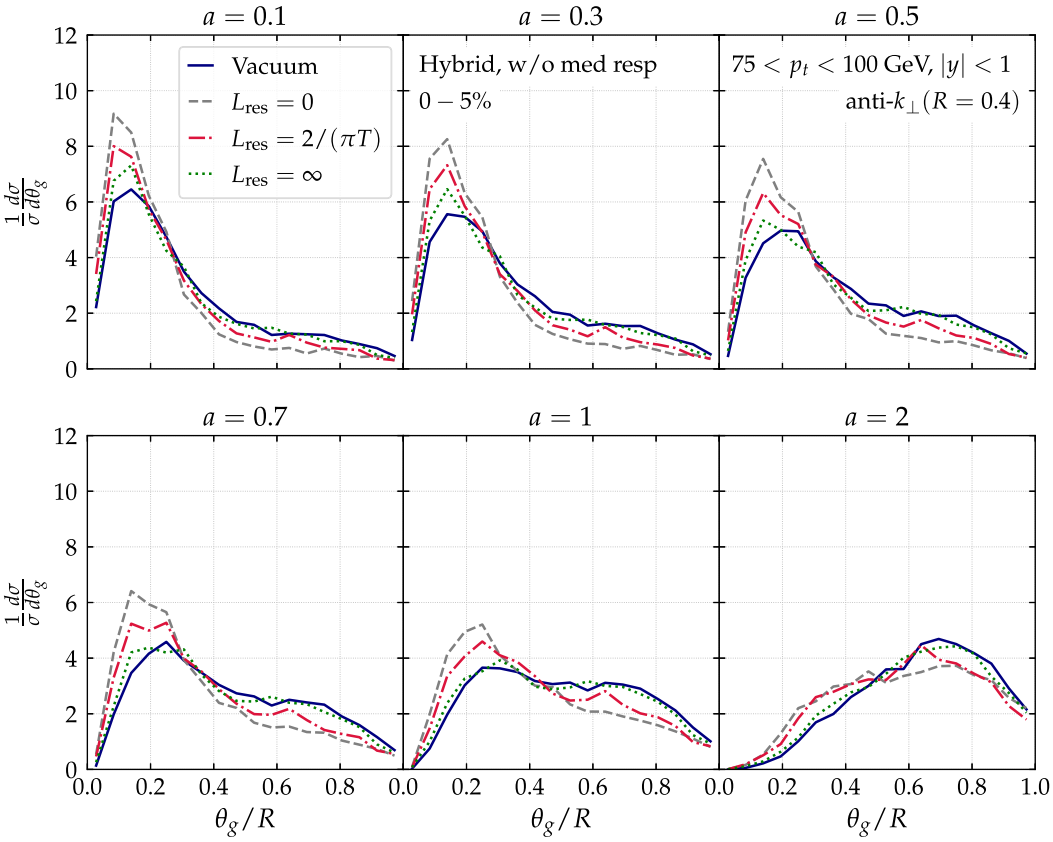


FIG. 12. θ_g distribution for various values of the DyG parameter a with the Hybrid and no medium response. The fully coherent energy loss ($L_{\rm res} = \infty$) scenario closely resembles the vacuum distribution.

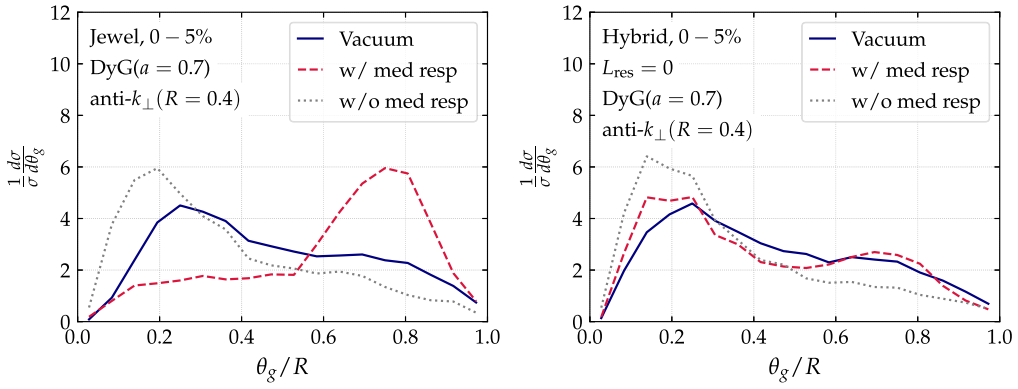


FIG. 13. θ_g distribution for the Jewel (left) and the Hybrid (with $L_{\rm res} = 0$) models including medium response for $a = 0.7$ in the DyG condition. The medium response introduces an enhancement of partons towards wider angles.

closely resemble the vacuum distributions. The interpretation of this result is quite transparent: in the $L_{\rm res} = \infty$ case only the parent parton loses energy and therefore the filtering effect is drastically reduced. Naturally, the orthogonal scenario where all individual splittings are resolved by the medium, i.e., $L_{\rm res} = 0$, exhibits the biggest modification with respect to the vacuum benchmark. The intermediate case of $L_{\rm res} = 2/(\pi T)$ is in quantitative agreement with the JetMed results. This allows us to conclude that the proposed observable is agnostic to the fine details of the energy loss mechanism or the physical origin of the resolution scale. If measured experimentally, this observable has enough discriminatory power to discard (or confirm) a fully coherent scenario ($L_{\rm res} = \infty$). However, it would be hard to disentangle the $L_{\rm res} = 0$ and $0 \leq L_{\rm res} \leq \infty$ cases. To that end, further constraints on the model from other observables are required, e.g., fit the parameters to describe the R_{AA} and predict the θ_g distribution.

Notice, that, in contrast to JetMed, this Monte Carlo uses a realistic hydrodynamical profile. Remarkably, the narrowing of the distribution is not washed out by either the medium expansion or its geometrical fluctuations. This is confirmed by the values of the Kolmogorov-Smirnov metric shown in the right panel of Fig. 11. That said, we observe that the optimal value of a shifts from $0.5 \lesssim a \lesssim 0.7$ in JetMed to $0.7 \lesssim a \lesssim 1$ in the Hybrid model with $L_{\rm res} = 2\pi/T$.

C. Impact of medium response

The final effect that we would like to quantify is the impact of the medium backreaction in our proposed observable. To that end, we make use of two independent models of medium response: the ones implemented in the Jewel and Hybrid models. Technically, in order to avoid double counting of the thermal particles' momenta, we implemented the GridSub1 method described in Ref. [8] for Jewel[9] and the background subtraction technique presented in Appendix A of Ref. [39] in the Hybrid case. Further, in both models hadronization is switched on. One last remark before presenting the results is that in Jewel the radiation for unresolved emissions is not considered. Then, there is no coherence angle θ_c in this model and the θ_g distribution would be mainly sensitive to the filtering effect due to the finite size of the medium. In some sense, it is closest to the $L_{\rm res} = 0$ in the Hybrid calculation that we have presented above.

In Fig. 13, we present the θ_g distributions for the optimal value of a according to the Kolmogorov-Smirnov tests performed in the previous section. Further, in the case of the Hybrid model, we fix $L_{\rm res} = 0$ given that it showed the biggest difference with respect to the vacuum baseline. First of all, the vacuum curves between the two models are in quantitative agreement as expected since they are both based on the Pythia parton shower. The quenched curves without medium response of the two jet quenching Monte Carlo generators also match. However, once medium response is taken into account the two results differ. Qualitatively, an enhancement of wide-angle splittings is observed. Since particles originated from the medium backreaction are inherently soft, they can only affect this DyG observable if they appear at large enough angles. Their contribution is sizably different in the explored models. More concretely, in the Jewel case medium response completely distorts the shape of the distribution and creates a bump at the edge of the jet cone, while in the Hybrid case the impact is more

[9]We use a grid resolution value of 0.05. We have checked that our results do not change when varying this parameter up and down by a factor of 2.

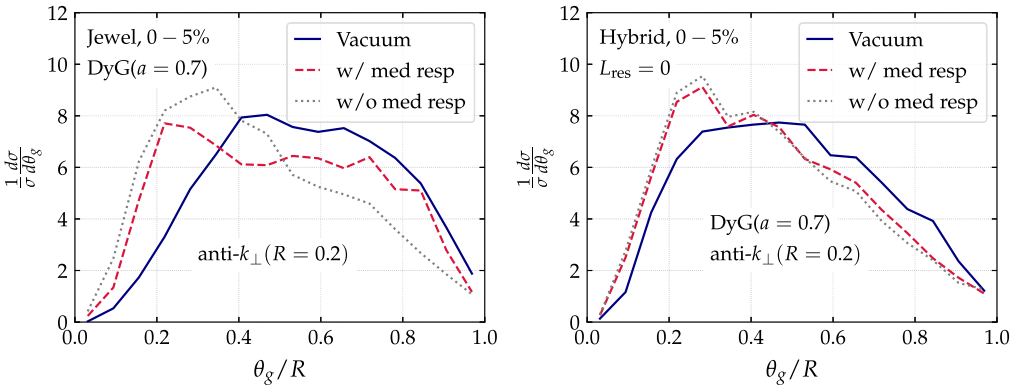


FIG. 14. Same as Fig. 13 but with $R = 0.2$ showing a significant reduction of wide angle emissions originated from medium response.

moderate but brings the medium and vacuum distributions closer. The signal created by the wake particles clearly pollutes the interpretation of the KS distance in terms of probing $L_{\rm res}$. Turning the argument around, these results suggest the potential of this observable to discriminate between different models of the medium backreaction. However, the large-angle domain is also contaminated by the fluctuating underlying event in heavy-ion collisions, as was shown in Ref. [32]. In this work, we are interested in designing a pQCD-dominated observable and therefore we proceed to present two possible ways of reducing the impact of soft physics, i.e., both the medium backreaction and the thermal background: (i) using smaller jet radii and (ii) selecting semiperipheral events.[10]

1. Jet radius dependence

Several experimental measurements of jet substructure [7,11,91,92] have used small-R jets in order to reduce the impact of combinatorial jets. From the theoretical point of view, describing the cone-size dependence of jet quenching is an active field of research [58,93]. The impact of reducing the jet radius from $R = 0.4$ to $R = 0.2$ on the θ_g distribution is shown in Fig. 14. Clearly, tagging soft, thermal particles in narrower jets is less probable than in Fig. 13. In the case of the Hybrid model this choice is extremely efficient in minimizing the influence of medium response. The last statement is true for all values of the dynamical grooming parameter a. On the other hand, reducing the jet radius is not enough to make the sensitivity to recoil particles in Jewel vanish. An extra cut on the z of the emissions *á la* SoftDrop would probably be helpful. Of course, by narrowing the phase space for emissions, quenching effects are also diminished and that is the reason why the vacuum and medium distributions look more alike than in the $R = 0.4$ case. Therefore, we conclude that, as expected, mitigating the impact of medium response by shrinking the jet radius comes at the price of a reduction in the discriminating power of the observable.

2. Centrality dependence

Another possibility to reduce the impact of medium response is to explore semiperipheral collisions where the medium is not as dense as when the two nuclei collide head on. Figure 15 demonstrates that moving to semiperipheral collisions does not reduce the medium response component as effectively as reducing the jet radius did. We therefore conclude that the combination that maximizes the sensitivity to color coherence effects is $R = 0.2$, 0–5% and $0.5 \leq a \leq 1$.

Studying the dependence of the θ_g distribution with centrality is interesting not only from the point of view of reducing the impact of medium response, but also to further constrain the resolution angle [58]. In a weak coupling picture the scaling of θ_c with respect to the length of the medium is well known to be $\theta_c \propto L^{-3/2}$. In contrast, if no coherent angle existed and the maximum of the θ_g distribution was driven by the filtering effect caused by the finite size of the medium, one would expect a $\theta_{\rm max} \propto L^{-1}$ scaling. Therefore, the centrality dependence of the θ_g distribution s peak location would be stronger for a θ_c dependent energy loss. In order to explicitly demonstrate this statement one would have to implement the geometry of the collision in JetMed in such a way that different centralities can be simulated. This task is left for future work.

[10]Another alternative that would effectively reduce the impact of medium response, but that we do not pursue in this work, would be to increase the jet p_t in the selection.

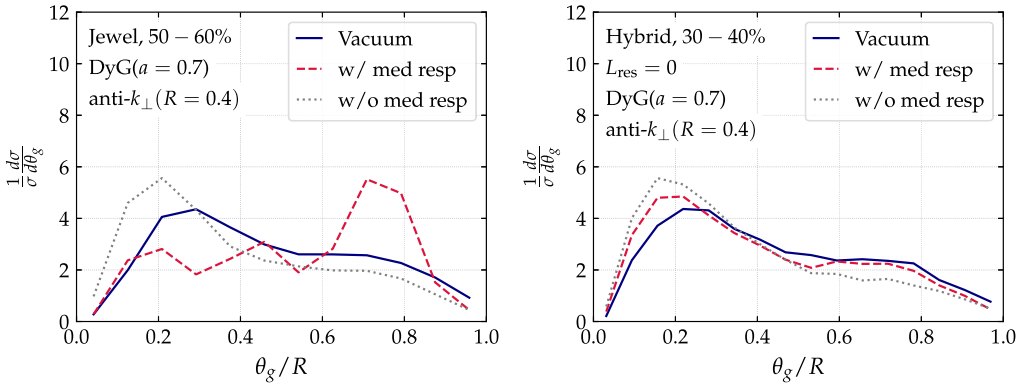


FIG. 15. Same as Fig. 13 but for 50–60% in Jewel and for 30–40% in Hybrid. The medium response contribution is reduced but less effectively than in Fig. 14.

IV. FINAL REMARKS AND OUTLOOK

In this paper, we presented a comprehensive analysis of a promising jet substructure observable in heavy-ion collisions: the dynamically groomed jet radius θ_g distribution. We have shown through analytic calculations grounded in pQCD that the medium-modified distribution is strongly sensitive to the coherence angle of the medium θ_c. These analytic calculations are supplemented by Monte Carlo calculations of this observable with JetMed which are in qualitative agreement with our analytic results. In summary, our study demonstrates the ability of the dynamically groomed jet radius to measure decoherence effects in the substructure of a jet traveling through a dense QCD medium. The numerical routines used in this work are provided in Ref. [51].

Given the simplifications inherent to our analytic and JetMed calculations, it is not our intention to provide quantitative predictions for the θ_g distribution to be measured in heavy-ion collisions. However, we have been able to pin down the dominant physical mechanisms at stake which drive the modification of this observable. In fact, there are several dynamical processes in the medium that converge at the QGP resolution angle and leave their imprint on the dynamically groomed jet radius. For instance, (i) the presence of the veto region leads to a narrowing of the θ_g distribution that is more pronounced, the larger the value of $a \leq 2$ is. On the contrary, (ii) the medium-induced branching kernel generates an enhancement of large-angle splittings due to transverse momentum broadening. The last two competing physical ingredients are eclipsed by (iii) differential energy loss. That is, when constructing a toy model for an in-medium parton shower that includes vacuum-like and medium-induced emissions as well as energy loss, we observe that narrow splittings are enhanced for all values of a with respect to the vacuum baseline. Notably, the inclusion of a fluctuating jet path length smooths out the transition at θ_c, but does not wash out the signal completely.

We have also explored the sensitivity of our results to the underlying theoretical modeling of the jet-medium interactions, using the Hybrid model that relies on strong coupling jet-medium interactions. This model also predicts a strong sensitivity of the observable to the medium resolution length, which is the strong coupling analog of the coherence angle. Further, we have studied the impact of medium response in this observable with both the Hybrid and Jewel models. The imprint of these soft particles on the θ_g distribution is an enhancement of wide angle splittings. However, the magnitude of the $\theta_g \approx R_{\text{jet}}$ peak significantly differs in the two descriptions of the medium. Since we are interested in a pQCD-dominated observable we explored two routes to reduce the medium response contribution: reducing the jet cone size and using semiperipheral events. Our findings indicate that the former option is more efficient than the latter.

We would like to emphasize that the main difference between the in-medium and vacuum θ_g distributions is not just a displacement in the peak position, but rather a significant modification of the shape of the distribution as a whole. That is the main reason why we quantify the in-medium to vacuum differences with the Kolmogorov-Smirnov distance and not with a simple ratio as is typically done experimentally for other jet substructure

observables such as the SoftDrop family in Ref. [11]. Thanks to the use of this metric, we were able to provide, both for the analytic calculation and for the Monte Carlo simulations, a reasonable window for the dynamical grooming parameter a and jet radius R to be used experimentally in order to maximize the effects of the coherence angle, while at the same time minimizing the nonperturbative contributions, such as hadronization, geometry fluctuations and medium response. This is a tremendous advantage of the dynamical grooming procedure which depends on a single free parameter a. As such, studying the physics probed by the observable as a function of this single parameter is straightforward. We found that the optimal values are $0.5 \lesssim a \lesssim 1$ and $R \lesssim 0.2$, regardless of the theoretical model of jet quenching.

We have shown that both weak and strong coupling models lead to similar trends in the θ_g distribution. In order to move forward and disentangle theoretical models of jet quenching, there is a crucial need for performing global analyses in which models are tested against both global jet energy loss (R_{AA} like) and jet substructure measurements, such as θ_g after dynamical grooming considered in this paper. A scan in terms of centrality classes or colliding system sizes is an interesting possibility to be explored in the future, given the theoretically well-defined path-length dependence of the critical angle. It will likely constrain more precisely the shape of the medium-modified phase space in Fig. 2 and reveal unambiguously the existence of a critical line at $\theta = \theta_c$ as well as its dependence on the physical properties of the medium. In addition, experimental data on the k_t of the hardest emission would provide complementary information to the θ_g measurement since it probes the orthogonal direction in phase space.

A natural extension of this work would be to go beyond the multiple soft scattering approximation of the parton-medium interaction. We plan to study the impact of rare, hard scatterings on the phase space of emissions within the improved opacity expansion in a forthcoming publication [50]. Further, our resummation could be extended to account for heavy quarks in order to quantify the potential of the θ_g distribution to expose the dead-cone effect in heavy-ion collisions [94–98].

ACKNOWLEDGMENTS

We are grateful to Daniel Pablos and Raghav Kunnawalkam Elayavalli for providing the Hybrid and Jewel samples, respectively. Enlightening conversations with Yacine Mehtar-Tani and Konrad Tywoniuk are acknowledged. We thank Liliana Apolinario, Leticia Cunqueiro, Laura Havener, Edmond Iancu, Yacine Mehtar-Tani, James Mulligan and Daniel Pablos for a careful reading of the manuscript and useful comments. P. C.'s work was supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics, under Contract No. DE-SC0012704. A. S. O.'s work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 788223, PanScales). A. T.'s work was supported by the Starting Grant from Trond Mohn Foundation (BFS2018REK01).

APPENDIX A: SUDAKOV WITH VETO REGION

The purpose of this appendix is to provide a semi-analytic formula for the Sudakov form factor that includes the veto region for vacuum-like emissions given by Eq. (28). Using Eq. (14), one can express $\Delta_{\notin\text{veto}}$ in terms of $\Delta_{\in\text{veto}}$ using

$$\ln(\Delta_{\notin\text{veto}}) = \ln(\Delta) - \ln(\Delta_{\in\text{veto}}). \tag{A1}$$

After replacing $\tilde{P}(z,\theta)$ by Eq. (6), we get for the in-veto contribution

$$\ln \Delta_{\in\text{veto}}(\kappa|a) = -2\bar{\alpha} \int_{\theta_{\min}}^{\theta_{\max}} \frac{\mathrm{d}\theta'}{\theta'} \ln \frac{z_{\max}(\kappa,\theta')}{z_{\min}(\kappa,\theta')} \tag{A2}$$

with

$$z_{\min}(\kappa,\theta') = \max\left[\kappa\left(\frac{R}{\theta'}\right)^a, \frac{2}{\theta'^2 p_t L}\right], \tag{A3}$$

$$z_{\max}(\kappa,\theta') = \min\left[1, \left(\frac{2\hat{q}}{p_t^3\theta'^4}\right)^{\frac{1}{3}}\right], \tag{A4}$$

$$\theta_{\min}(\kappa) = \max\left[R\kappa^{\frac{1}{a}}, \theta_c, \text{ if } \left(a > \frac{4}{3}\right) : \left(\frac{2\hat{q}}{(\kappa R^a p_t)^3}\right)^{\frac{1}{4-3a}}\right], \tag{A5}$$

$$\theta_{\max}(\kappa) = \min\left[R, \text{ if } \left(a \le \frac{4}{3}\right) : \left(\frac{2\hat{q}}{(\kappa R^a p_t)^3}\right)^{\frac{1}{4-3a}}\right]. \tag{A6}$$

APPENDIX B: REALISTIC JET SPECTRUM

To make our analytic predictions closer to reality we have to include the jet spectrum. It enters into our calculation through the quark/gluon ratio and through the spectrum power n in the energy loss component of Eq. (27). We use the dijet parametrization from Ref. [59] at $\sqrt{s} = 5.02$ TeV, $|\eta| < 2.8$ and $R = 0.4$:

TABLE I. Parametrization of the jet spectrum for both proton PDFs and using the nuclear PDF for Pb.

	i	a [mb]	p_{t0} [GeV]	b	c
PDF	q	0.008	20.8	4.7	0.08
	g	0.007	26.4	5.1	0.17
Nuclear PDFs	q	0.001	32.8	4.7	0.1
	g	0.0005	44.4	5.2	0.15

$$\frac{\mathrm{d}\sigma_i}{\mathrm{d}p_t} = a\left(\frac{p_{t0}}{p_t}\right)^{-\left(b+c\ln\frac{p_t}{p_{t0}}\right)}, \tag{B1}$$

where the subscript i indicates the flavor of the initiating parton and (a, b, p_{t0}, c) are free parameters. This leads to the coefficients presented in Table I.

The impact of the nuclear parton distribution functions (PDFs) on the observable at play is shown in Fig. 16. We observe that the quark/gluon fraction is barely modified in this p_t window.

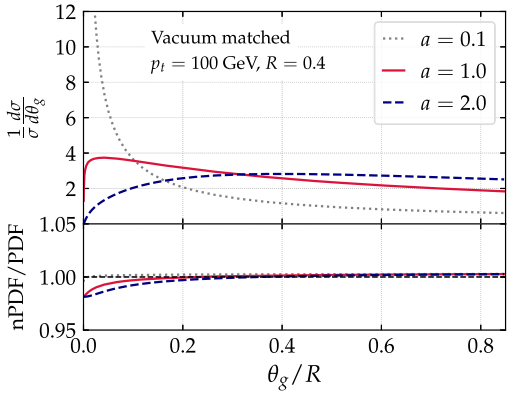


FIG. 16. θ_g distribution for vacuum splittings with realistic quark-gluon fractions using nuclear PDFs. The ratio to the vacuum PDF is presented in the bottom panel showing a very mild impact of the nPDFs.

[1] H. A. Andrews *et al.*, J. Phys. G **47**, 065102 (2020).
[2] L. Cunqueiro and A. M. Sickles, Prog. Part. Nucl. Phys. **124**, 103940 (2022).
[3] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Rev. C **90**, 024908 (2014).
[4] M. Aaboud *et al.* (ATLAS Collaboration), Eur. Phys. J. C **77**, 379 (2017).
[5] P. Caucal, E. Iancu, A. H. Mueller, and G. Soyez, J. High Energy Phys. 10 (2020) 204.
[6] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Lett. B **730**, 243 (2014).
[7] S. Acharya *et al.* (ALICE Collaboration), J. High Energy Phys. 10 (2018) 139.
[8] R. Kunnawalkam Elayavalli and K. C. Zapp, J. High Energy Phys. 07 (2017) 141.
[9] Y.-T. Chien and I. Vitev, J. High Energy Phys. 05 (2016) 023.
[10] M. S. Abdallah *et al.* (STAR Collaboration), Phys. Rev. C **105**, 044906 (2022).
[11] S. Acharya *et al.* (ALICE Collaboration), Phys. Rev. Lett. **128**, 102001 (2022).
[12] J. Adam *et al.* (STAR Collaboration), Phys. Lett. B **811**, 135846 (2020).
[13] S. Acharya *et al.* (ALICE Collaboration), J. High Energy Phys. 05 (2022) 061.
[14] A. M. Sirunyan *et al.* (CMS Collaboration), Phys. Rev. D **98**, 092014 (2018).
[15] G. Aad *et al.* (ATLAS Collaboration), Phys. Rev. D **101**, 052007 (2020).
[16] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, J. High Energy Phys. 09 (2013) 029.
[17] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, J. High Energy Phys. 05 (2014) 146.
[18] V. N. Gribov and L. N. Lipatov, Deep inelastic ep-scattering in a perturbation theory., Technical Report No. IPTI-381-71, Inst. of Nuclear Physics, Leningrad, 1972.
[19] G. Altarelli and G. Parisi, Nucl. Phys. **B126**, 298 (1977).
[20] Y. L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977).
[21] A. J. Larkoski, S. Marzani, and J. Thaler, Phys. Rev. D **91**, 111501 (2015).
[22] P. Cal, K. Lee, F. Ringer, and W. J. Waalewijn, arXiv: 2106.04589.
[23] R. Baier, Y. L. Dokshitzer, A. H. Mueller, S. Peigne, and D. Schiff, Nucl. Phys. **B484**, 265 (1997).
[24] B. G. Zakharov, JETP Lett. **65**, 615 (1997).
[25] J. Casalderrey-Solana, Y. Mehtar-Tani, C. A. Salgado, and K. Tywoniuk, Phys. Lett. B **725**, 357 (2013).
[26] P. Caucal, E. Iancu, and G. Soyez, J. High Energy Phys. 10 (2019) 273.
[27] A. M. Sirunyan *et al.* (CMS Collaboration), Phys. Rev. Lett. **120**, 142302 (2018).
[28] Y. Mehtar-Tani and K. Tywoniuk, J. High Energy Phys. 04 (2017) 125.
[29] Y.-T. Chien and I. Vitev, Phys. Rev. Lett. **119**, 112301 (2017).
[30] G. Milhano, U. A. Wiedemann, and K. C. Zapp, Phys. Lett. B **779**, 409 (2018).
[31] N.-B. Chang, S. Cao, and G.-Y. Qin, Phys. Lett. B **781**, 423 (2018).
[32] J. Mulligan and M. Ploskon, Phys. Rev. C **102**, 044913 (2020).

[33] Y. Mehtar-Tani, C. A. Salgado, and K. Tywoniuk, Phys. Rev. Lett. **106**, 122002 (2011).
[34] J. Casalderrey-Solana and E. Iancu, J. High Energy Phys. 08 (2011) 015.
[35] Y. Mehtar-Tani and K. Tywoniuk, Phys. Lett. B **744**, 284 (2015).
[36] K. Rajagopal, A. V. Sadofyev, and W. van der Schee, Phys. Rev. Lett. **116**, 211603 (2016).
[37] F. Ringer, B.-W. Xiao, and F. Yuan, Phys. Lett. B **808**, 135634 (2020).
[38] P. Caucal, E. Iancu, A. H. Mueller, and G. Soyez, Phys. Rev. Lett. **120**, 232001 (2018).
[39] J. Casalderrey-Solana, G. Milhano, D. Pablos, and K. Rajagopal, J. High Energy Phys. 01 (2020) 044.
[40] J. H. Putschke *et al.*, arXiv:1903.07706.
[41] A. Majumder, Phys. Rev. C **88**, 014909 (2013).
[42] Y. He, T. Luo, X.-N. Wang, and Y. Zhu, Phys. Rev. C **91**, 054908 (2015); **97**, 019902(E) (2018).
[43] L. Apolinário, A. Cordeiro, and K. Zapp, Eur. Phys. J. C **81**, 561 (2021).
[44] Y. Mehtar-Tani, A. Soto-Ontoso, and K. Tywoniuk, Phys. Rev. D **101**, 034004 (2020).
[45] F. D'Eramo, K. Rajagopal, and Y. Yin, J. High Energy Phys. 01 (2019) 172.
[46] P. Caucal and Y. Mehtar-Tani, arXiv:2109.12041.
[47] P. Caucal, A. Soto-Ontoso, and A. Takacs, J. High Energy Phys. 07 (2021) 020.
[48] J. Mulligan (ALICE Collaboration), Proc. Sci., HardProbes2020 (**2021**) 141 [arXiv:2009.07172].
[49] R. Ehlers (ALICE Collaboration), Proc. Sci., HardProbes2020 (**2021**) 146 [arXiv:2009.12247].
[50] J. Barata, P. Caucal, A. Soto-Ontoso, A. Takacs, and K. Tywoniuk (to be published).
[51] https://github.com/albaontoso/DyG-HI-analytics.
[52] M. Gyulassy, P. Levai, and I. Vitev, Phys. Rev. Lett. **85**, 5535 (2000).
[53] S. Caron-Huot and C. Gale, Phys. Rev. C **82**, 064902 (2010).
[54] X. Feal and R. Vazquez, Phys. Rev. D **98**, 074029 (2018).
[55] C. Andres, L. Apolinário, and F. Dominguez, J. High Energy Phys. 07 (2020) 114.
[56] C. Andres, F. Dominguez, and M. Gonzalez Martinez, J. High Energy Phys. 03 (2021) 102.
[57] J. a. Barata, Y. Mehtar-Tani, A. Soto-Ontoso, and K. Tywoniuk, J. High Energy Phys. 09 (2021) 153.
[58] Y. Mehtar-Tani, D. Pablos, and K. Tywoniuk, Phys. Rev. Lett. **127**, 252301 (2021).
[59] A. Takacs and K. Tywoniuk, J. High Energy Phys. 10 (2021) 038.
[60] J. a. Barata, F. Domínguez, C. A. Salgado, and V. Vila, J. High Energy Phys. 05 (2021) 148.
[61] S. P. Adhya, C. A. Salgado, M. Spousta, and K. Tywoniuk, Eur. Phys. J. C **82**, 20 (2022).
[62] E. Blanco, K. Kutak, W. Placzek, M. Rohrmoser, and K. Tywoniuk, Eur. Phys. J. C **82**, 355 (2022).
[63] M. Rohrmoser, Comput. Phys. Commun. **276**, 108343 (2022).
[64] C. Andres, N. Armesto, H. Niemi, R. Paatelainen, and C. A. Salgado, Phys. Lett. B **803**, 135318 (2020).
[65] F. A. Dreyer, G. P. Salam, and G. Soyez, J. High Energy Phys. 12 (2018) 064.
[66] S. Acharya *et al.* (ALICE Collaboration), Phys. Rev. C **101**, 034911 (2020).
[67] J.-P. Blaizot, E. Iancu, and Y. Mehtar-Tani, Phys. Rev. Lett. **111**, 052001 (2013).
[68] J.-P. Blaizot, F. Dominguez, E. Iancu, and Y. Mehtar-Tani, J. High Energy Phys. 06 (2014) 075.
[69] J.-P. Blaizot and Y. Mehtar-Tani, Ann. Phys. (Amsterdam) **368**, 148 (2016).
[70] J.-P. Blaizot, L. Fister, and Y. Mehtar-Tani, Nucl. Phys. **A940**, 67 (2015).
[71] J.-P. Blaizot, Y. Mehtar-Tani, and M. A. C. Torres, Phys. Rev. Lett. **114**, 222002 (2015).
[72] R. Baier, Y. L. Dokshitzer, A. H. Mueller, and D. Schiff, J. High Energy Phys. 09 (2001) 033.
[73] C. A. Salgado and U. A. Wiedemann, Phys. Rev. D **68**, 014008 (2003).
[74] L. Fister and E. Iancu, J. High Energy Phys. 03 (2015) 082.
[75] Y. Mehtar-Tani and K. Tywoniuk, Phys. Rev. D **98**, 051501 (2018).
[76] J. Brewer, J. G. Milhano, and J. Thaler, Phys. Rev. Lett. **122**, 222301 (2019).
[77] Y.-L. Du, D. Pablos, and K. Tywoniuk, J. High Energy Phys. 21 (2020) 206.
[78] Y.-L. Du, D. Pablos, and K. Tywoniuk, Phys. Rev. Lett. **128**, 012301 (2022).
[79] J. Brewer, Q. Brodsky, and K. Rajagopal, J. High Energy Phys. 02 (2022) 175.
[80] S. P. Adhya, C. A. Salgado, M. Spousta, and K. Tywoniuk, J. High Energy Phys. 07 (2020) 150.
[81] P. Caucal, E. Iancu, and G. Soyez, J. High Energy Phys. 04 (2021) 209.
[82] C. A. Salgado and U. A. Wiedemann, Phys. Rev. Lett. **89**, 092303 (2002).
[83] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, Comput. Phys. Commun. **191**, 159 (2015).
[84] P. Skands, S. Carrazza, and J. Rojo, Eur. Phys. J. C **74**, 3024 (2014).
[85] J. Casalderrey-Solana, D. C. Gulhan, J. G. Milhano, D. Pablos, and K. Rajagopal, J. High Energy Phys. 10 (2014) 019; 09 (2015) 175(E).
[86] K. C. Zapp, F. Krauss, and U. A. Wiedemann, J. High Energy Phys. 03 (2013) 080.
[87] K. C. Zapp, Eur. Phys. J. C **74**, 2762 (2014).
[88] M. Cacciari, G. P. Salam, and G. Soyez, J. High Energy Phys. 04 (2008) 063.
[89] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, J. High Energy Phys. 08 (1997) 001.
[90] Z. Hulcher, D. Pablos, and K. Rajagopal, J. High Energy Phys. 03 (2018) 010.
[91] S. Acharya *et al.* (ALICE Collaboration), Phys. Lett. B **802**, 135227 (2020).
[92] ATLAS Collaboration, Measurement of suppression of large-radius jets and its dependence on substructure in Pb + Pb at 5.02 TeV by ATLAS detector, Technical Report No. ATLAS-CONF-2019-056, CERN, Geneva, 2019.
[93] D. Pablos, Phys. Rev. Lett. **124**, 052301 (2020).
[94] L. Cunqueiro, D. Napoletano, and A. Soto-Ontoso (to be published).

[95] L. Cunqueiro and M. Płoskoń, Phys. Rev. D **99**, 074027 (2019).
[96] S. Acharya *et al.* (ALICE Collaboration), Nature (London) **605**, 440 (2022).
[97] Y. L. Dokshitzer, V. A. Khoze, and S. I. Troian, J. Phys. G **17**, 1602 (1991).
[98] N. Armesto, C. A. Salgado, and U. A. Wiedemann, Phys. Rev. D **69**, 114003 (2004).

Article V.

5.5 A unified picture of medium-induced radiation

Johannes Hamre Isaksen, Adam Takacs and Konrad Tywoniuk

JHEP, **02**, 056 (2023)

Paper about a systematic calculation of medium-induced emissions in a finite-size medium, and finite energy. Furthermore, solving medium-induced cascade with this novel kernel. Keywords: medium-induced emissions, Bethe-Heitler emissions, harmonic oscillator approximation, BDMPS-Z spectrum, GLV spectrum, fragmentation function.

Published for SISSA by Springer

Received: *June 20, 2022*
Revised: *January 13, 2023*
Accepted: *February 1, 2023*
Published: *February 15, 2023*

A unified picture of medium-induced radiation

Johannes Hamre Isaksen, Adam Takacs and Konrad Tywoniuk

Department of Physics and Technology, University of Bergen, Allegaten 55, 5007 Bergen, Norway

E-mail: johannes.isaksen@uib.no, adam.takacs@uib.no, konrad.tywoniuk@uib.no

Abstract: We revisit the picture of jets propagating in the quark-gluon plasma. In addition to vacuum radiation, partons scatter on the medium constituents resulting in induced emissions. Analytical approaches to including these interactions have traditionally dealt separately with multiple, soft, or rare, hard scatterings. A full description has so far only been available using numerical methods. We achieve full analytical control of the relevant scales and map out the dominant physical processes in the full phase space. To this aim, we extend existing expansion schemes for the medium-induced spectrum to the Bethe-Heitler regime. This covers the whole phase space from early to late times, and from hard splittings to emissions below the thermal scale. Based on the separation of scales, a space-time picture naturally emerges: at early times, induced emissions start to build from rare scatterings with the medium. At a later stage, induced emissions due to multiple soft scatterings result in a turbulent cascade that rapidly degrades energy down to, and including, the Bethe-Heitler regime. We quantify the impact of such an improved picture, compared to the current state-of-the-art factorization that includes only soft scatterings, by both analytical and numerical methods for the medium-induced energy distribution function. Our work serves to improve our understanding of jet quenching from small to large systems and for future upgrades of Monte Carlo generators.

Keywords: Jets and Jet Substructure, Quark-Gluon Plasma

ArXiv ePrint: 2206.02811

https://doi.org/10.1007/JHEP02(2023)156

JHEP02(2023)156

Contents

1 Introduction **1**

2 Heuristic discussion of the medium-induced spectrum **4**

3 Spectrum of medium-induced emissions **9**
- 3.1 General formalism 9
- 3.2 Opacity expansion (OE) 12
- 3.3 Resummed opacity expansion (ROE) 14
- 3.4 Improved opacity expansion (IOE) 18
- 3.5 Summary of the regimes and the induced emission spectrum 23

4 Resumming multiple emissions in the medium **25**
- 4.1 The necessity of multiple emissions 25
- 4.2 Resummation of multiple emissions 26
- 4.3 Analytic solutions of the evolution equation 29

5 Numerical evaluation of the medium cascade **33**

6 Conclusions and outlook **37**

A All order formulas for medium-induced spectrum in the soft limit **38**
- A.1 Opacity expansion 38
- A.2 Resummed opacity expansion 39
- A.3 Improved opacity expansion 41

B General formulas for the emission rate in the soft limit **43**

C Medium-induced spectrum and rate with finite-z corrections **43**
- C.1 Opacity expansion 45
- C.2 Resummed opacity expansion 46
- C.3 Improved opacity expansion 47

D Numerical implementation of the evolution equation **50**

E HTL potential **51**

1 Introduction

Short-lived droplets of hot and dense nuclear matter, called the quark-gluon plasma (QGP), are produced in relativistic heavy-ion collisions at RHIC and LHC. Embedded in the same high-energy collisions, hard QCD processes are also present, resulting in the production of collimated sprays of energetic particles that are commonly referred to as jets [1, 2]. Jets are well-understood, perturbative objects within perturbative QCD and they are described

JHEP02(2023)156

up to high precision in proton-proton collisions [3–5]. During their propagation, however, jet particles can interact with the surrounding nuclear matter. The modification of jet features, therefore, reflects the properties of the QGP created in heavy-ion collisions [6–11]. Currently, a vigorous experimental program dedicated to quantifying jet modifications is ongoing at both RHIC and LHC, focusing on a broad set of observables which includes measurements of the modification of the jet spectrum, jet substructure observables, and jet correlations [12–16] (for a selection of predictions see refs. [17–27]).

High-energy jets are particularly suitable probes of the QGP because their energy scale $Q_{\rm jet}$ is much larger than the typical momentum scale of the medium $Q_{\rm med}$. If this is the case, the impact of medium modifications should therefore not affect the internal structure of the jet, which would still rely on perturbative QCD [28–30]. A key ingredient when considering jet modifications is the radiation induced by scatterings with the deconfined medium constituents. Such emissions typically appear at scales comparable to $Q_{\rm med}$. Emission at scales much higher than $Q_{\rm med}$, on the other hand, are unaffected by the medium, resulting in a factorized picture between vacuum and medium processes [29, 31]. Medium-induced emissions redistribute the original jet parton energy to multiple, soft particles over large angles, including out of the jet cone. This leads to a net jet energy loss which, in turn, is manifested as a suppression of the jet spectrum (for an updated discussion of jet quenching see refs. [25, 26], and for applications to substructure see refs. [18, 20, 24]). Consequently, medium-induced emissions are a crucial component of jet energy loss and thus of phenomenological studies of jet observables in heavy-ion collisions.

The medium-induced emission spectrum was formulated a long time ago [32–34]. Previous solutions were limited to either (i) expanding in the number of scatterings (referred to as the opacity expansion) [35–38], or (ii) considering multiple soft scatterings (called the harmonic oscillator approximation) [32, 33, 39, 40]. Meanwhile, several works focused on the underlying scales that separate the limiting cases [32, 41–46]. The full problem has also been tackled by numerical techniques [47, 48] (or more recently in refs. [49–52]). Not long ago, analytical techniques were developed that provided a unified description of the multiple, soft and rare, hard scatterings in a dense medium [53–57], which better match the full numerical solutions. The main challenge, common to both the numeric and analytic approaches, resides in dealing with multiple interactions with the underlying medium.

In this paper, we revisit the different analytic approaches to resumming multiple interactions for calculating the medium-induced emission spectrum. These include the opacity expansion (OE) and the improved opacity expansion (IOE), which includes harmonic oscillator approximation as the leading term. Moreover, we rigorously derive the resummed opacity expansion (ROE) for the first time, which extends the description of the spectrum to low energy emissions in the Bethe-Heitler regime. We provide a novel unified picture of these resummation schemes by identifying their relevant emergent scales and demonstrating their respective regions of validity. For example, we show that the single scattering approximation, contained in the leading order of OE, is valid even for a big medium, where one would expect more than one scattering if the emitted energy is high enough. We show that the full phase space of medium-induced emissions, spanning from the maximal jet energy to the thermal scale, is covered by a union of these expansions, see also ref. [46]. Each of the

JHEP02(2023)156

expansions is also associated with the corresponding physical scattering processes, and thus we reinterpret the frequently used terms such as GLV emissions, coherent scatterings, and Bethe-Heitler region in a unified framework. Our framework goes beyond previous attempts to describe all regimes of medium-induced bremsstrahlung by presenting a resummation framework that can be systematically improved and that is valid both in the dilute and dense regimes.

As a next step, we identify the regimes where not only multiple interactions are important, but also multiple emissions [58, 59]. These conditions are met for sufficiently soft emissions in a large medium. The previously established hierarchy of emergent scales plays a crucial role in mapping out early, rare, and relatively hard emissions and a successive cascade of soft splittings. In this context, hard medium-induced splittings can be thought of as extra sources, in addition to the parent parton, for the full cascade. This description is realized analytically in a novel scheme that combines a fixed order expansion of rare emissions with an all-order resummation of soft splittings. The resulting energy distribution links the asymptotic early and late time behaviors for which analytical solutions exist. Finally, we resum multiple induced emissions numerically, using the previously obtained precise determination of the in-medium splitting rates, to calculate the energy distribution function. We highlight the interplay of rare hard scatterings, coherent soft splittings, and Bethe-Heitler emissions in a finite medium, providing a state-of-the-art resummation.

Our reorganized picture helps not only with the physical understanding of induced emissions, but provides a fast and efficient way to calculate the medium-induced spectrum, which is a key ingredient for estimating jet energy loss. It also serves to inform Monte Carlo algorithms simulating full jet evolution inside the medium about how to implement multiple medium-induced emissions and how to combine them with vacuum-like emissions, e.g. see in ref. [29].

The paper is structured as follows. In section 2, as an introduction, we discuss the structure of the induced spectrum in the various regimes using heuristic arguments, and we show how the radiation in the different regimes is related to single soft, multiple soft and single hard scatterings with the medium, see figure 3. The spectrum is calculated in detail in section 3. We revisit the opacity expansion and the improved opacity expansion schemes, and put on a firm footing a novel resummation scheme, dubbed resummed opacity expansion, which is valid for emissions below the Bethe-Heitler scale. Improving on previous discussions, we provide formulas for the spectrum at arbitrary order and calculate it exactly up to second order in all the expansions, allowing us for the first time to establish regions where they converge. Finally, in section 4 we consider the problem of multiple emissions. We analyze induced particles coming from the full phase space and confirm the importance of considering multiple emissions, especially in the soft sector. In order to facilitate an improved analytical understanding of the problem, we finally suggest a resummation scheme of multiple emissions by iterating in rare, hard emissions and including an arbitrary number of soft splittings. This is compared to the full numerical results. We conclude with an outlook in section 6. The appendix contains lots of useful formulas, including the rate of emissions and finite-z corrections that are important for phenomenology. The code we have developed to calculate the kernels and solve for the energy distribution is provided at https://github.com/adam-takacs/kernels.git.

JHEP02(2023)156

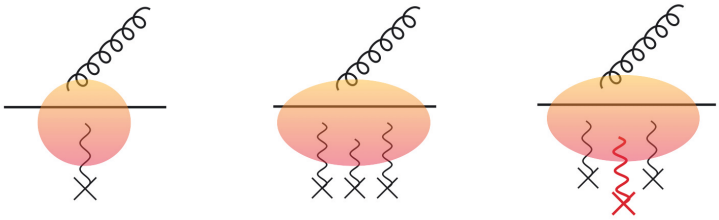


Figure 1. Three regimes of the radiative spectrum in dense media, $L \gg \lambda$: the Bethe-Heitler regime $\omega < \omega_{\rm BH}$ (leftmost), the BDMPS-Z regime $\omega_{\rm BH} < \omega < \omega_c$ (middle), and the hard GLV regime $\omega_c < \omega$ (rightmost figure). The size of the blob represents the typical formation time of the emission, which is $t_{\rm f} < \lambda$ in the leftmost panel and $t_{\rm f} > \lambda$ in the center and rightmost panels. Black, wavy lines represent soft interactions with the medium, where the momentum transfer is of the order of the medium scale $|\boldsymbol{q}| \lesssim \mu$, while the red, wavy line represents a hard scattering event, with $|\boldsymbol{q}| \gg \mu$.

2 Heuristic discussion of the medium-induced spectrum

The spectrum of gluon emissions induced by scatterings on a deconfined medium plays a central role in the phenomenology of jet quenching. However, a full understanding of all regimes have so far been lacking analytically and was previously only achievable through numerical methods. Here, we present a unified view of all relevant medium scales and their related regimes. Similar heuristic discussions have previously been presented in, e.g., [8, 32, 46, 53]. In section 3 we will provide rigorous derivations of the findings argued for here.

We formulate the spectrum of induced emissions by focusing on the relevant length scales:

- The **size** of the medium L (or the length of propagation $t < L$).
- The **mean free path** of the medium $\lambda \sim \frac{1}{n\sigma_{\rm tot}}$, which combines the density n and the scattering strength $\sigma_{\rm tot} \sim \int \mathrm{d}\sigma$, and it describes the distance between typical scatterings.
- The **formation time** of an emission $t_{\rm f} = \frac{2\omega}{\boldsymbol{k}^2}$, where ω is the energy and $\boldsymbol{k}$ is the transverse momentum of the emission.

In addition, the relation between the in-medium screening scale μ and the range of available transverse momenta $|\boldsymbol{k}|$ is also important.

The *opacity* $\chi \equiv L/\lambda$ characterizes the denseness of the medium. If the opacity is small ($L \ll \lambda$), the medium is "dilute", or weakly interacting, while it is "dense", or strongly interacting if $L \gg \lambda$.[1] The dilute medium barely consists of scattering centers, however, in the dense medium one should account for an arbitrary number of interactions.

[1]From this perspective, we have "fixed" L and vary λ. Naturally, we could also have identified these two regimes as a "large" and "small" media, where we "fix" the mean free path and vary L instead.

JHEP02(2023)156

A big part of this paper will be about calculating the spectrum of medium-induced emissions. For reference, the vacuum spectrum reads,

$$\omega \frac{\mathrm{d}I}{\mathrm{d}\omega} \sim \alpha_s \int \frac{\mathrm{d}\boldsymbol{k}^2}{\boldsymbol{k}^2}\,, \tag{2.1}$$

where currently we did not specify the limits of the transverse momentum integral. This contains the well-known soft ($\omega \to 0$) and collinear ($\boldsymbol{k}^2 \to 0$) divergences. In contrast, the collinear divergence in the medium spectrum is removed by the need to exchange transverse momentum with the medium.

Motivated by this, we introduce a heuristic model that captures some of the features of the medium-induced spectrum, given by

$$\omega \frac{\mathrm{d}I}{\mathrm{d}\omega} \sim \alpha_s L \int \mathrm{d}^2\boldsymbol{k}\, \sigma(\boldsymbol{k}) \sim \alpha_s^3 n L \int \frac{\mathrm{d}\boldsymbol{k}^2}{(\boldsymbol{k}^2+\mu^2)^2}\,, \tag{2.2}$$

where n is the medium density and μ a screening mass. The behavior at high-$\boldsymbol{k}$, $\sigma \sim \boldsymbol{k}^{-4}$, reproduces the expected Coulomb tail. The factor L arises since the emission can take place anywhere along the medium length. For a more accurate description of medium-induced emissions, we refer the reader to section 3.

In our effective description, we focus on the hierarchy among the introduced scales and show the separation of different scattering regions. Firstly, in the $t_\mathrm{f} \gg L$ limit, the formation of the emission extends beyond the medium, where one naturally should expect vacuum physics to dominate.[2] We will hence not consider this possibility here. The remaining cases are listed below:

Dilute media ($t_\mathrm{f} \leq L \ll \lambda$). In case of a low medium opacity, we expect that roughly one scattering occurs. This process will typically transfer a momentum of order of the Debye mass to the emitted gluon, or $\langle \boldsymbol{k}^2 \rangle \sim \mu^2$, leading to $t_\mathrm{f} = 2\omega/\mu^2$. The formation of the gluon has to take place inside the medium, giving rise to the characteristic energy scale in the dilute regime, namely

$$\bar{\omega}_c = \frac{1}{2}\mu^2 L\,. \tag{2.3}$$

This separates two regimes of emissions that are sourced via different scattering processes: on the one side soft gluons with $\omega < \bar{\omega}_c$, generated via a soft scattering with the medium $\langle \boldsymbol{k}^2 \rangle \lesssim \mu^2$. Hard gluons with $\omega > \bar{\omega}_c$ can also be generated, however those demand a large momentum exchange with the medium, $\langle \boldsymbol{k}^2 \rangle > \mu^2$, which is comparatively rare. Let us now consider how the spectrum behaves in these two distinct regimes.

According to our discussion above, the soft production should be dominated by soft transverse momentum exchanges with the medium. Hence, we expect that the spectrum of emitted gluons goes as

$$\omega \frac{\mathrm{d}I}{\mathrm{d}\omega}\bigg|_{\omega<\bar{\omega}_c} \sim \alpha_s L \int_0^\infty \mathrm{d}\boldsymbol{k}^2 \frac{\alpha_s^2 n}{(\boldsymbol{k}^2+\mu^2)^2} \sim \alpha_s \frac{L}{\lambda}\,, \tag{2.4}$$

[2] For such soft emissions, medium effects can influence the color coherence properties leading to a modification of the phase space [60, 61].

JHEP02(2023)156

where the integral is dominated by $\boldsymbol{k}^2 \lesssim \mu^2$. This integral gives the proportionality with the inverse mean free path i.e. $n/\mu^2 \sim 1/\lambda$, resulting in an overall factor of medium opacity L/λ. This parametric estimate misses an important logarithmic factor $\sim \ln \bar{\omega}_c/\omega$, see a further discussion in section 3.2, which signals that the simplifications pertaining to the "soft" regime break down at $\omega \approx \bar{\omega}_c$.

For hard emissions, $\omega > \bar{\omega}_c$, we instead get that

$$\omega \frac{\mathrm{d}I}{\mathrm{d}\omega}\bigg|_{\omega>\bar{\omega}_c} \sim \alpha_s^3 nL \int_{\omega/L}^{\infty} \frac{\mathrm{d}\boldsymbol{k}^2}{\boldsymbol{k}^4} \sim \alpha_s \frac{L}{\lambda}\frac{\bar{\omega}_c}{\omega}\,, \tag{2.5}$$

where we used $t_{\mathrm{f}} = 2\omega/\boldsymbol{k}^2 < L$, and neglected the screening mass μ^2 in this parametric regime, since $\langle \boldsymbol{k}^2\rangle \gg \mu^2$. Compared to the soft regime from eq. (2.4) it is suppressed by an additional power of $\bar{\omega}_c/\omega \ll 1$. The complete spectrum in the dilute regime is sketched in figure 2 (left).

Dense media with long formation time ($\boldsymbol{\lambda \ll t_{\mathrm{f}} \ll L}$). In a dense medium we should expect that typically many scatterings occur during the emission process, which is illustrated in the middle of figure 1. This demands a more sophisticated model than what we suggested in eq. (2.2). Nevertheless, we can approximate the total transferred transverse momentum by $\langle \boldsymbol{k}^2\rangle \sim \hat{q}t$, which resembles a random walk for t time in two dimensions, with $\hat{q}$ playing the role of a diffusion constant.[3] This constant determines the typical transverse momentum accumulated per unit length, or $\hat{q} \sim n \sim \mu^2/\lambda$. In this case the formation time becomes

$$t_{\mathrm{f}} = \sqrt{\frac{2\omega}{\hat{q}}}\,. \tag{2.6}$$

This is often called the coherence length, since during the formation time, interference effects between multiple scattering with the medium are active and the gluon feels only one effective scattering center. The accumulated transverse momentum during the splitting process is in this case $\langle \boldsymbol{k}^2\rangle = \sqrt{2\omega\hat{q}}$, which is the celebrated Landau-Pomeranchuk-Migdal (LPM) effect.

Again, comparing the formation time to the medium length, leads to the characteristic energy scale in the dense regime, namely

$$\omega_c = \frac{1}{2}\hat{q}L^2\,, \tag{2.7}$$

and thus $\omega < \omega_c$ for multiple soft scatterings. The maximal possible momentum accumulated via multiple soft scatterings is denoted $\langle \boldsymbol{k}^2\rangle \sim Q_s^2 = \hat{q}L$. The other limiting scale of the multiple scattering regime arises when considering the minimal formation time in this hierarchy, i.e. $t_{\mathrm{f}} > \lambda$, giving rise the scale $\omega > \omega_{\mathrm{BH}}$, (see later in eq. (2.10)). In this case, the accumulated transverse momentum squared reduces to a single soft scattering $\langle \boldsymbol{k}^2\rangle \sim \hat{q}\lambda \sim \mu^2$.

[3] Arbitrary dense medium, would result in overlapping scatterings that description if beyond the scope of this paper. Multiple independent scatterings require well separated scattering centers ($1/\mu \ll \lambda$) see in ref. [32].

JHEP02(2023)156

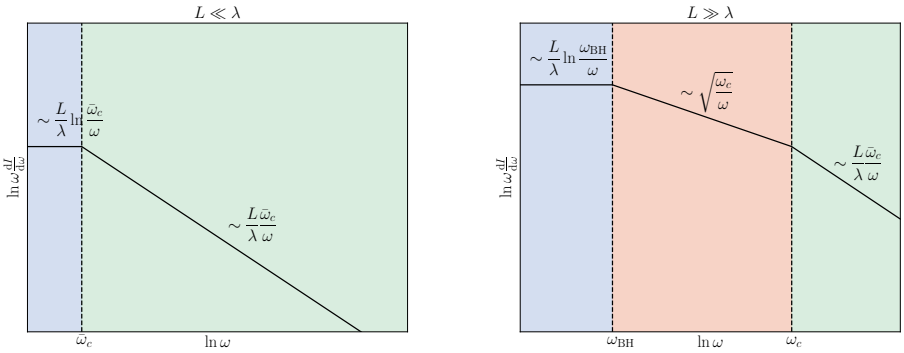


Figure 2. A sketch of the spectrum of medium-induced gluons in a dilute medium $L \ll \lambda$ (left) and in a dense medium $L \gg \lambda$ (right).

In the multiple soft scattering regime, characterized by $\omega_{\mathrm{BH}} \ll \omega \ll \omega_c$, the mean free path has to be replaced by the formation time in eq. (2.4), leading to

$$\omega \frac{\mathrm{d}I}{\mathrm{d}\omega}\bigg|_{\omega_{\mathrm{BH}} \ll \omega \ll \omega_c} \sim \alpha_s \frac{L}{t_{\mathrm{f}}} \sim \alpha_s \sqrt{\frac{\hat{q}L^2}{\omega}} \,. \tag{2.8}$$

This is also often referred to as the BDMPS-Z spectrum in the soft limit.

For hard gluon emissions, $\omega \gg \omega_c$, we also have to demand that $\langle \boldsymbol{k}^2 \rangle \gg Q_s$. In other words, only a hard scattering can provide sufficient transverse momentum to fulfill all the conditions. The relevant contribution is therefore captured by eq. (2.5) and, remarkably, the spectrum in this limit is identical to the hard tail in the dilute regime, namely

$$\omega \frac{\mathrm{d}I}{\mathrm{d}\omega}\bigg|_{\omega \gg \omega_c} \sim \alpha_s \frac{L}{\lambda} \frac{\bar{\omega}_c}{\omega} \,. \tag{2.9}$$

This demonstrates that, even in a dense medium, hard emissions mostly are driven by single, rare hard scattering events. An illustration of this can be seen on the right in figure 1.

Dense media with short formation time ($t_{\mathrm{f}} \ll \lambda \ll L$). The picture we just described should hold as long as there indeed is time for multiple scatterings during the emission process, namely that $t_{\mathrm{f}} > \lambda$. However, when the formation time is short the parton will only have time to scatter once before it splits. This is illustrated on the left in figure 1. The transverse scale is typically soft $\langle \boldsymbol{k}^2 \rangle \sim \mu^2$, and thus $t_{\mathrm{f}} = \frac{2\omega}{\mu^2}$. This regime is characterized by $t_{\mathrm{f}} \ll \lambda$, or equivalently as a condition on the energy $\omega \ll \omega_{\mathrm{BH}}$, where we have defined the scale

$$\omega_{\mathrm{BH}} = \frac{1}{2}\mu^2 \lambda \,. \tag{2.10}$$

Note that $\bar{\omega}_c(\lambda) \equiv \omega_{\mathrm{BH}}$. In this case the spectrum becomes

$$\omega \frac{\mathrm{d}I}{\mathrm{d}\omega}\bigg|_{\omega < \omega_{\mathrm{BH}}} \sim \alpha_s \frac{L}{\lambda} \,, \tag{2.11}$$

which is similar to the result in eq. (2.4) and it is sometimes referred to as Bethe-Heitler region because of the QED analogue. The $t_{\mathrm{f}} \ll \lambda$ condition is satisfactory but not necessary

JHEP02(2023)156

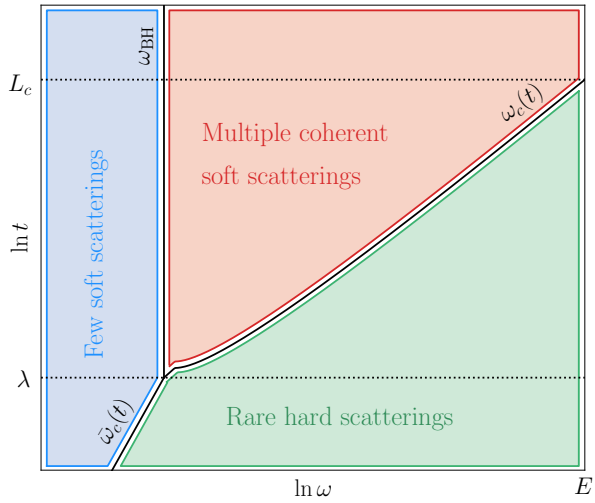


Figure 3. The phase space for medium-induced emissions designated to the leading scattering processes. The main length scales of the problem are λ and the critical length L_c, corresponding to the characteristic energies $\omega_{\rm BH}$ and E, see text for further details.

for BH emissions. There are BH emissions with $t_{\rm f} > \lambda$ but only with one real scattering. We show this more rigorously later in section 3. Our heuristic analysis fails to capture the additional logarithmic term which in this case comes with $\ln \omega_{\rm BH}/\omega$, see section 3.3. A sketch of the spectrum in a dense medium can be found in figure 2 (right).

Summary. Bringing together the heuristic arguments of this section, we show a sketch of the emission spectrum for dilute and dense media in figure 2. The full emission phase space is divided by three lines corresponding to the emergent scales: $\bar{\omega}_c(t) = \frac{1}{2}\mu^2 t$ in the dilute regime ($L < \lambda$), and $\omega_{\rm BH} = \frac{1}{2}\mu^2\lambda$ and $\omega_c(t) = \frac{1}{2}\hat{q}t^2$ in the dense regime ($L > \lambda$) and they are shown in figure 3. The $\omega_c(t)$ line is not completely straight because $\hat{q}$ in general is ω dependent (see in section 3.4). When the medium size is of the order of the mean free path they all collapse to the same value, i.e. $\bar{\omega}_c(\lambda) = \omega_c(\lambda) = \omega_{\rm BH}$. Typically, we adopt a notation where the scales written without the t-argument denote their respective values at L, e.g. $\omega_c \equiv \omega_c(L)$.

These scales delineate three distinct regimes of scattering processes and thus induced emissions that were discussed in the preceding paragraphs. The areas between these scales are governed by few soft, multiple soft and rare hard interactions with the medium, as depicted with colors in figure 3 and discussed above. We also show two length scales: the mean free path λ and the critical medium length L_c. The mean free path marks the time where multiple scatterings appear. The critical medium length indicates where rare, hard scatterings will no longer have an effect, that is where $\omega_c(t) = E$, leading to $L_c = \sqrt{2E/\hat{q}}$.[4]

[4]When considering a finite splitting fraction z the exact definition turns out to be $L_c = \sqrt{E/(2\hat{q})}$.

JHEP02(2023)156

As the parton moves through the medium, at each instant t the emission rate $\omega \mathrm{d}I/(\mathrm{d}\omega \mathrm{d}t)$ is different since the phase space for available emissions changes. The emission spectrum $\omega \mathrm{d}I/\mathrm{d}\omega$, evaluated at a given t (typically $t = L$), includes the accumulated range of processes that occurred up to that time. The main goal of this work is to consider multiple medium-induced emissions in this system of dynamically evolving scales presented in figure 3. In the next section, section 3, we will formally derive the results we only have argued for in this section. Finally, in section 4 we will tackle the issue of multiple emissions in this scheme.

3 Spectrum of medium-induced emissions

In this section, we derive the all order emission spectrum induced by elastic scatterings on a deconfined medium. Of the three expansions we present here, the opacity expansion and the improved opacity expansion have been discussed in depth in previous works, see refs. [35, 36] and [53–57] respectively. The resummed opacity expansion has been argued for before, see [36, 46, 62], but is here derived rigorously for the first time. This paper strives to be a comprehensive reference for all of the expansions, and hence they are all presented in detail. Furthermore, we extend previous calculations to all orders and present results to order $N = 2$ for the opacity expansion and $N_r = 2$ for the resummed opacity expansion, and extract the relevant limits analytically. This provides valuable insight into the underlying structure of the expansions in different regimes.

3.1 General formalism

Currently, we consider the emission of a gluon with energy ω from a parent parton with energy E in the soft limit, i.e. $\omega \ll E$. The soft limit is used in this section because it is much more clear and readable. For a description beyond the strictly soft limit we refer to appendix C, which includes novel results.

Our starting point is the definition of the spectrum of medium-induced gluons [32–34, 36],

$$\omega \frac{\mathrm{d}I}{\mathrm{d}\omega} = \frac{2\alpha_s C_R}{\omega^2} \mathrm{Re} \int_0^\infty \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_1 \, \partial_{\boldsymbol{x}} \cdot \partial_{\boldsymbol{y}} \left[\mathcal{K}(\boldsymbol{x}, t_2; \boldsymbol{y}, t_1) - \mathcal{K}_0(\boldsymbol{x}, t_2; \boldsymbol{y}, t_1) \right]_{\boldsymbol{x}=\boldsymbol{y}=0} , \tag{3.1}$$

where C_R is the Casimir color factor of the emitting particle ($C_R = C_F$ for a quark and $C_R = N_c$ for a gluon).[5] The three-point correlator $\mathcal{K}$ solves the Schrödinger equation

$$\left[i\partial_t + \frac{\partial_{\boldsymbol{x}}^2}{2\omega} + iv(\boldsymbol{x}, t) \right] \mathcal{K}(\boldsymbol{x}, t; \boldsymbol{y}, t_0) = i\delta(t - t_0)\delta(\boldsymbol{x} - \boldsymbol{y}) , \tag{3.2}$$

where the potential $v(\boldsymbol{x}, t)$ describes scatterings in a thermal or quasi-particle like background,[6]

$$v(\boldsymbol{x}, t) = \int_{\boldsymbol{q}} \sigma(\boldsymbol{q}, t) \left(1 - \mathrm{e}^{i\boldsymbol{q}\cdot\boldsymbol{x}} \right) . \tag{3.3}$$

[5] This expression can be derived directly from the fully z dependent spectrum in eq. (C.1), see the discussion in appendix C.

[6] Throughout, we adopt a shorthand notation, so that $\int_p = \int \frac{\mathrm{d}^4 p}{(2\pi)^4}$, $\int_{\boldsymbol{p}} = \int \frac{\mathrm{d}^2 \boldsymbol{p}}{(2\pi)^2}$, and $\int_{\boldsymbol{x}} = \int \mathrm{d}^2 \boldsymbol{x}$.

JHEP02(2023)156

Here $\sigma(\boldsymbol{q},t) = N_c n(t) \mathrm{d}^2\sigma_{\mathrm{el}}/\mathrm{d}^2\boldsymbol{q}$ is proportional to the in-medium elastic scattering cross section, where $n(t) \sim T^3$ is the density of scattering centers.[7] The color factor N_c appears because, in this limit, only the emitted gluon picks up transverse momentum in the medium. The potential can be extracted from an effective theory that accounts for both large and small momentum exchanges with the medium [63]. In the main part of this paper, we will use the Gyulassy-Wang potential [64] that both contains a hard Coulomb-tail and implements screening in the infrared $\mu^2 \gg \boldsymbol{q}^2$,

$$\frac{\mathrm{d}^2\sigma_{\mathrm{el}}}{\mathrm{d}^2\boldsymbol{q}} = \frac{g^4}{(\boldsymbol{q}^2+\mu^2)^2}\,, \tag{3.4}$$

where μ is a screening mass of the order of the Debye mass of a thermal medium. We also provide the spectrum in appendix E with the LO hard thermal loop (HTL) potential [65]. Also, in eq. (3.2) we neglect quark and gluon thermal masses, which corresponds to taking the high-energy limit ($E \gg m_{q,g}^4/\hat{q} \sim \omega_{\mathrm{BH}}$), see e.g. refs. [41, 62] for further discussion.

When the medium is not present, $v = 0$, one recovers the propagation of a single parton in vacuum, $\mathcal{K}(\boldsymbol{x},t_2;\boldsymbol{y},t_1) \equiv \mathcal{K}_0(\boldsymbol{x}-\boldsymbol{y},t_2-t_1)$ with $\mathcal{K}_0(\boldsymbol{x},t) = \frac{\omega}{2\pi i t}\exp[i\,\omega\boldsymbol{x}^2/(2t)]$. To only capture medium effects, the vacuum term is explicitly subtracted in eq. (3.1).

The emission spectrum in eq. (3.1) is the result of a path integral formalism in which arbitrarily many soft and hard scatterings are included. It does, however, not account for longitudinal momentum ($\sim$ energy) exchange with the medium. Having written the expression as a spectrum we also implicitly assume that the creation of the initial parton is factorized from the induced process (for example it was created in a highly virtual vacuum process). Finally, the medium averages leading to the simple form of the three-point correlator, as in eq. (3.2), assumes independent scatterings on the medium. This parametrically holds if the size of the potential is much smaller than the mean free path, i.e. $\mu^{-1} \ll \lambda$, where the typical exchanged momentum is $|\delta\boldsymbol{q}| \sim \mu$ [32].

We should also note that eq. (3.1) emerges as the result of a momentum integral of the differential spectrum $\mathrm{d}I/(\mathrm{d}\omega\,\mathrm{d}^2\boldsymbol{k})$ in the soft limit [39, 57], with no kinematical constraint on the transverse momentum ($\boldsymbol{k}$) integral similarly to refs. [34, 58]. A more careful treatment of the kinematics would be important especially if one is interested in emissions inside or out of a given cone [26, 39, 51].

Let us now cast the equation for the spectrum in an equivalent form. On many occasions it is more practical to work in transverse momentum space,

$$\mathcal{K}(\boldsymbol{p},t_2;\boldsymbol{p}_0,t_1) = \int_{\boldsymbol{x},\boldsymbol{y}} \mathrm{e}^{-i\boldsymbol{p}\cdot\boldsymbol{x}+i\boldsymbol{p}_0\cdot\boldsymbol{y}}\mathcal{K}(\boldsymbol{x},t_2;\boldsymbol{y},t_1)\,. \tag{3.5}$$

The vacuum propagator $\mathcal{K}_0$ then becomes a plane wave, i.e. $\mathcal{K}_0(\boldsymbol{p},t) = \exp[-i\,\boldsymbol{p}^2 t/(2\omega)]$. In this representation, the solution to the Schrödinger equation (3.2) can be written as the

[7] We include the number density of the scattering centers $n(t)$ into $v(\boldsymbol{x},t)$ and $\sigma(\boldsymbol{q},t)$ similarly to the previous works in refs. [53, 54, 57].

JHEP02(2023)156

recursive equation

$$\mathcal{K}(\boldsymbol{p},t_2;\boldsymbol{p}_0,t_1) = (2\pi)^2\delta(\boldsymbol{p}-\boldsymbol{p}_0)\mathcal{K}_0(\boldsymbol{p};t_2-t_1) - \int_{t_1}^{t_2}\mathrm{d}s\int_{\boldsymbol{q}}\mathcal{K}_0(\boldsymbol{p};t_2-s)v(\boldsymbol{q},s)\mathcal{K}(\boldsymbol{p}-\boldsymbol{q},s;\boldsymbol{p}_0,t_1)\,, \tag{3.6}$$

where now

$$v(\boldsymbol{q},s) = (2\pi)^2\delta(\boldsymbol{q})\Sigma(s) - \sigma(\boldsymbol{q},s)\,, \tag{3.7}$$

ensures probability conservation. Here, $\Sigma(s) \equiv \int_{\boldsymbol{q}}\sigma(\boldsymbol{q},s)$ can be interpreted as the inverse of the (local) mean free path λ along a trajectory of a propagating parton, or

$$\lambda(s) = \frac{1}{\Sigma(s)}\,. \tag{3.8}$$

In these expressions, we have assumed that the integral over the elastic scattering cross section exists. In many cases, e.g. for the HTL potential [65], one needs to introduce an IR regulator. However, $v(\boldsymbol{q},s)$ in eq. (3.7) is not sensitive to this IR regulation and therefore the expansion in eq. (3.6) is well-defined. We have provided a further discussion of the HTL potential in appendix E.

The medium-induced spectrum now reads

$$\omega\frac{\mathrm{d}I}{\mathrm{d}\omega} = \frac{2\alpha_s C_R}{\omega^2}\mathrm{Re}\int_0^\infty \mathrm{d}t_2\int_0^{t_2}\mathrm{d}t_1\int_{\boldsymbol{p},\boldsymbol{q}}\boldsymbol{p}\cdot\boldsymbol{q}\left[\mathcal{K}(\boldsymbol{p},t_2;\boldsymbol{q},t_1) - (2\pi)^2\delta(\boldsymbol{p}-\boldsymbol{q})\mathcal{K}_0(\boldsymbol{p},t_2-t_1)\right]. \tag{3.9}$$

The vacuum contribution can then be removed by inserting eq. (3.6) into eq. (3.9), yielding

$$\omega\frac{\mathrm{d}I}{\mathrm{d}\omega} = \frac{4\alpha_s C_R}{\omega}\mathrm{Re}\,i\int_0^L \mathrm{d}t_2\int_0^{t_2}\mathrm{d}t_1\int_{\boldsymbol{p},\boldsymbol{p}_0,\boldsymbol{q}}\frac{\boldsymbol{p}\cdot\boldsymbol{p}_0}{\boldsymbol{p}^2}v(\boldsymbol{q},t_2)\mathcal{K}(\boldsymbol{p}-\boldsymbol{q},t_2;\boldsymbol{p}_0,t_1)\,, \tag{3.10}$$

where we regulated the integral over the latter time coordinate using an adiabatic turn-off at infinity (see also ref. [46]). The other time integrals are limited by the extent of the medium L. Noticing, that

$$\int_{\boldsymbol{p}}\frac{p^i}{\boldsymbol{p}^2}v(\boldsymbol{p}-\boldsymbol{k},s) = \frac{k^i}{\boldsymbol{k}^2}\Sigma(\boldsymbol{k}^2,s)\,, \tag{3.11}$$

where $\Sigma(\boldsymbol{k}^2,s) = \int_{\boldsymbol{q}}\sigma(\boldsymbol{q},s)\Theta(\boldsymbol{q}^2-\boldsymbol{k}^2)$,[8] we obtain

$$\omega\frac{\mathrm{d}I}{\mathrm{d}\omega} = \frac{4\alpha_s C_R}{\omega}\mathrm{Re}\,i\int_0^L \mathrm{d}t_2\int_0^{t_2}\mathrm{d}t_1\int_{\boldsymbol{p},\boldsymbol{p}_0}\Sigma(\boldsymbol{p}^2,t_2)\frac{\boldsymbol{p}\cdot\boldsymbol{p}_0}{\boldsymbol{p}^2}\mathcal{K}(\boldsymbol{p},t_2;\boldsymbol{p}_0,t_1)\,. \tag{3.12}$$

While the above results are valid for any medium potential, in this work we will focus on the GW scattering potential, defined in eq. (3.4). In this case, we find that

$$\Sigma(\boldsymbol{k}^2,s) = \frac{\hat{q}_0(s)}{\boldsymbol{k}^2+\mu^2}\,, \tag{3.13}$$

where $\hat{q}_0(s) = 4\pi\alpha_s^2 N_c n(s)$ is a measure of the scattering density. Currently, we consider a medium of constant density, $n(s) = n_0$.

[8] Also $\Sigma(0,s) = \Sigma(s)$, consistent with the definition in eq. (3.7).

JHEP02(2023)156

The spectrum, given by eq. (3.1) or eq. (3.12), can be evaluated using numerical techniques [47–52] or by employing analytical approximations. As we will see the approximate approaches rely on expanding the problem as a series, which will give the true answer at infinite order. The different series have different radii of convergence, and none of them will alone converge for all L and ω, meaning more than one have to be employed. In most cases, however, the first order expansion is sufficient to provide an accurate approximation of the all order result. In the following we discuss three well-defined approaches that together provide an accurate description of the true problem for all L and ω, called: the opacity expansion, the resummed opacity expansion, and the improved opacity expansion. We will derive these, discuss their limits and their regions of validity.

We also point out that the medium parameters for the numerical evaluations in figures 4–6 (right) are chosen to maximally separate the relevant scales and to illustrate the main features of the spectrum. They are also similar to the ones used in phenomenological studies [24, 26, 57]. It is worth pointing out that, although this particular choice violates the assumption of non-overlapping scattering centers and should be treated with care, changing the values of the parameters would not alter the qualitative picture of separating different regimes in the (ω, t) plane.

3.2 Opacity expansion (OE)

The opacity expansion of the spectrum arises when inserting eq. (3.6) directly into eq. (3.12), and was developed in refs. [36, 66].[9] The truncation of this series at a given order n in the medium scattering potential gives the $N = n$ term, which is by definition proportional to $(L/\lambda)^n$ (see eq. (A.2)). Physically this means, at $N = n$ one counts n number of scatterings (both with and without momentum exchange) on the full elastic potential. The relevant energy scales that arise are $\bar{\omega}_c = \frac{1}{2}\mu^2 L$, and $\frac{L}{\lambda}\bar{\omega}_c = \frac{1}{2}\hat{q}_0 L^2$ as discussed in section 2. A general formula for the spectrum at any order is derived in appendix A, and with finite-z corrections in appendix C. These results are used in the following calculations, and we will refer to the appendices for more details.

First order ($N = 1$). The spectrum at first order of opacity is well known [35, 36]. Since eq. (3.12) already includes at least one scattering, we obtain the $N = 1$ term by replacing the full propagator $\mathcal{K}$ by the vacuum one. We then find,

$$\omega \frac{\mathrm{d}I^{N=1}}{\mathrm{d}\omega} = 8\pi\bar{\alpha}\frac{L}{\lambda}\frac{\bar{\omega}_c}{\omega}\int_{\boldsymbol{p}} \tilde{\Sigma}(\boldsymbol{p}^2)\,\mathrm{Re}\,i\int_0^1 \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_0\, \mathrm{e}^{-i\boldsymbol{p}^2(t_1-t_0)}\,, \tag{3.14}$$

where we have switched to dimensionless integration variables by defining $\boldsymbol{p}^2 \to \boldsymbol{p}^2 L/(2\omega)$ and $t \to t/L$, and where $\tilde{\Sigma}(\boldsymbol{p}^2) = (\boldsymbol{p}^2 + \bar{\omega}_c/\omega)^{-1}$. This expression can also be obtained from the general $N = n$ result in eq. (A.2). After simplifications, the spectrum becomes

$$\omega \frac{\mathrm{d}I^{N=1}}{\mathrm{d}\omega} = 2\bar{\alpha}\frac{L}{\lambda}\frac{\bar{\omega}_c}{\omega}\int_0^\infty \mathrm{d}p\, \frac{1}{p + \frac{\bar{\omega}_c}{\omega}}\frac{p - \sin p}{p^2}\,, \tag{3.15}$$

[9] To be precise, expanding our formulas order by order in opacity reproduces the expansion defined in ref. [36], which reproduces ref. [66] in the "incoherent" limit.

JHEP02(2023)156

where $\bar{\alpha} = \alpha_s C_R/\pi$ and $\lambda = \mu^2/\hat{q}_0$. It also agrees with eq. (6.7) in ref. [36] (see also in ref. [39]). We recognize the dependence on the medium opacity L/λ and the ratio $\bar{\omega}_c/\omega$. The remaining integral can be done analytically, but the resulting expression is not very illuminating. However, the limiting behavior can readily be extracted,

$$\omega \frac{\mathrm{d}I^{N=1}}{\mathrm{d}\omega} \simeq \begin{cases} 2\bar{\alpha}\frac{L}{\lambda}\left(\ln\frac{\bar{\omega}_c}{\omega} - 1 + \gamma_E\right), & \text{for } \omega \ll \bar{\omega}_c\,, \\ \frac{\pi}{2}\bar{\alpha}\frac{L}{\lambda}\frac{\bar{\omega}_c}{\omega}, & \text{for } \omega \gg \bar{\omega}_c\,. \end{cases} \tag{3.16}$$

This agrees well with the heuristic discussion in section 2. In particular, we identify a logarithmic behavior $\sim \ln\frac{\bar{\omega}_c}{\omega}$ in the infrared. Notice the different expansion structures in the soft $\sim \bar{\alpha}\frac{L}{\lambda}$ and in the hard $\sim \bar{\alpha}\frac{L}{\lambda}\frac{\bar{\omega}_c}{\omega}$ limits, which we will come back to.

Second order ($N = 2$). The calculation for $N = 2$ follows in a similar way, leading to

$$\begin{aligned} \omega \frac{\mathrm{d}I^{N=2}}{\mathrm{d}\omega} = -8\pi\bar{\alpha}\left(\frac{L}{\lambda}\right)^2 \frac{\bar{\omega}_c}{\omega} \int_{\boldsymbol{p}_2,\boldsymbol{p}_1} \tilde{\Sigma}(\boldsymbol{p}_2^2)\frac{\boldsymbol{p}_2\cdot\boldsymbol{p}_1}{\boldsymbol{p}_2^2}\tilde{v}(\boldsymbol{p}_2-\boldsymbol{p}_1) \\ \times \mathrm{Re}\, i\int_0^1 \mathrm{d}t_2 \int_0^{t_2}\mathrm{d}t_1 \int_0^{t_1}\mathrm{d}t_0\, \mathrm{e}^{-i\boldsymbol{p}_2^2(t_2-t_1)}\mathrm{e}^{-i\boldsymbol{p}_1^2(t_1-t_0)}\,, \end{aligned} \tag{3.17}$$

with dimensionless integration variables, and where $\tilde{v}(\boldsymbol{p}) = (2\pi)^2\delta(\boldsymbol{p}) - \frac{\bar{\omega}_c}{\omega}\tilde{\sigma}(\boldsymbol{p})$. In the GW model, $\tilde{\sigma}(\boldsymbol{p}) = 4\pi/(\boldsymbol{p}^2 + \frac{\bar{\omega}_c}{\omega})^2$. After inserting $\tilde{v}$, doing the time integrals and simplifying this can be written as

$$\omega \frac{\mathrm{d}I^{N=2}}{\mathrm{d}\omega} = -4\bar{\alpha}\left(\frac{L}{\lambda}\right)^2 \frac{\bar{\omega}_c}{\omega}\left[\mathcal{I}_1\left(\frac{\bar{\omega}_c}{\omega}\right) - \frac{\bar{\omega}_c}{\omega}\mathcal{I}_2\left(\frac{\bar{\omega}_c}{\omega}\right)\right], \tag{3.18}$$

where we have defined the integrals

$$\mathcal{I}_1\left(\frac{\bar{\omega}_c}{\omega}\right) = \int_0^\infty \mathrm{d}p\, \frac{1}{p+\frac{\bar{\omega}_c}{\omega}}\frac{1-\cos p - \frac{p}{2}\sin p}{p^3}\,, \tag{3.19}$$

$$\begin{aligned} \mathcal{I}_2\left(\frac{\bar{\omega}_c}{\omega}\right) = \int_0^\infty \mathrm{d}p_2 \int_0^\infty \mathrm{d}p_1\, \frac{p_1}{p_2+\frac{\bar{\omega}_c}{\omega}}\frac{1}{\left[\left(p_2+p_1+\frac{\bar{\omega}_c}{\omega}\right)^2 - 4p_2p_1\right]^{3/2}} \\ \times \frac{1}{p_2-p_1}\left[\frac{1}{p_1^2}(1-\cos p_1) - \frac{1}{p_2^2}(1-\cos p_2)\right]. \end{aligned} \tag{3.20}$$

The $\mathcal{I}_1$ integral can be done analytically, but $\mathcal{I}_2$ is more complicated. It can be shown that it is much smaller than $\mathcal{I}_1$ in the soft limit. In the hard limit, $\mathcal{I}_1$ and $\mathcal{I}_2$ cancel at the order of $\mathcal{O}(\frac{\bar{\omega}_c}{\omega})$, leaving a positive contribution going as $\mathcal{O}(\frac{\bar{\omega}_c}{\omega})^2$. In summary,

$$\omega \frac{\mathrm{d}I^{N=2}}{\mathrm{d}\omega} \simeq \begin{cases} -\bar{\alpha}\left(\frac{L}{\lambda}\right)^2, & \text{for } \omega \ll \bar{\omega}_c\,, \\ \sim \bar{\alpha}\left(\frac{L}{\lambda}\right)^2\left(\frac{\bar{\omega}_c}{\omega}\right)^2, & \text{for } \omega \gg \bar{\omega}_c\,. \end{cases} \tag{3.21}$$

We notice that the $N = 2$ is proportional to $\bar{\alpha}(\frac{L}{\lambda})^2$ in the soft limit, and goes like $\bar{\alpha}(\frac{L}{\lambda}\frac{\bar{\omega}_c}{\omega})^2$ in the hard limit. This immediately implies that $N = 2$ is always subleading to $N = 1$ if the medium is dilute $L \ll \lambda$ or if the emission is hard $\omega \gg \frac{L}{\lambda}\bar{\omega}_c$. Given the structure of the expansion, we expect the previous statement to hold at arbitrary $N = n$ order. This is in agreement with the earlier, heuristic observation in refs. [41, 67].

JHEP02(2023)156

Based on the limits for $N = 1$ and $N = 2$, given by eqs. (3.16) and (3.21), in the regimes where the expansion holds the all order OE spectrum is expected to take the form

$$\omega \frac{\mathrm{d}I}{\mathrm{d}\omega} = \begin{cases} \bar{\alpha} \sum_{n=1}^{\infty} \left(\frac{L}{\lambda}\right)^n h_n \left(\frac{\omega}{\bar{\omega}_c}\right), & \omega \ll \bar{\omega}_c\,, \\ \bar{\alpha} \sum_{n=1}^{\infty} \left(\frac{L}{\lambda}\frac{\bar{\omega}_c}{\omega}\right)^n \tilde{h}_n \left(\frac{\bar{\omega}_c}{\omega}\right), & \omega \gg \bar{\omega}_c\,, \end{cases} \tag{3.22}$$

where the OE coefficients $h_n, \tilde{h}_n$ are finite and can be calculated order by order. Note that we have not strictly proven this for all orders, although our $N = 1$ and $N = 2$ results strongly indicate this structure. In the soft limit $\omega \ll \bar{\omega}_c$, the OE expansion converges rapidly, defining the expected "naive" radius of convergence $L/\lambda < 1$. However, in the hard limit $\omega \gg \bar{\omega}_c$ there is convergence even if the medium is big, provided $\frac{L}{\lambda}\frac{\bar{\omega}_c}{\omega} < 1$. The full region of convergence is shown in green in the left panel of figure 4. Outside of this region we expect higher orders to grow uncontrollably and hence the OE is not valid when truncated at any finite order.

The resulting spectrum from eqs. (3.15) and (3.18) is shown in the right panel of figure 4 for different propagation lengths (labeled with t). For short lengths $t < \lambda$ the OE is valid for all ω. For $t > \lambda$, the OE is only valid if $\frac{t}{\lambda}\frac{\bar{\omega}_c}{\omega} < 1$ (see also the green region on the left panel). We note that the $N = 2$ correction becomes important at $t > \lambda$ and $\omega \approx \frac{t}{\lambda}\bar{\omega}_c$ (the latter constraint shown as bullets in the figure). For larger media, the grey bullets, representing the minimal energy for achieving convergence, moves to higher values, and the truncated OE series at ω smaller than this becomes ill-defined. This can be seen, for instance, in the upper line in figure 4 (right) for a medium length of $t = 4\,\mathrm{fm}$. We have also compared to a full numerical evaluation of the spectrum from refs. [51, 68]. The figure shows that this indeed is well approximated by the OE in its region of validity, as we have argued. Our curves are not expected to hold in the limit $\omega \ll 1\,\mathrm{GeV}$, where several important effects were not taken into account such as thermal masses, realistic 2–2 elastic scatterings, and other non-perturbative effects. We still plot the curves down to very small ω to compare the different expansion schemes.

3.3 Resummed opacity expansion (ROE)

Next, we turn to dense media, $L \gg \lambda$, where multiple scattering have to be accounted for. However, as depicted to the left in figure 1 and discussed in section 2, for soft emissions with short formation times, a single scattering still gives the leading contribution to the spectrum. This defines the so-called Bethe-Heitler regime named after the QED analogue of this process. A qualitative argument of this regime was first given in ref. [36] and later developed in ref. [46], see also in ref. [62] which coined the name "resummed opacity expansion" (ROE).[10]

[10] In ref. [32], authors derive the opacity expansion from the all-order formula eq. (3.1) in a similar way as we did by expanding $\mathcal{K}$. Accidentally, in one of their intermediate steps in section 4, they kept the virtual interactions resummed, which corresponds to our ROE. Back then, however, they did not realize the importance of that formula and they expanded it to reproduce the OE.

JHEP02(2023)156

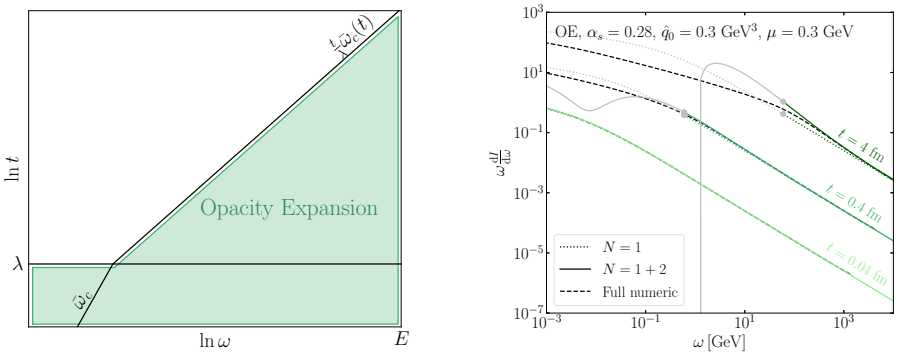


Figure 4. *Left:* the sketch of the region of validity (and convergence) of the opacity expansion for different propagation length t and emission energy ω. *Right:* the induced emission spectrum for gluons in the opacity expansion. The gray part of the curves denotes regions, where the expansion is not valid. Using the parameters presented, $\lambda = 0.06\,\mathrm{fm}$. The full numeric solution is also presented with dashed lines.

Here, we provide for the first time a consistent framework for dealing with an expansion of the *real* scatterings with the medium, whereby real we mean interactions with a finite transverse momentum exchange. At the same time, an all-order resummation of the corresponding *virtual* interactions, with zero transverse momentum exchange, is performed.

Dividing eq. (3.6) by the vacuum propagator leaves us with

$$\frac{\mathcal{K}(\boldsymbol{p},t;\boldsymbol{p}_0,t_0)}{\mathcal{K}_0(\boldsymbol{p};t-t_0)} = (2\pi)^2\delta(\boldsymbol{p}-\boldsymbol{p}_0) - \int_{t_0}^{t}\mathrm{d}s\int_{\boldsymbol{q}} v(\boldsymbol{q},s)\frac{\mathcal{K}(\boldsymbol{p}-\boldsymbol{q},s;\boldsymbol{p}_0,t_0)}{\mathcal{K}_0(\boldsymbol{p};s-t_0)}\,. \tag{3.23}$$

Next, taking a derivative with respect to the latest time results in

$$\frac{\partial}{\partial t}\frac{\mathcal{K}(\boldsymbol{p},t;\boldsymbol{p}_0,t_0)}{\mathcal{K}_0(\boldsymbol{p};t-t_0)} + \Sigma(t)\frac{\mathcal{K}(\boldsymbol{p},t;\boldsymbol{p}_0,t_0)}{\mathcal{K}_0(\boldsymbol{p};t-t_0)} = \int_{\boldsymbol{q}}\sigma(\boldsymbol{q})\frac{\mathcal{K}(\boldsymbol{p}-\boldsymbol{q},t;\boldsymbol{p}_0,t_0)}{\mathcal{K}_0(\boldsymbol{p};t-t_0)}\,, \tag{3.24}$$

where $\Sigma(t) = \int_{\boldsymbol{q}}\sigma(\boldsymbol{q},t)$, as before. This motivates defining the elastic Sudakov factor,

$$\Delta(t,t_0) \equiv \mathrm{e}^{-\int_{t_0}^{t}\mathrm{d}s\,\Sigma(s)} = \mathrm{e}^{-(t-t_0)\Sigma}\,, \tag{3.25}$$

where the last equality holds for media with constant density.[11] This represents the probability of no elastic scattering occurring between times t_0 and t. Integrating out the time, we arrive at a slightly modified iterative equation

$$\begin{aligned}\mathcal{K}(\boldsymbol{p},t;\boldsymbol{p}_0,t_0) &= (2\pi)^2\delta(\boldsymbol{p}-\boldsymbol{p}_0)\Delta(t,t_0)\mathcal{K}_0(\boldsymbol{p};t-t_0)\\ &\quad + \int_{t_0}^{t}\mathrm{d}s\frac{\Delta(t,t_0)}{\Delta(s,t_0)}\int_{\boldsymbol{q}}\mathcal{K}_0(\boldsymbol{p};t-s)\sigma(\boldsymbol{q},s)\mathcal{K}(\boldsymbol{p}-\boldsymbol{q},s;\boldsymbol{p}_0,t_0)\,.\end{aligned} \tag{3.26}$$

[11] For medium potentials with unscreened soft divergences, such as the HTL potential, one has to modify this prescription to include an IR regulator. We refer to appendix E for a further discussion.

JHEP02(2023)156

Compared to the standard opacity expansion in eq. (3.6), this looks very similar. However, the expansion is not in the potential $v(\boldsymbol{q}, s)$, which contains both a "real" and a "virtual" part, but in the potential scattering $\sigma(\boldsymbol{q}, s)$ which comes from the term that provides a finite momentum transfer in the process. The virtual contributions, where no net momentum was exchanged, are accounted for to all orders in the Sudakov factor. This is why this expansion referred as a *resummed* opacity expansion.

The relevant scale that appears at high opacity is $\omega_{\text{BH}} = \frac{1}{2}\mu^2\lambda$, as we discussed in section 2. Interestingly, at low opacity $L \ll \lambda$, the scale changes to $\bar{\omega}_c = \frac{1}{2}\mu^2 L$, which we recognize from the opacity expansion. In this regime the ROE is actually equivalent to the OE if one gathers all the terms up to the same order in $(\frac{L}{\lambda})^n$. However, as we will see, the terms are reshuffled in the ROE compared to the OE.

It is possible to reach a general formula for the resummed opacity expansion at arbitrary order. This was done in appendix A.2, and we refer to that section for detailed calculations.

First order ($N_r = 1$). The first order can be obtained from eq. (A.9) with $n = 1$, and reads

$$\omega \frac{\mathrm{d}I^{N_r=1}}{\mathrm{d}\omega} = -8\pi\bar{\alpha}\frac{L}{\lambda}\frac{\bar{\omega}_c}{\omega}\int_{\boldsymbol{p}} \tilde{\Sigma}(\boldsymbol{p}^2)\,\mathrm{Im}\,T\left(\boldsymbol{p}^2 - i\frac{L}{\lambda}\right), \tag{3.27}$$

in re-scaled, dimensionless variables. Equation (3.27) corresponds to the formula (eq. (4.6)) in ref. [46], but is here derived more rigorously. Here, we have defined the function

$$T(x) = \int_0^1 \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_0\, \mathrm{e}^{-ix(t_1-t_0)} = \frac{1 - ix - \mathrm{e}^{-ix}}{x^2}. \tag{3.28}$$

The real and imaginary parts of $T(\boldsymbol{p}^2 - i\chi)$ are given in eq. (A.10). After doing the angular integral this becomes

$$\omega \frac{\mathrm{d}I^{N_r=1}}{\mathrm{d}\omega} = -2\bar{\alpha}\frac{L}{\lambda}\frac{\bar{\omega}_c}{\omega}\int_0^\infty \mathrm{d}p\,\frac{1}{p + \frac{\bar{\omega}_c}{\omega}}\,\mathrm{Im}\,T\left(p - i\frac{L}{\lambda}\right). \tag{3.29}$$

At low opacity $L \ll \lambda$, the function $T(p - iL/\lambda)$ becomes

$$-\mathrm{Im}\,T(p)|_{L\ll\lambda} = \frac{p - \sin(p)}{p^2}, \tag{3.30}$$

making it equivalent to the OE result in eq. (3.15). The limiting behavior in the relevant limits of eq. (3.29) can be extracted, leading to

$$\omega \frac{\mathrm{d}I^{N_r=1}}{\mathrm{d}\omega} \simeq \begin{cases} 2\bar{\alpha}\frac{L}{\lambda}\left(\ln\left(\frac{\bar{\omega}_c}{\omega}\right) - 1 + \gamma_E\right) - \bar{\alpha}\left(\frac{L}{\lambda}\right)^2\left(1 - \pi\frac{\omega}{\bar{\omega}_c}\right), & \text{for } \omega \ll \bar{\omega}_c\,, \\ \frac{\pi\bar{\alpha}}{2}\frac{L}{\lambda}\frac{\bar{\omega}_c}{\omega} - \frac{\pi}{6}\bar{\alpha}\left(\frac{L}{\lambda}\right)^2\frac{\bar{\omega}_c}{\omega}, & \text{for } \omega \gg \bar{\omega}_c\,. \end{cases} \tag{3.31}$$

At leading order in $\mathcal{O}(\frac{L}{\lambda})$ this is the same as the $N = 1$ opacity expansion, presented in eq. (3.16). However, in contrast to the OE, subleading "$N = 2$"-like terms $\sim (\frac{L}{\lambda})^2$ appear, which only will be relevant when compared to higher-order contributions at $N_r = 2$.

JHEP02(2023)156

In the high opacity limit $L \gg \lambda$, we have to extract the relevant limit of $T(p - i\chi)$ in a careful way, yielding

$$-\mathrm{Im}\, T(p - i\chi)|_{L\gg\lambda} \simeq \frac{p}{(L/\lambda)^2 + p^2}\,. \tag{3.32}$$

Changing the integration variable to $q = p\lambda/L$, we observe that $\omega_{\rm BH}$ replaces $\bar{\omega}_c$ as the relevant scale, and eq. (3.29) becomes

$$\omega\frac{\mathrm{d}I^{N_r=1}}{\mathrm{d}\omega} \simeq 2\bar{\alpha}\frac{L}{\lambda}\frac{\omega_{\rm BH}}{\omega}\int_0^\infty \mathrm{d}q\,\frac{1}{q+\frac{\omega_{\rm BH}}{\omega}}\frac{q}{1+q^2} = 2\bar{\alpha}\frac{L}{\lambda}\frac{\omega_{\rm BH}}{\omega}\frac{\frac{\pi}{2}+\frac{\omega_{\rm BH}}{\omega}\ln\left(\frac{\omega_{\rm BH}}{\omega}\right)}{1+\left(\frac{\omega_{\rm BH}}{\omega}\right)^2}\,. \tag{3.33}$$

Finally, one can extract the soft and hard limits of this expression, which are given by

$$\omega\frac{\mathrm{d}I^{N_r=1}}{\mathrm{d}\omega} \simeq \begin{cases} 2\bar{\alpha}\frac{L}{\lambda}\left(\ln\left(\frac{\omega_{\rm BH}}{\omega}\right)+\frac{\pi}{2}\frac{\omega}{\omega_{\rm BH}}\right), & \text{for } \omega \ll \omega_{\rm BH}\,, \\ \pi\bar{\alpha}\frac{L}{\lambda}\frac{\omega_{\rm BH}}{\omega}, & \text{for } \omega \gg \omega_{\rm BH}\,. \end{cases} \tag{3.34}$$

The soft limit agrees with the heuristic discussion in section 2. Strikingly, we see that the behavior in the soft and hard limit takes exactly the same form as for $N = 1$ except that $\bar{\omega}_c$ has been replaced by $\omega_{\rm BH}$ (note that $\bar{\omega}_c(L = \lambda) = \omega_{\rm BH}$).

Second order ($N_r = 2$). The second order is found from eq. (A.9) with $n = 2$, and reads

$$\begin{aligned}\omega\frac{\mathrm{d}I^{N_r=2}}{\mathrm{d}\omega} &= 8\pi\bar{\alpha}\left(\frac{L}{\lambda}\right)^2\left(\frac{\bar{\omega}_c}{\omega}\right)^2\int_{\boldsymbol{p}_2,\boldsymbol{p}_1}\tilde{\Sigma}(\boldsymbol{p}_2^2)\frac{\boldsymbol{p}_2\cdot\boldsymbol{p}_1}{\boldsymbol{p}_2^2}\tilde{\sigma}(\boldsymbol{p}_2-\boldsymbol{p}_1) \\ &\quad\times\frac{1}{\boldsymbol{p}_2^2-\boldsymbol{p}_1^2}\left(\mathrm{Re}\,T(\boldsymbol{p}_1^2-i\chi)-\mathrm{Re}\,T(\boldsymbol{p}_2^2-i\chi)\right)\,.\end{aligned} \tag{3.35}$$

After going to polar coordinates and doing the angular integrals, this becomes

$$\begin{aligned}\omega\frac{\mathrm{d}I^{N_r=2}}{\mathrm{d}\omega} &= 4\bar{\alpha}\left(\frac{L}{\lambda}\right)^2\left(\frac{\bar{\omega}_c}{\omega}\right)^2\int_0^\infty\mathrm{d}p_2\int_0^\infty\mathrm{d}p_1\,\frac{1}{p_2+\frac{\bar{\omega}_c}{\omega}}\frac{p_1}{\left[\left(p_1+p_2+\frac{\bar{\omega}_c}{\omega}\right)^2-4p_1p_2\right]^{3/2}} \\ &\quad\times\frac{1}{p_2-p_1}\left(\mathrm{Re}\,T(p_1-i\chi)-\mathrm{Re}\,T(p_2-i\chi)\right)\,.\end{aligned} \tag{3.36}$$

We study this expression separately in the low- and high-opacity limits.

In the low opacity limit $L \ll \lambda$, the spectrum becomes

$$\begin{aligned}\omega\frac{\mathrm{d}I^{N_r=2}}{\mathrm{d}\omega} &\simeq 4\bar{\alpha}\left(\frac{L}{\lambda}\right)^2\left(\frac{\bar{\omega}_c}{\omega}\right)^2\int_0^\infty\mathrm{d}p_2\int_0^\infty\mathrm{d}p_1\,\frac{1}{p_2+\frac{\bar{\omega}_c}{\omega}}\frac{p_1}{\left[\left(p_1+p_2+\frac{\bar{\omega}_c}{\omega}\right)^2-4p_1p_2\right]^{3/2}} \\ &\quad\times\frac{1}{p_2-p_1}\left(\frac{1-\cos p_1}{p_1^2}-\frac{1-\cos p_2}{p_2^2}\right)\,,\end{aligned} \tag{3.37}$$

where again the only relevant energy scale is $\bar{\omega}_c$, as it is in the OE. The double momentum integral can be recognized as $\mathcal{I}_2$ from $N = 2$ of the opacity expansion. The soft and hard limits are

$$\omega\frac{\mathrm{d}I^{N_r=2}}{\mathrm{d}\omega} \simeq \begin{cases} \pi\bar{\alpha}\left(\frac{L}{\lambda}\right)^2\frac{\omega}{\bar{\omega}_c}, & \text{for } \omega \ll \bar{\omega}_c\,, \\ \frac{\pi}{6}\bar{\alpha}\left(\frac{L}{\lambda}\right)^2\frac{\bar{\omega}_c}{\omega}, & \text{for } \omega \gg \bar{\omega}_c\,. \end{cases} \tag{3.38}$$

JHEP02(2023)156

Summing up the two first orders of the ROE and OE we see that $N_r = 1 + 2$ agrees with $N = 1 + 2$, but only when keeping the subleading $\sim (L/\lambda)^2$ terms at order $N_r = 1$. As mentioned before, the opacity expansion is arranged so that the order $N = n$ only contains terms where the opacity scales as $\sim \chi^n$, where $\chi = L/\lambda$. The resummed opacity expansion also includes all of the same terms, but they are spread out over different orders of the expansion due to the resummation contained in the Sudakov factor. The orders $N_r < n$ do contain terms going as χ^n. To get the right term at order χ^n in the ROE one therefore has to keep the subleading corrections going as χ^n at all previous orders of the expansion. For this reason the opacity expansion is more convenient to use in the low opacity limit, as it does not mix orders of opacity.

In the high opacity limit $L \gg \lambda$, we get

$$\omega \frac{\mathrm{d}I^{N_r=2}}{\mathrm{d}\omega} \simeq 4\bar{\alpha}\frac{L}{\lambda}\left(\frac{\omega_{\mathrm{BH}}}{\omega}\right)^2 \int_0^\infty \mathrm{d}p_2 \int_0^\infty \mathrm{d}p_1 \frac{1}{p_2 + \frac{\bar{\omega}_c}{\omega}} \frac{p_1}{\left[\left(p_1 + p_2 + \frac{\bar{\omega}_c}{\omega}\right)^2 - 4p_1p_2\right]^{3/2}} \times \frac{p_2 + p_1}{(1 + p_2^2)(1 + p_1^2)} \,. \tag{3.39}$$

The soft and hard limits of this expression are given by

$$\omega \frac{\mathrm{d}I^{N_r=2}}{\mathrm{d}\omega} \simeq \begin{cases} \pi\bar{\alpha}\frac{L}{\lambda}\frac{\omega}{\omega_{\mathrm{BH}}}, & \text{for } \omega \ll \omega_{\mathrm{BH}} \,, \\ \pi\bar{\alpha}\frac{L}{\lambda}\frac{\omega_{\mathrm{BH}}}{\omega}, & \text{for } \omega \gg \omega_{\mathrm{BH}} \,, \end{cases} \tag{3.40}$$

where similarly to $N_r = 1$, the relevant scale is now ω_{BH}. Both $N_r = 1$ and 2 goes as $\sim \bar{\alpha}\frac{L}{\lambda}$, however, in the soft limit $N_r = 1$ dominates, while in the hard limit $\mathrm{d}I^{N_r=2} \sim \mathrm{d}I^{N_r=1}$. This shows that ROE is quickly convergent if $\omega \ll \omega_{\mathrm{BH}}$, while the expansion appears to break down for harder emissions. We expect this structure to appear to all orders in $N_r = n$. The resulting validity of the expansion is shown in the left of figure 5 in blue. Based on our findings, the expansion scheme for the ROE at high opacity $L \gg \lambda$ is

$$\omega \frac{\mathrm{d}I}{\mathrm{d}\omega} = \bar{\alpha}\frac{L}{\lambda}\sum_{n=0}^{\infty} f_n\left(\frac{\omega}{\omega_{\mathrm{BH}}}\right) \,, \tag{3.41}$$

where f_n is a finite function that can be obtained order by order for $\omega \ll \omega_{\mathrm{BH}}$.

The resulting spectrum is shown in the right of figure 5 for different propagation lengths. For short times ($t < \lambda$), the ROE is valid for all ω and it gives the same spectrum as the OE (compare to the right panel of figure 4). For longer propagation the ROE is only valid if $\omega < \omega_{\mathrm{BH}}$, which is denoted with bullets in the figure. Outside of the valid region, the curves turn to gray (see also the left panel). Based on the figure, $N_r = 2$ has negligible contribution to the spectrum until $t \approx \lambda$ or $\omega \approx \omega_{\mathrm{BH}}$. Again, the dashed line represents the full numerical evaluation of the spectrum from refs. [51, 68] which is well approximated by the ROE in its region of validity.

3.4 Improved opacity expansion (IOE)

The final expansion scheme we consider is the improved opacity expansion, introduced in refs. [53–57]. We saw that the ROE at high opacity does not work for energies much higher than ω_{BH}. This makes sense since at higher energies the formation time t_{f} of the

JHEP02(2023)156

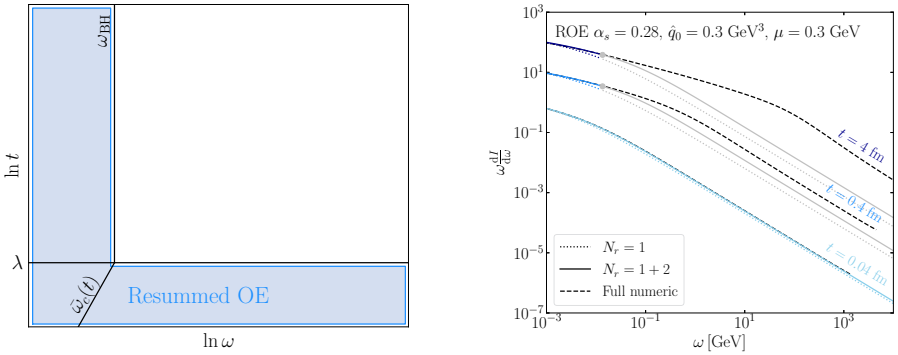


Figure 5. *Left:* the sketch of region of validity (and convergence) of the resummed opacity expansion for different propagation length t and emission energy ω. *Right:* the induced emission spectrum for gluons in the resummed opacity expansion. The gray part of the curves denotes regions, where the expansion is not valid. With the parameters presented, $\lambda = 0.06\,\text{fm}$. The full numeric solution is also presented with dashed lines.

emission becomes bigger than the mean free path, implying that the parton will scatter many times on the medium. The main motivation for the improved opacity expansion is to resum multiple soft scatterings, while account perturbatively for rare, hard scatterings. This is achieved by introducing a scale Q^2 that separates soft and hard scatterings in the scattering potential,

$$v(\boldsymbol{x},t) \approx v_{\rm HO}(\boldsymbol{x},t) + \delta v(\boldsymbol{x},t)\,, \tag{3.42}$$

where $v_{\rm HO}(\boldsymbol{x},t) = \frac{\hat{q}(t)}{4}\boldsymbol{x}^2$ and $\delta v(\boldsymbol{x},t) = \frac{\hat{q}_0(t)}{4}\boldsymbol{x}^2 \ln\frac{1}{Q^2\boldsymbol{x}^2}$. Equation (3.42) is the $\mu|\boldsymbol{x}| \ll 1$ expansion of eq. (3.3) with the GW potential. The first term is referred to as the harmonic oscillator approximation (HO), where the jet quenching parameter is

$$\hat{q}(t) = \hat{q}_0(t)\ln\frac{Q^2}{\mu_*^2}\,, \tag{3.43}$$

where $\mu_*^2 = \frac{\mu^2}{4}\mathrm{e}^{-1+2\gamma_E}$ for the GW potential. The logarithm in $\hat{q}$ comes from the fact that the typical exchanged momentum $\langle\boldsymbol{k}^2\rangle = L\int_{\boldsymbol{q}}\boldsymbol{q}^2\sigma(\boldsymbol{q})$ is divergent and thus it has to be regulated resulting in the leading logarithmic form in eq. (3.43) (see in ref. [42]). As long as $Q^2/\mu_*^2 \gg 1/(Q^2\boldsymbol{x}^2)$, the HO term dominates over δv, and the latter can be treated as a perturbation. This provides a big advantage, since the multiple scattering in the HO approximation can be resummed analytically.

The separation scale Q^2 has to be fixed in a meaningful way to not to interfere with the expansion [55]. A natural choice that achieves this is evaluating Q^2 at the typical transverse momentum of the emission $\boldsymbol{k}^2 \sim \hat{q}t_{\rm f}$, that yields

$$Q_r^2(\omega) = \sqrt{\omega\hat{q}(\omega)}\,, \tag{3.44}$$

which constitutes an implicit equation for $Q_r^2(\omega)$, and for $\hat{q}(\omega) \equiv \hat{q}(Q_r(\omega))$, see eq. (3.43).[12]

[12]Equation (3.44) has a solution only if $\omega > 2e\frac{\mu_*^4}{\hat{q}_0} \simeq 0.925\omega_{\rm BH}$. When this is satisfied, and $L > \lambda$, then $Q_r^2 > \mu_*^2$ by default. This is the necessary condition for the convergence of the IOE. The IOE therefore breaks down for $\omega \lesssim \omega_{\rm BH}$.

JHEP02(2023)156

The IOE corresponds to expanding the full medium solution for $\mathcal{K}(\boldsymbol{x};\boldsymbol{y})$ around the harmonic oscillator, in contrast to the conventional opacity expansion where one expands around the vacuum solution. It can be cast in the iterative equation

$$\begin{aligned}\mathcal{K}(\boldsymbol{x},t_2;\boldsymbol{y},t_1) &= \mathcal{K}_{\rm HO}(\boldsymbol{x},t_2;\boldsymbol{y},t_1)\\ &\quad - \int_{t_1}^{t_2}\mathrm{d}s\int_{\boldsymbol{z}}\mathcal{K}_{\rm HO}(\boldsymbol{x},t_2;\boldsymbol{z},s)\delta v(\boldsymbol{z},s)\mathcal{K}(\boldsymbol{z},s;\boldsymbol{y},t_1)\,. \end{aligned}\tag{3.45}$$

Here, $\mathcal{K}_{\rm HO}(\boldsymbol{x};\boldsymbol{y})$ is itself the solution to an iterative equation, namely

$$\begin{aligned}\mathcal{K}_{\rm HO}(\boldsymbol{x},t_2;\boldsymbol{y},t_1) &= \mathcal{K}_0(\boldsymbol{x}-\boldsymbol{y},t_2-t_1)\\ &\quad - \int_{t_1}^{t_2}\mathrm{d}s\int_{\boldsymbol{z}}\mathcal{K}_0(\boldsymbol{x}-\boldsymbol{z},t_2-s)v_{\rm HO}(\boldsymbol{z},s)\mathcal{K}_{\rm HO}(\boldsymbol{z},s;\boldsymbol{y},t_1)\,. \end{aligned}\tag{3.46}$$

The formal solution for $\mathcal{K}_{\rm HO}(\boldsymbol{x},t_2;\boldsymbol{y},t_1)$ can also be cast as a path integral, namely

$$\mathcal{K}_{\rm HO}(\boldsymbol{x},t_2;\boldsymbol{y},t_1) = \int_{\boldsymbol{r}(t_1)=\boldsymbol{y}}^{\boldsymbol{r}(t_2)=\boldsymbol{x}}\mathcal{D}\boldsymbol{r}\,\mathrm{e}^{\,i\int_{t_1}^{t_2}\mathrm{d}s\left[\frac{\omega}{2}\dot{\boldsymbol{r}}^2+iv_{\rm HO}(\boldsymbol{r},s)\right]}\,,\tag{3.47}$$

which has a well-known analytical solution in a static medium,

$$\mathcal{K}_{\rm HO}\left(\boldsymbol{x},t_2;\boldsymbol{y},t_1\right) = \frac{\omega\Omega}{2\pi i\,\sin(\Omega\Delta t)}\mathrm{e}^{\frac{i\omega\Omega}{2\sin(\Omega\Delta t)}\left[\cos(\Omega\Delta t)\,(\boldsymbol{x}^2+\boldsymbol{y}^2)-2\boldsymbol{x}\cdot\boldsymbol{y}\right]}\,,\tag{3.48}$$

where $\Delta t\equiv t_2-t_1$ and $\Omega=\frac{1-i}{2}\sqrt{\hat{q}(\omega)/\omega}$ is the characteristic oscillator frequency.

Inserting this expansion into the equation for the medium-induced spectrum eq. (3.1) separates it into two parts,

$$\omega\frac{\mathrm{d}I^{\rm HO}}{\mathrm{d}\omega} = \frac{2\alpha_s C_R}{\omega^2}\mathrm{Re}\int_0^\infty\mathrm{d}t_2\int_0^{t_2}\mathrm{d}t_1\,\partial_{\boldsymbol{x}}\cdot\partial_{\boldsymbol{y}}\left[\mathcal{K}_{\rm HO}(\boldsymbol{x},t_2;\boldsymbol{y},t_1)-\mathcal{K}_0(\boldsymbol{x},t_2;\boldsymbol{y},t_1)\right]_{\boldsymbol{x}=\boldsymbol{y}=0}\,,\tag{3.49}$$

$$\begin{aligned}\omega\frac{\mathrm{d}I^{\rm IOE}}{\mathrm{d}\omega} &= -\frac{2\alpha_s C_R}{\omega^2}\mathrm{Re}\int_0^\infty\mathrm{d}t_2\int_0^{t_2}\mathrm{d}s\int_0^s\mathrm{d}t_1\int_{\boldsymbol{z}}\\ &\quad\times\partial_{\boldsymbol{x}}\cdot\partial_{\boldsymbol{y}}[\mathcal{K}_{\rm HO}(\boldsymbol{x},t_2;\boldsymbol{z},s)\delta v(\boldsymbol{z},s)\mathcal{K}(\boldsymbol{z},s;\boldsymbol{y},t_1)]_{\boldsymbol{x}=\boldsymbol{y}=0}\,. \end{aligned}\tag{3.50}$$

The first term gives rise to the well-known HO spectrum, while the second constitutes an expansion in hard splittings around the harmonic oscillator. The IOE spectrum can be simplified further, giving

$$\omega\frac{\mathrm{d}I^{\rm IOE}}{\mathrm{d}\omega} = \frac{2\bar{\alpha}}{\omega}\mathrm{Re}\,i\int_0^L\mathrm{d}t_2\int_0^{t_2}\mathrm{d}t_1\int_{\boldsymbol{x}}\mathrm{e}^{-i\frac{\omega\Omega}{2}\tan\left(\Omega(L-t_2)\right)\boldsymbol{x}^2}\delta v(\boldsymbol{x})\frac{\boldsymbol{x}}{\boldsymbol{x}^2}\cdot\partial_{\boldsymbol{y}}\mathcal{K}(\boldsymbol{x},t_2;\boldsymbol{y},t_1)|_{\boldsymbol{y}=0}\,,\tag{3.51}$$

where $\mathcal{K}(\boldsymbol{x};\boldsymbol{y})$ should be iterated using eq. (3.45) in order to generate higher orders of the expansion. A general formula for the improved opacity expansion at arbitrary order can also be derived, which was done in section A.3, see also in ref. [55].

JHEP02(2023)156

Harmonic oscillator (HO). The harmonic oscillator approximation resums all coherent soft scatterings during the formation of the emission. The relevant scale is

$$\omega_c \equiv \frac{1}{2}\hat{q}(\omega_c)L^2\,, \tag{3.52}$$

where the scale in the jet quenching parameter is set to ω_c. This scale was already identified in section 2. The HO approximation is expected to be valid for $\omega_{\rm BH} \ll \omega \ll \omega_c$.

We derive here the familiar harmonic oscillator spectrum in a new way. The vacuum contribution can easily be subtracted by inserting eq. (3.45) into eq. (3.1), which gives

$$\begin{aligned}\omega\frac{\mathrm{d}I^{\rm HO}}{\mathrm{d}\omega} = &-\frac{2\alpha_s C_R}{\omega^2}\mathrm{Re}\int_0^\infty \mathrm{d}t_2\int_0^{t_2}\mathrm{d}s\int_0^s \mathrm{d}t_1\int_{\boldsymbol{z}}\\ &\times\partial_{\boldsymbol{x}}\cdot\partial_{\boldsymbol{y}}\left[\mathcal{K}_0(\boldsymbol{x},t_2;\boldsymbol{z},s)v_{\rm HO}(\boldsymbol{z},s)\mathcal{K}_{\rm HO}(\boldsymbol{z},s;\boldsymbol{y},t_1)\right]_{\boldsymbol{x}=\boldsymbol{y}=0}\,.\end{aligned} \tag{3.53}$$

Using the fact that $\int_s^\infty \mathrm{d}t_2\,\partial_{\boldsymbol{x}}\mathcal{K}_0(\boldsymbol{x},t_2;\boldsymbol{z},s)|_{\boldsymbol{x}=0} = -i\frac{\omega}{\pi}\frac{\boldsymbol{z}}{\boldsymbol{z}^2}$, the spectrum becomes

$$\omega\frac{\mathrm{d}I^{\rm HO}}{\mathrm{d}\omega} = \frac{\bar{\alpha}\hat{q}(\omega)}{2\omega}\mathrm{Re}\,i\int_0^L \mathrm{d}t_2\int_0^{t_2}\mathrm{d}t_1\int_{\boldsymbol{z}}\boldsymbol{z}\cdot\partial_{\boldsymbol{y}}\mathcal{K}_{\rm HO}(\boldsymbol{z},t_2;\boldsymbol{y},t_1)|_{\boldsymbol{y}=0}\,. \tag{3.54}$$

This can be further simplified by using that

$$\begin{aligned}\int_{\boldsymbol{z}}\boldsymbol{z}\cdot\boldsymbol{\partial}_{\boldsymbol{y}}\mathcal{K}_{\rm HO}(\boldsymbol{z},t_2;\boldsymbol{y},t_1)|_{\boldsymbol{y}=0} &= -\frac{(\omega\Omega)^2}{2\pi\sin^2(\Omega(t_2-t_1))}\int_{\boldsymbol{z}}\boldsymbol{z}^2\mathrm{e}^{i\frac{\omega\Omega}{2}\boldsymbol{z}^2\cot(\Omega(t_2-t_1))}\,,\\ &= \frac{2}{\cos^2(\Omega(t_2-t_1))}\,.\end{aligned} \tag{3.55}$$

The time integration can be now be dealt with straightforwardly, yielding

$$\int_0^L \mathrm{d}t_2\int_0^{t_2}\mathrm{d}t_1\,\frac{1}{\cos^2(\Omega(t_2-t_1))} = -\frac{\ln\cos\Omega L}{\Omega^2}\,, \tag{3.56}$$

and thus the spectrum becomes

$$\omega\frac{\mathrm{d}I^{\rm HO}}{\mathrm{d}\omega} = 2\bar{\alpha}\,\ln|\cos\Omega L|\,, \tag{3.57}$$

which is the familiar BDMPS-Z spectrum [32, 69]. The limits of this are

$$\omega\frac{\mathrm{d}I^{\rm HO}}{\mathrm{d}\omega} \simeq \begin{cases}\bar{\alpha}\sqrt{\frac{2\omega_c}{\omega}}, & \text{for } \omega\ll\omega_c\,,\\ \frac{\bar{\alpha}}{6}\left(\frac{\omega_c}{\omega}\right)^2, & \text{for } \omega\gg\omega_c\,.\end{cases} \tag{3.58}$$

The soft limit agrees with the discussion in section 2, while the hard limit is subleading compared to the OE $N=1$ in eq. (3.16). Defining $\hat{q}$ with a logarithm extends the region of validity, which was also found in ref. [42], leading to the curved $\omega_c(t)$ line in figure 3.

Next-to-harmonic oscillator (NHO). Using the definition in eq. (3.51) and the results of section A.3, the first order of the improved opacity expansion can be written as

$$\omega\frac{\mathrm{d}I^{\rm NHO}}{\mathrm{d}\omega} = \frac{2\bar{\alpha}}{\pi}\frac{L}{\lambda}\frac{\bar{\omega}_c}{\omega}\mathrm{Re}\int_0^1\mathrm{d}s\int_{\boldsymbol{u}}\frac{1}{2}\ln\left(\frac{\omega}{\bar{\omega}_c}\frac{\mu^2}{2Q^2}\frac{1}{\boldsymbol{u}^2}\right)\mathrm{e}^{\frac{i}{2}f(s)\boldsymbol{u}^2}\,, \tag{3.59}$$

JHEP02(2023)156

where we have defined the function $f(s) = \sigma\sqrt{\omega_c/\omega}[\cot(\sigma s\sqrt{\omega_c/\omega}) - \tan(\sigma(1-s)\sqrt{\omega_c/\omega})]$, and $\sigma = \frac{1-i}{\sqrt{2}}$. After doing the $\boldsymbol{u}$ integral this becomes

$$\omega\frac{\mathrm{d}I^{\mathrm{NHO}}}{\mathrm{d}\omega} = 2\bar{\alpha}\frac{L}{\lambda}\frac{\bar{\omega}_c}{\omega}\mathrm{Re}\, i\int_0^1 \mathrm{d}s\,\frac{1}{f(s)}\left[1-\gamma_E+\ln\left(-i\frac{\omega}{\bar{\omega}_c}\frac{\mu_*^2}{Q^2}f(s)\right)\right]. \tag{3.60}$$

The limits of this expression can readily be extracted. In the soft limit, $\omega \ll \omega_c$, we have $f(s) \to 2i\sigma\sqrt{\omega_c/\omega}$ while in the hard limit, $\omega \gg \omega_c$, it becomes $f(s) \to 1/s$. These simplifications make it possible to do the last time integration. Hence, the extracted limiting behavior is,

$$\omega\frac{\mathrm{d}I^{\mathrm{NHO}}}{\mathrm{d}\omega} \simeq \begin{cases} \bar{\alpha}\sqrt{\frac{2\omega_c}{\omega}}\,\frac{1}{2\ln Q^2/\mu_*^2}\left(\frac{\pi}{4}+\gamma_E+\ln\left(\frac{\sqrt{\hat{q}\omega}}{\sqrt{2}Q^2}\right)\right), & \text{for } \omega \ll \omega_c\,, \\ \frac{\pi\bar{\alpha}}{2}\frac{L}{\lambda}\frac{\bar{\omega}_c}{\omega}, & \text{for } \omega \gg \omega_c\,. \end{cases} \tag{3.61}$$

In the soft limit $\omega \ll \omega_c$, N^{n}HO terms will take the form of the HO by using Q_r, and thus

$$\omega\frac{\mathrm{d}I}{\mathrm{d}\omega} = \bar{\alpha}\sqrt{\frac{\hat{q}(\omega)L^2}{\omega}}\left(1+\frac{1}{2}\frac{a_0}{\ln Q_r^2/\mu_*^2}+\mathcal{O}\left(\frac{1}{\ln Q_r^2/\mu_*^2}\right)^2\right), \tag{3.62}$$

where we added the HO term, and used eq. (3.44). The choice of $Q = Q_r(\omega)$ is effective, when the medium is big enough $L \gg \lambda$. It is clear that the expansion parameter of the IOE is $\ln^{-1}(Q_r^2/\mu_*^2) \ll 1$ in the soft limit. Therefore, N^{n}HO terms can be absorbed into an effective jet transport parameter,

$$\hat{q}_{\mathrm{eff}}(Q^2) = \hat{q}_0\ln\left(\frac{Q_r^2}{\mu_*^2}\right)\left[1+\frac{a_0}{\ln Q_r^2/\mu_*^2}+\frac{a_1}{\ln^2 Q_r^2/\mu_*^2}+\ldots\right]. \tag{3.63}$$

The coefficients $a_0 = 1.016$ and $a_1 = 0.316$ of the expansion and higher-order terms up to N^{2}HO were found in ref. [55].

In the hard limit of eq. (3.61), one can see that the IOE reproduces the hard limit of $N = 1$ in the OE from eq. (3.16). Furthermore, it is bigger than the HO contribution in eq. (3.57) and thus NHO dominates for $\omega \gg \omega_c$.

As mentioned above, the HO is meaningful if $L > \lambda$ and $\omega > \omega_{\mathrm{BH}}$. Furthermore, the IOE is expected to converge if $v_{\mathrm{HO}} > \delta v$ or equivalently $\omega > \omega_{\mathrm{BH}}$. However, as the hard limits of the first order of the IOE and the OE are equal, and the OE is valid down to $\bar{\omega}_c$, it is reasonable to assume that also the IOE is valid down to $\bar{\omega}_c$. Therefore, the region of validity will be extended to all L and $\omega > \min(\omega_{\mathrm{BH}}, \bar{\omega}_c)$, as shown in the left of figure 6 in red.

The spectra obtained with the IOE from eq. (3.57) and eq. (3.61) are shown in the right panel of figure 6 for different propagation lengths. At early times $t < \lambda$, the HO approximation is highly suppressed, due to the absence of multiple scattering. However, the contribution from the NHO makes the total agree with the hard limit of $N = 1$ OE and $N_r = 1$ ROE (cf. figures 4 and 5). The deviation close to $\omega \approx \bar{\omega}_c$ (gray bullet) arises since Q_r was chosen to reproduce the HO spectrum which is strictly valid for $L \gg \lambda$. For later times $t > \lambda$, both the HO and NHO will give sizable contributions, where the HO dominates if $\omega < \omega_c$ (red bullets) and NHO dominates if $\omega > \omega_c$. The HO approximation breaks down if $\omega < \omega_{\mathrm{BH}}$ (gray bullets). The dashed line is the full numerical solution from refs. [51, 68] and the IOE well captures it in its region of validity.

JHEP02(2023)156

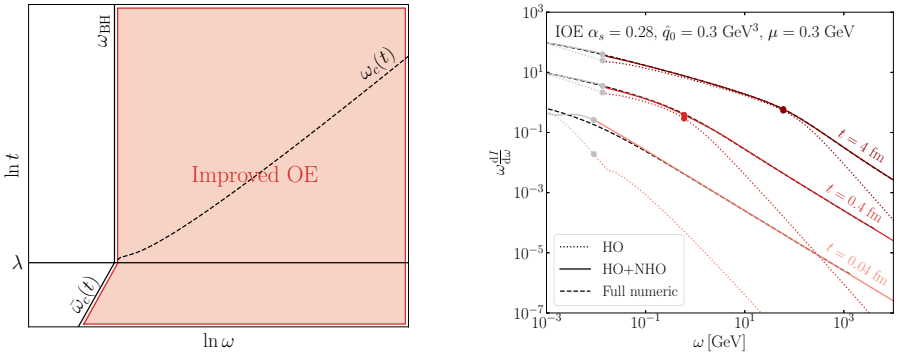


Figure 6. *Left:* the sketch of region of validity (and convergence) of the improved opacity expansion for different propagation length t and emission energy ω. *Right:* the induced emission spectrum for gluons in the improved opacity expansion. The gray part of the curves denotes regions where the expansion is not valid. Using the parameters presented, $\lambda = 0.06\,\text{fm}$. The full numeric solution is also presented with dashed lines.

3.5 Summary of the regimes and the induced emission spectrum

In this section, we have presented three distinct perturbative expansions (OE, ROE, and IOE) that provide different ways of calculating the induced emission spectrum $\omega \frac{dI}{d\omega}$ in their respective regimes of convergence. The OE and IOE are expansion schemes that were first developed in previous works, while the ROE is rigorously derived in this section for the first time. Their regions of validity are sketched previously in figures 4–6, and at least one of the expansions is valid at every point in the phase space (ω, t). Here ω is the emitted energy and t is the propagated length (L is the maximal length of the medium and E is the energy of the emitting particle). As a consequence, our description of the spectrum is complete in the full phase space, as one can always use one of the expansions to reach an approximation of the true spectrum, and one can reach better accuracy by including higher orders. Note that the expansions are overlapping: for $L < \lambda$ both OE and ROE are valid, and for $\omega > \omega_c$ both IOE the OE can be used.

The results presented obtained so far within the unified resummation framework are valid in both dilute and dense regimes and can be systematically improved to arbitrary high order in the expansions. For practical purposes, however, a handy and efficient interpolation formula such suffice to capture the relevant features to high precision. This would be very useful for other applications, such as resumming multiple emissions in sections 4 and 5. To describe the spectrum in the whole phase space, we use (to first order)

$$\frac{dI^{\rm Full}}{d\omega} = \begin{cases} \frac{dI^{\rm ROE}}{d\omega}\,, & \omega < \min(\omega_{\rm BH}, \bar{\omega}_c(t))\,, \\ \frac{dI^{\rm IOE}}{d\omega}\,, & \text{otherwise}\,. \end{cases} \tag{3.64}$$

Based on figures 4–6 (and the all order expansion formulas), the first-order terms already capture the most important effects. We stress that this is arguable the most straightforward

JHEP02(2023)156

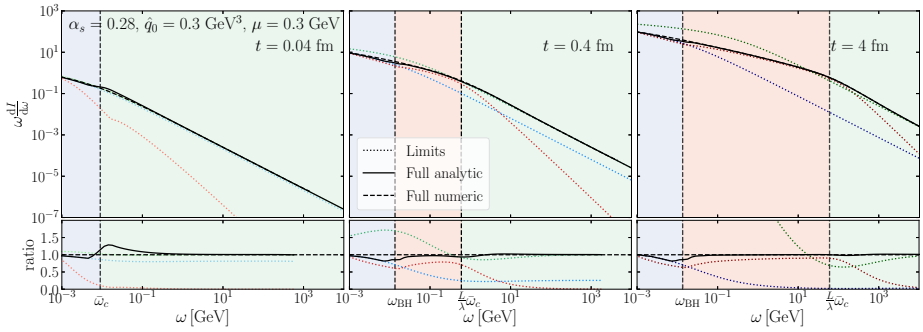


Figure 7. Summary of the induced emission spectrum for gluons, combined from the three expansion schemes at different propagation length (different panels). The black lines are our final forms from eq. (3.64) that uses $N_r = 1$ and HO+NHO. The shaded areas denote the leading scattering process and the corresponding dotted lines are the limiting $N = 1$, HO and $N_r = 1$ contributions. The dashed lines are the numerical solution of eq. (3.1) from ref. [51].

interpolation scheme. However, it turns out that it gives a good description in almost the whole phase space, deviating maximally 30% from the exact numerical results around the Bethe-Heitler energy in the dilute regime, see figure 7 (left, lower panel).[13]

To summarize, the limiting behaviour of the spectrum in different regions of the phase space is

$$\omega \frac{\mathrm{d}I}{\mathrm{d}\omega}\bigg|_{L\ll\lambda} = \begin{cases} 2\bar{\alpha}\frac{L}{\lambda}\left(\ln\frac{\bar{\omega}_c}{\omega} - 1 + \gamma_E\right), & \text{for } \omega \ll \bar{\omega}_c\,, \\ \frac{\pi}{2}\bar{\alpha}\frac{L}{\lambda}\frac{\bar{\omega}_c}{\omega}, & \text{for } \bar{\omega}_c \ll \omega\,, \end{cases} \tag{3.65}$$

for $L \ll \lambda$, and

$$\omega \frac{\mathrm{d}I}{\mathrm{d}\omega}\bigg|_{L\gg\lambda} = \begin{cases} 2\bar{\alpha}\frac{L}{\lambda}\ln\left(\frac{\omega_{\mathrm{BH}}}{\omega}\right), & \text{for } \omega \ll \omega_{\mathrm{BH}}\,, \\ \bar{\alpha}\sqrt{\frac{2\omega_c}{\omega}}, & \text{for } \omega_{\mathrm{BH}} \ll \omega \ll \omega_c\,, \\ \frac{\pi}{2}\bar{\alpha}\frac{L}{\lambda}\frac{\bar{\omega}_c}{\omega}, & \text{for } \omega_c \ll \omega\,, \end{cases} \tag{3.66}$$

for $L \gg \lambda$. This agrees with the formulas from the heuristic discussion in section 2. In figure 7, we evaluated eq. (3.64) (black curve) up to $N_r = 1$ and HO+NHO for different times. The dotted curves are the limits of $N_r = 1$, HO and $N = 1$, shown in blue, red and green respectively. The regions are shaded with the same colors as in figure 3, visualizing the regions of the distinct scattering processes. At the transition point $\min(\omega_{\mathrm{BH}}, \bar{\omega}_c(t))$, the spectrum is not completely smooth, and the difference is expected to vanish as one goes to higher orders in the perturbative expansion. We defined a switching function that makes the transition smoother, which is described in appendix D.

In figure 7, we have also plotted the full numerical spectra from refs. [51, 68] with dashed lines. The excellent agreement with our curves corroborates the validity of our formula in

[13]The interpolation can be further inspected in figure 11, where we plot the spectrum on a semilog scale. The exact details of matching the ROE regime with the IOE involves a smooth interpolation function, that avoids blowing up the logarithmic dependence of $\hat{q}$. This is described in detail in appendix D.

JHEP02(2023)156

eq. (3.64), where we used $N_r = 1$ and HO+NHO in the plot. In the ratio panel, one can see, how the different regions are adding up to give an overall very accurate description of the full spectrum. Moreover, as we saw in figures 4–6, by including higher orders e.g. $N_r = 2$, the description becomes more accurate, smoothing the transition around ω_{BH}. We leave the study of these higher-order corrections and the corresponding uncertainties for future studies.

4 Resumming multiple emissions in the medium

Section 3 presents a theoretical framework consisting of different perturbative expansions (namely the opacity expansion (OE), the resummed OE (ROE), and the improved OE (IOE)) to describe the full phase space (ω, t) of the medium-induced gluon emission spectrum. The emitted energy ω is limited by the energy of the emitter $\omega \ll E$, and the propagation time (length) t is in turn limited by the medium length $t < L$. Our effective formalism accounts for arbitrarily many scatterings that can be arbitrarily hard or soft. Finally, eq. (3.64) describes the emission spectrum up to arbitrary precision and recovers the full solution of eq. (3.1) that has only been achieved numerically before [48, 49, 51].

It is now time to explore what consequences the full induced spectrum instills on a parton propagating through the medium. In this section, we will present analytical solutions of the evolution equation of the medium-induced cascade. We present the inclusive gluon energy distribution for different medium lengths, and we will focus on gluons for transparency. The numerical solutions are presented in section 5. Vacuum emissions belong outside of the scope of the current work.

4.1 The necessity of multiple emissions

Multiple emissions have to be taken into account whenever the multiplicity of gluons is large. We define the multiplicity of gluons above the energy ω in terms of the spectrum $\mathrm{d}I/\mathrm{d}\omega$, as

$$N(\omega) = \int_{\omega}^{\infty} \mathrm{d}\omega' \frac{\mathrm{d}I}{\mathrm{d}\omega'} \,. \tag{4.1}$$

The upper limit of the integral is taken to infinity because we will currently assume that the energy of the emitter E is much larger than the largest available medium energy scale. Also, for our current purposes, it suffices to consider the leading behavior of the spectrum in the various scattering regimes presented in figure 3.

At low opacity $L \ll \lambda$ and starting from $\omega < \omega_{\text{BH}}$, we find

$$N(\omega) \simeq \bar{\alpha} \frac{L}{\lambda} \ln^2 \frac{\bar{\omega}_c}{\omega} + \frac{\pi \bar{\alpha}}{2} \frac{L}{\lambda} \,, \tag{4.2}$$

where $\bar{\omega}_c = \frac{1}{2}\mu^2 L$ and we have only kept the leading terms. The maximal multiplicity in the hard regime, see the second term, is always small for perturbative splittings with $\bar{\alpha} \ll 1$. Furthermore, the multiplicity in the soft (Bethe-Heitler) regime, given by the first term, becomes large only at very small energies, i.e. $\omega < \mathrm{e}^{\sqrt{1/(\bar{\alpha}L/\lambda)}} \bar{\omega}_c$. We can therefore safely neglect multiple emissions at low opacities.

JHEP02(2023)156

For dense media, $L \gg \lambda$, we discuss the three pertinent cases. As usual, we will denote $\omega_c = \frac{1}{2}\hat{q}L^2$. Then, for $\omega \gg \omega_c$ (rare hard scattering regime), we find

$$N(\omega) \simeq \bar{\alpha}\frac{L}{\lambda}\frac{\bar{\omega}_c}{\omega}\,, \tag{4.3}$$

where we again neglected subleading terms. Hence hard emissions, described by the OE expansion, can safely be considered to be rare.

Next, for $\omega_{\rm BH} \ll \omega \ll \omega_c$ we get

$$N(\omega) \simeq 2^{\frac{3}{2}}\bar{\alpha}\sqrt{\frac{\omega_c}{\omega}}\,, \tag{4.4}$$

where we introduced the leading behavior of the spectrum and neglected the multiplicity from hard emissions, following the discussion above. The multiplicity becomes large $N(\omega) \gg 1$ at energies $\omega \ll \bar{\alpha}^2\omega_c$ and thus multiple emission becomes dominant. For large enough medium length $L \gg \lambda$, there is a significant phase space allowing for multiple emission ($\bar{\alpha}_s^2\omega_c(t) > \omega_{\rm BH}$), resulting in the power enhancement.

Finally, for soft gluon energies $\omega < \omega_{\rm BH} \ll \omega_c$, the multiplicity is

$$N(\omega) \simeq \bar{\alpha}\frac{L}{\lambda}\ln^2\frac{\omega_{\rm BH}}{\omega} + 2^{\frac{3}{2}}\bar{\alpha}\sqrt{\frac{\omega_c}{\omega_{\rm BH}}}\,, \tag{4.5}$$

where, again, only the leading terms from each regime were kept. The second term in eq. (4.5) scales as $\sim (L/\lambda)\sqrt{\hat{q}/\hat{q}_0}$. Based on the discussion above, this term is already large and the multiplicity continues to grow only logarithmically for small ω, and therefore multiple emissions are going to happen.

4.2 Resummation of multiple emissions

Considering multiple emissions in a medium poses a tremendous theoretical challenge. The situation is quite analogous to the description of multiple gluon emissions in vacuum. Similar to QCD jets, the main challenge when considering medium effects lies in dealing with intricate interference effects between subsequent emissions, see e.g. refs. [70–72]. However, when considering multiple soft emissions, that occur quasi-instantaneously, these effects can safely be neglected [58, 59].[14]

In order to clarify the framework that we work in, let us briefly recall the main arguments for neglecting interference effects for a set of multiple induced emissions. For the time being, we stick to emissions in the HO region which dominate the multiplicity. The typical time it takes an emission to form, often referred to as a *formation time* (or in some works branching time), of a soft gluon is $t_{\rm f} \sim \sqrt{\omega/\hat{q}}$. This time is much smaller than the extent of the medium $t_{\rm f} \ll L$ as long as $\omega \ll \omega_c$.

Another relevant quantity is the time between two subsequent emissions. This is related to the no emission probability (or Sudakov factor). For a leading particle with energy E, the first emission is produced at time $t_{\rm f1}$, with energy ω_1. A second, strongly ordered

[14] We can also extend this logic for the semi-hard emissions which are included in our formalism, since they are rare occurrences and therefore the resummation has no effect, see section 4.3.

JHEP02(2023)156

emission ($\omega_2 \ll \omega_1$) of the original parton forms much quicker $t_{\rm f2} \ll t_{\rm f1}$. The time between the two emissions $t_{\rm split}$ can be estimated with $\int_{t_{\rm f1}}^{t_{\rm split}} {\rm d}t \int_{\omega_2}^{\omega_1} {\rm d}\omega \, \frac{{\rm d}I}{{\rm d}\omega {\rm d}t} \sim 1$, that is basically the probability of not having emissions between $\omega_2 < \omega < \omega_1$, resulting in

$$t_{\rm split} \sim t_{\rm f1} + \frac{1}{\bar{\alpha}} t_{\rm f2} \,. \tag{4.6}$$

Hence, our rough estimate implies that typically $t_{\rm split} \gg t_{\rm f2}$ (for $\bar{\alpha} \ll 1$). Therefore, the formation of emissions is short compared to the time that separates emissions and thus emissions form independently. This motivates the resummation of multiple independent emissions in terms of a rate equation.

Similar analysis can be done for hard emissions $\omega \gg \omega_c$, for which $t_{\rm split}$ is very long due to the unlikeliness of hard scatterings, and therefore the emissions are formed independently.

For soft emissions $\omega \ll \omega_{\rm BH}$, $t_{\rm split} \sim t_{\rm f1} + \frac{1}{\bar{\alpha}} \ln^{-2} \frac{\lambda}{t_{\rm f2}}$ and thus emissions form independently. Close to the boundary in cases where $t_{\rm split} \approx t_{\rm f1} + t_{\rm f2}$, a more complicated structure appears in terms of resummation, as emissions might overlap. In this case, interference effects between the two emissions have to be included. A similar thing happens in vacuum for wide angle soft emissions, which result in angular ordering and in non-global effects for which the resummation has been understood just recently [73]. We will use the rate equation to account for emissions with any ω. However, it will not necessarily account correctly for interference among them and further study is needed in the future.

One question still remains open, namely the choice of the time scale used in the rate equation. In case of two emissions, the second emission experiences a shorter medium, of the scale $\sim L - t_{\rm f1} - t_{\rm split}$. We know, however, that in the soft limit $\omega \ll \omega_c$, the formation time is $t_{\rm f} \ll L$, and therefore, the length degradation should not matter for a large medium [58, 59]. For small media, or for emissions with comparable formation times, these corrections can become significant. It is an unresolved question how to incorporate these corrections into a rate equation see e.g. ref. [71].[15] However, as argued above, the corrections to the rate coming from finite-size effects can be treated in a perturbative fashion. While these issues merit further studies, perhaps within a Monte Carlo approach, we consider them to go beyond our present scope and we assume that all emissions experience the same length L. This matches the approximation in most of the current energy-loss models.

In this section, we will focus on the single-inclusive energy distribution of partons carrying energy xE after traveling length t in the medium, where E is the initial energy. It is defined as

$$D(x,t) \equiv x \frac{{\rm d}N}{{\rm d}x} \,. \tag{4.7}$$

The formalism can easily be extended to account for parton flavors, see e.g. [75], but for now, we restrict our attention to a pure gluon cascade.

In section 3 we focused on emissions of soft gluons with energies $\omega \ll E$. Now we will consider generic splitting processes where a parton with flavor index $a = q, g$ and initial energy E shares its energy with two daughter partons, with energies zE and $(1-z)E$ and flavor indices b and c, respectively, for $0 < z < 1$. The spectrum of such splittings ${\rm d}I_{ba}/{\rm d}z$

[15] See also section 4 in ref. [74], where modifications of the rate due to finite formation time were studied.

JHEP02(2023)156

is given by eq. (C.1). The general features of figure 3 remain the same with the substitution $\omega \to z(1-z)E$.[16] For further details, see the discussion in appendix C.

As was shown in refs. [43, 59], for sufficiently soft emissions, with formation times much smaller than the medium length, interference effects are suppressed and one can consider multiple emissions as occurring independently. The evolution equation for the energy distribution, that accounts for an arbitrary number of induced emissions, is given by

$$\frac{\partial}{\partial t}D(x,t) = \int_x^1 \mathrm{d}z\, \mathcal{K}\left(z, \frac{x}{z}E, t\right) D\left(\frac{x}{z}, t\right) - \int_0^1 \mathrm{d}z\, z\, \mathcal{K}\left(z, xE, t\right) D(x,t)\,. \tag{4.8}$$

The initial condition is a single gluon carrying energy E, hence $D(x,0) = \delta(1-x)$. The splitting kernel $\mathcal{K}(z,E,t)$ is the rate of emissions off a particle with energy E,

$$\mathcal{K}(z,E,t) = 2\,\frac{\mathrm{d}I_{gg}}{\mathrm{d}z\mathrm{d}t}\bigg|_E\,, \tag{4.9}$$

and the rate with full z-dependence for the $g \to gg$ splitting can be found in eq. (C.1).[17] The first term in eq. (4.8) is a real emission describing an emitted gluon with energy fraction x (gain term), while the second is a virtual emission that does not change the energy of the emitter (loss term). Both terms contribute to cancelling out the apparent divergence at $z \to 1$. The evolution equation conserves the total energy contained in the spectrum,

$$\int_0^1 \mathrm{d}x\, D(x,t) = 1\,, \tag{4.10}$$

which can be confirmed directly from (4.8).

The leading parametric behavior of the splitting kernels can be derived by taking appropriate limits and is presented in section 3, cf. eqs. (3.16), (3.34), and (3.58). This results in,

$$\mathcal{K}(z,E,t)|_{t\ll\lambda} = \begin{cases} \frac{2\bar{\alpha}}{z(1-z)}\frac{1}{\lambda}\ln\left(\frac{\bar{\omega}_c(t)}{z(1-z)E}\right) & \text{for } z(1-z)E \ll \bar{\omega}_c(t)\,, \\ \frac{\bar{\alpha}\pi}{2}\frac{\hat{q}_0 t}{[z(1-z)]^2 E} & \text{for } \bar{\omega}_c(t) \ll z(1-z)E\,, \end{cases} \tag{4.11}$$

for $t \ll \lambda$, and

$$\mathcal{K}(z,E,t)|_{t\gg\lambda} = \begin{cases} \frac{2\bar{\alpha}}{z(1-z)}\frac{1}{\lambda}\ln\left(\frac{\omega_{\mathrm{BH}}}{z(1-z)E}\right) & \text{for } z(1-z)E \ll \omega_{\mathrm{BH}}\,, \\ \bar{\alpha}\sqrt{\frac{\hat{q}}{[z(1-z)]^3 E}} & \text{for } \omega_{\mathrm{BH}} \ll z(1-z)E \ll \omega_c(t)\,, \\ \frac{\bar{\alpha}\pi}{2}\frac{\hat{q}_0 t}{[z(1-z)]^2 E} & \text{for } \omega_c(t) \ll z(1-z)E\,, \end{cases} \tag{4.12}$$

for $t \gg \lambda$, where $\bar{\omega}_c(t) = \frac{1}{2}\mu^2 t$ and $\omega_c = \frac{1}{2}\hat{q}t^2$. For the analytical estimates in this section we neglect the running of $\hat{q}$, but this will be included in the numerics presented in section 5.

[16] Hence the upper limit of ω in figure 3 should now be $E/4$.

[17] In the soft limit, the kernel is closely related to the spectrum discussed in the previous section or, more precisely, the rate $\mathrm{d}I/(\mathrm{d}\omega\mathrm{d}t)$, calculated in appendix B. Importantly, for the gluon splitting kernel, the divergences in $z \to 0$ and $z \to 1$ are folded together in the limit $\omega \to 0$, hence the additional symmetry factor in eq. (4.9).

JHEP02(2023)156

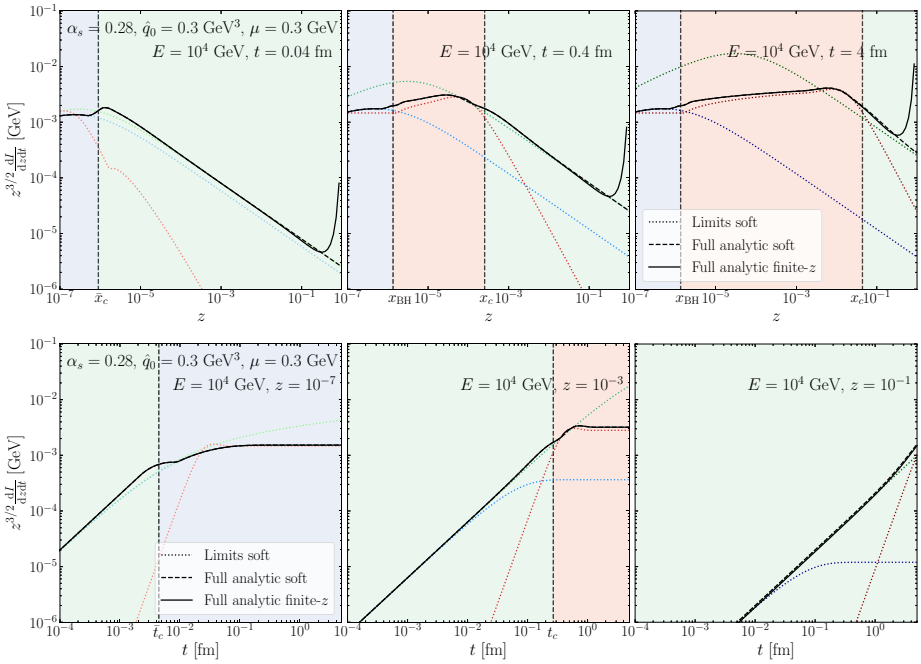


Figure 8. The medium-induced rate from eqs. (4.9), and (3.64) for gluons for different emitted energy and propagation time (black dashed lines). The black solid lines include finite-z corrections. The dotted lines are the $N_r = 1$, HO and $N = 1$ lines according to figure 7. The color shading corresponds to figure 3i.e. denoting the dominant scattering processes. The parameters correspond to $\lambda = 0.06\,\text{fm}$.

The rate $\mathrm{d}I/(\mathrm{d}z\mathrm{d}t)$ is plotted in figure 8, in a similar manner to the spectrum in figure 7, and rescaled by a factor $z^{3/2}$ to highlight the behavior at small z. The panels in the upper part show the z-dependence for three different t (early, mid-, and late times), and has a very similar structure to that of the spectrum. The color shading corresponds to figure 3, i.e. the dominant scattering processes. The dotted lines are the $N_r = 1$, HO and $N = 1$ lines from section 3 as in figure 7, where $\omega \to zE$ was used. Similarly, the dashed line is the full solution in the soft limit with $\omega \to zE$. The solid black lines are from eq. (5.3) and they contain finite-z corrections from appendix C. The finite-z corrections change the rate for hard emissions around $z \sim 1$. The panels in the lower part show the time dependence of the rate for a fixed emitted energy.

4.3 Analytic solutions of the evolution equation

The evolution equation (4.8) is readily solved by numerical evaluation, which will be discussed in section 5. Here we will discuss limiting cases where analytical solutions can be found. We can find such solutions at early times (considering only one emission) and at late times (considering many soft emissions). These are by now well-known limiting

JHEP02(2023)156

cases. Finally, we also consider the novel case of the evolution equation at intermediate times, where both rare, hard emissions and multiple, soft emissions can occur in sequence according to their respective allowed phase space of emissions, given in figure 3.

To simplify our discussion, in this section we will neglect the Bethe-Heitler regime. We will nevertheless include it in the full numerical solutions presented in section 5.

Early time evolution. At an early stage of medium propagation, the leading parton has little time to interact with the medium which also translates into a small probability of splitting. Given our previous discussion, the natural medium scale to compare with is the mean free path λ. Hence, at $t \lesssim \lambda$ we consider a single splitting, leading to

$$\begin{aligned} D(x,t) &\simeq \delta(1-x)\left[1-\int_0^t \mathrm{d}s \int_0^1 \mathrm{d}z\, z\mathcal{K}(z,xE,s)\right] + \int_0^t \mathrm{d}s \int_x^1 \mathrm{d}z\, \mathcal{K}\left(z,\frac{x}{z}E,s\right)\delta\left(1-\frac{x}{z}\right) \\ &= \int_0^t \mathrm{d}s\, x\mathcal{K}(x,E,s)\,, \end{aligned} \tag{4.13}$$

where we dropped the term proportional to $\delta(1-x)$, which is only important for energy conservation, cf. eq. (4.10). Using the results in eq. (4.11), we find that

$$D(x,t) \simeq \begin{cases} 2\bar{\alpha}\frac{t}{\lambda}\frac{1}{1-x}\ln\left(\frac{\bar{\omega}_c(t)}{x(1-x)E}\right) & \text{for } x \ll \bar{x}_c\,, \\ \frac{\pi\bar{\alpha}}{4}\frac{\hat{q}_0}{E}\frac{t^2}{x(1-x)^2} & \text{for } \bar{x}_c \ll x \ll 1-\bar{x}_c\,, \end{cases} \tag{4.14}$$

where we have defined $\bar{x}_c = \bar{\omega}_c(t)/E$. For $x \ll \bar{x}_c$, corresponding to the single soft or Bethe-Heitler scattering regime, a characteristic $D \sim \ln 1/x$ structure appears that is similar to the DGLAP energy distribution in vacuum. On the contrary, for $x \gg \bar{x}_c$, corresponding to the single hard scattering regime, $D \sim 1/x$, and the two regimes are separated by $\bar{x}_c$. The limiting case of eq. (4.14) is shown in figure 9 with dashed lines for different times (different panels). Since we, as we move onward, will largely neglect the description of the infared regime, we only used the single hard scattering (green) contribution that is valid $x > \bar{x}_c$.

Formally, the early-time expansion breaks down when $t > \lambda$ which is also the characteristic time when multiple interactions with the medium become important. Finally, we do not expect the early time solution to hold for $x < \bar{\alpha}^2\omega_c/E$ where multiple emissions play an important role.

Late time evolution. The evolution equation can be solved exactly if one assumes coherent scatterings dominate for all momentum fractions x and all times t. This approximation is most sound when $t \gg L_c$, as seen in figure 3. The analytical solution neglects the Bethe-Heitler, which will make the solution less reliable for very small $x < \omega_{\mathrm{BH}}/E$. We will call this solution D_0 as it serves as a baseline for subsequent calculations. It is the solution to the evolution equation

$$\frac{\partial}{\partial t}D_0(x,t) = \int_x^1 \mathrm{d}z\, \mathcal{K}_{\mathrm{coh}}\left(z,\frac{x}{z}E,t\right) D_0\left(\frac{x}{z},t\right) - \int_0^1 \mathrm{d}z\, z\mathcal{K}_{\mathrm{coh}}\left(z,xE,t\right) D_0(x,t) \tag{4.15}$$

JHEP02(2023)156

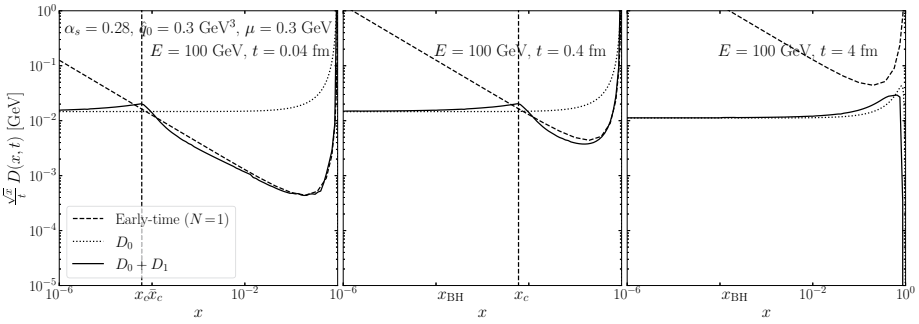


Figure 9. Analytic solutions of the energy distribution in different approximations (eqs. (4.13), (4.16) and (4.26)). With these parameters $\lambda = 0.06\,\text{fm}$, and thus the different panels are $t < \lambda$, $\lambda < t < L_c$, and $t \lesssim L_c$.

where $\mathcal{K}_{\rm coh}$ is the soft limit of the harmonic oscillator regime, see the middle term in eq. (4.12). The solution from ref. [76] is

$$D_0(x,\tau) = \frac{\tau}{\sqrt{x}(1-x)^{3/2}} \mathrm{e}^{-\pi \frac{\tau^2}{1-x}}\,, \tag{4.16}$$

where we dropped the $\delta(x)$ zero mode term, which is only important for energy conservation. We have also defined the re-scaled evolution variable which absorbs the energy scale,

$$\tau = \bar{\alpha}\sqrt{\frac{\hat{q}_0}{E}}t\,. \tag{4.17}$$

The solution has the $D_0 \sim \tau/\sqrt{x}$ shape characteristic of the turbulent cascade. Equation (4.16) is shown in figure 9 with dotted lines for different times.

Intermediate time evolution. To reach an approximate solution for intermediate times it is useful to cast the evolution in different variables. Following ref. [77], and defining $\xi = x/z$ in the gain term and $\xi = xz$ in the loss term the evolution can be rewritten as

$$\frac{\partial}{\partial\tau} D(x,\tau) = \int_x^1 \mathrm{d}\xi\, P(x,\xi,\tau) D(\xi,\tau) - D(x,\tau)\int_0^x \mathrm{d}\xi\, P(\xi,x,\tau)\,, \tag{4.18}$$

where $P(x,\xi,\tau) = \frac{1}{\bar{\alpha}}\sqrt{\frac{E}{\hat{q}_0}}\frac{x}{\xi^2}\mathcal{K}(\frac{x}{\xi},\xi E,t)$ and τ is defined in (4.17). This form of the evolution equation clearly shows the cancellation of divergences between the gain and loss terms at $\xi \to x$.

At early times, $t < \lambda$, the emissions are governed by single interactions with the medium. In the regime $\lambda < t < L_c$ you also have to take into account that coherent emissions play a role in the soft regime $\omega < \omega_c(t)$ (neglecting Bethe-Heitler emissions). However, the region above $\omega_c(t)$ will contain emissions from single hard scatterings, see eq. (4.12) (third line). In that case the kernel will be modified to include both a soft and a hard component, divided by the $\omega_c(t)$ separation line, as follows

$$P(x,\xi,\tau) = \theta_c(x,\xi,\tau) P_{\rm coh}(x,\xi,\tau) + \theta_h(x,\xi,\tau) P_{\rm hard}(x,\xi,\tau)\,, \tag{4.19}$$

JHEP02(2023)156

where the coherent and hard emission kernels were defined in eq. (4.12). After changing variables they become

$$P_{\rm coh}(x,\xi,\tau) = \sqrt{\frac{\xi}{x}}\frac{1}{(\xi-x)^{3/2}} \qquad \text{for } x \ll x_c \quad \text{or} \quad \xi - x_c \ll x\,,$$
$$P_{\rm hard}(x,\xi,\tau) = \frac{\pi}{2\bar{\alpha}}\frac{\xi}{x}\frac{1}{(\xi-x)^2}\tau \qquad \text{for } x_c \ll x \ll \xi - x_c\,, \tag{4.20}$$

where $x_c \equiv \omega_c(t)/E = \tau^2/(2\bar{\alpha}^2)$ and we have assumed that $x_c \ll 1$ in the limits. These conditions can be encoded in a set of Heaviside theta-functions to make sure that each kernel is used solely in its regime of validity, namely[18]

$$\theta_c(x,\xi,\tau) = \Theta(x_c - x) + \Theta(x_c + x - \xi) - \Theta(x_c - x)\Theta(x_c + x - \xi)\,,$$
$$\theta_h(x,\xi,\tau) = \Theta(x - x_c)\Theta(\xi - x_c - x)\,. \tag{4.21}$$

The now τ-dependent separation line $x_c(\tau)$ distinguishes the different regimes.

The full solution can be written as

$$D(x,\tau) = D_0(x,\tau) + \delta D(x,\tau)\,, \tag{4.22}$$

where $D_0(x,\tau)$ is a solution to the coherent, soft kernel defined in eq. (4.16), and $\delta D(x,\tau)$ is a correcting factor.

Inserting this into the evolution equation (4.18), we get

$$\frac{\partial}{\partial\tau}D(x,\tau) = \int_x^1 \mathrm{d}\xi\, P(x,\xi,\tau)D_0(\xi,\tau) - D_0(x,\tau)\int_0^x \mathrm{d}\xi\, P(\xi,x,\tau)$$
$$+ \int_x^1 \mathrm{d}\xi\, P(x,\xi,\tau)\delta D(\xi,\tau) - \delta D(x,\tau)\int_0^x \mathrm{d}\xi\, P(\xi,x,\tau)\,. \tag{4.23}$$

Taking into account that $\theta_c + \theta_h = 1$, one can rewrite the kernel as

$$P(x,\xi,\tau) = P_{\rm coh}(x,\xi,\tau) + \delta P(x,\xi,\tau)\,, \tag{4.24}$$

where $\delta P \equiv (P_{\rm hard} - P_{\rm coh})\theta_h$. Inserting the new kernel into eq. (4.23), the term $\partial D_0/\partial\tau$ cancels, and we are left with an iterative formula for δD,

$$\delta D(x,\tau) = D_1(x,\tau)$$
$$+ \int_0^\tau \mathrm{d}\sigma \int_x^1 \mathrm{d}\xi\, P(x,\xi,\sigma)\delta D(\xi,\sigma) - \int_0^\tau \mathrm{d}\sigma\delta D(x,\sigma)\int_0^x \mathrm{d}\xi\, P(\xi,x,\sigma)\,. \tag{4.25}$$

Here we have defined the leading term in the correction as

$$D_1(x,\tau) = \int_0^\tau \mathrm{d}\sigma \int_x^1 \mathrm{d}\xi\, \delta P(x,\xi,\sigma)D_0(\xi,\sigma) - \int_0^\tau \mathrm{d}\sigma D_0(x,\sigma)\int_0^x \mathrm{d}\xi\, \delta P(\xi,x,\sigma)\,, \tag{4.26}$$

[18] In the case where x_c is not small the conditions are slightly more complicated, and the hard regime is encoded in $\theta_h(x,\xi,\tau) = \Theta\left(x - \frac{\xi}{2}\left(1 - \sqrt{1 - \frac{4x_c}{\xi}}\right)\right)\Theta\left(\frac{\xi}{2}\left(1 + \sqrt{1 - \frac{4x_c}{\xi}}\right) - x\right)$, and $\theta_c = 1 - \theta_h$.

JHEP02(2023)156

which is given entirely in terms of known functions. To capture the main modifications with respect to the purely coherent solution $D_0(x,\tau)$, it is sufficient to keep only $D_1(x,\tau)$. This is sound because $D_0(x,t=0)=\delta(1-x)$, implying that $\delta D(x,t=0)=0$. Therefore, all terms going as $\sim\delta D$ start out small. At later times, single hard emissions are rare, and thus δD becomes less and less important.

Based on this discussion it is reasonable to assume that one can approximate the intermediate time solution by the sum of the two leading terms D_0+D_1. For this to be true it must be checked that it reproduces the correct behavior at early and late times. The early times expansion is

$$\begin{aligned}\lim_{\tau\to 0}(D_0(x,\tau)+D_1(x,\tau)) &\simeq \int_0^\tau \mathrm{d}\sigma\int_x^1 \mathrm{d}\xi\, P_{\mathrm{coh}}(x,\xi,\sigma)\delta(1-\xi)\\ &\quad+\int_0^\tau \mathrm{d}\sigma\int_x^1 \mathrm{d}\xi\,\left[P_{\mathrm{hard}}(x,\xi,\tau)-P_{\mathrm{coh}}(x,\xi,\tau)\right]\delta(1-\xi)\\ &=\frac{\pi}{4\bar{\alpha}}\frac{\tau^2}{x(1-x)^2}\,,\end{aligned}\tag{4.27}$$

where we kept the leading term and ignored virtual terms containing $\delta(1-x)$. Therefore, the sum D_0+D_1 reproduces the hard part of the early time expansion given in eq. (4.14). Moreover, at late times ($t>L_c$), the phase space for hard ($\omega>\omega_c$) emissions vanishes, and thus $D_1\to 0$. Hence, at late times the intermediate time solution simply goes to the late time solution D_0. Consequently, we expect that the sum of the two first terms D_0+D_1 to provide a decent approximation of the true solution at all times. One can systematically calculate corrections to this solution by iterating eq. (4.25).

The the early time solution eq. (4.13) (dashed), the soft limit of the HO approximation (4.16) (dotted), and the first correction D_0+D_1 are shown in figure 9 with full lines for different times. For short lengths (left panel), D_0+D_1 closely resembles the early time solution, as expected. For late times D_0+D_1 reduces to D_0, as there is not any phase space for hard emissions left. The D_0 presents the small x tail $D_0\sim\sqrt{x}$ characteristic for turbulence [76]. At intermediate times we see that D_0+D_1 goes to D_0 at low x, while at high x there is a suppression due to the lack of coherent scatterings at early times. Qualitatively, figure 9 resembles the full numerical solution shown in figure 10. In order to compare the two figures, we have marked the value of $x_{\mathrm{BH}}\equiv\omega_{\mathrm{BH}}/E$ in figure 9. We have also not included the color coding in this figure since it does not include the physics from all the relevant regimes represented in figure 3.

5 Numerical evaluation of the medium cascade

In section 3 we presented an effective framework that describes medium-induced emissions up to arbitrary precision. By using this framework we showed in section 4 how different scattering processes contribute to multiple induced emissions. Based on the properties of the medium (e.g. length, mean free path), not all induced emissions are necessary to resum (or to consider many of them). For example, induced emissions from hard scatterings are not as important to resum as emissions from multiple soft scatterings. We developed a simple analytic model to include a single hard emission correction to the resummation of

JHEP02(2023)156

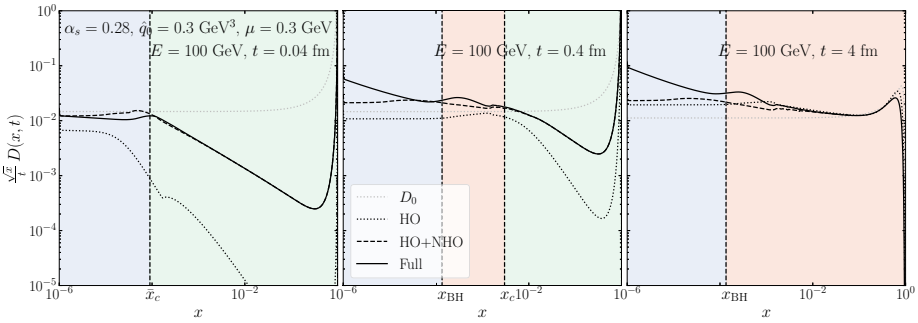


Figure 10. The energy distribution solution of the medium-induced emission evolution eq. (4.8) at different times (different panels) and using different scattering mechanisms to calculate the emission rate in eq. (4.9) (different line styles). The vertical dashed lines and the shaded areas separate different scattering regions, where on the first panel $t < \lambda$ and on the last panel $t > L_c$.

the multiple soft ones. In section 4, we made several simplifications e.g. we simplified the emission kernel $\mathcal{K}$, and neglected Bethe-Heitler emissions. It is important to understand how the energy distribution behaves also without using these simplifications and to understand the error introduced by employing them. That is the subject of this section.

In this section, we evaluate eq. (4.8) numerically using the kernel in our new framework. For more details about the numerical implementation, see appendix D and the complementary code [78]. The resulting energy distribution is presented in figure 10, where the different lines stem from using emission kernels $\mathcal{K}(z, E, t)$ at varying levels of approximation. The three panels correspond to three different time stages in the evolution: $t < \lambda$, $\lambda < t < L_c$, and $t > L_c$. As a reminder, there are no rare hard splittings for $t > L_c$, which is evident on the rightmost panel of figure 10.

The evolution starts at $t = 0$, with a single gluon of energy distribution $D(x, 0) = \delta(1-x)$ with $E = 100\,\text{GeV}$ energy. We solely use gluons during the evolution for simplicity, however our formalism is valid for other flavors too. The kernels we use include finite-z corrections, see appendix C.3 for more details. The vertical dashed lines in figure 10 separate the regions where different scattering processes dominate, which are the same regions as in figure 3. The phase space for emissions is determined by comparing $z(1-z)E$ with $\bar{\omega}_c(t), \omega_{\rm BH}$, and $\omega_c(t)$ in the relevant regions, as it is described in appendix C. The turbulent cascade solution $D_0(x, t)$ from eq. (4.16) is also shown as a baseline, and was discussed in more detail in section 4.3.

It is important to note that vacuum emissions are not included in the current study, and thus many important effects (e.g. vacuum fragmentation, medium resolution, color coherence), which are essential if one wishes to compare to measurements, will not be discussed in this work. We refer the interested readers to refs. [29, 44, 45, 61, 74] for further details.

Harmonic oscillator. A simple and much studied method of solving the evolution equation eq. (4.8) is by using the harmonic oscillator (HO) approximation of the emission

JHEP02(2023)156

rate (see for example in refs. [20, 23, 29, 76, 77, 79–81]). The kernel is then simply given by

$$\mathcal{K}(z,E,t) \approx 2\left.\frac{\mathrm{d}^2 I^{\mathrm{HO}}}{\mathrm{d}z\mathrm{d}t}\right|_E , \tag{5.1}$$

where the HO approximation was discussed in section 3.4 and in the numerics we used eq. (C.27). In this approximation, the induced emissions originate from multiple soft elastic scatterings with the medium, and we expect this process to dominate for $t \gg \lambda$ and $x_{\mathrm{BH}} \ll x \ll x_c(t)$ (red region in figure 3).

The energy distribution obtained by using the HO approximation is shown in figure 10 with black dotted lines. The HO spectrum exhibits the well know turbulent cascade behavior [76, 77], resulting in a characteristic tail $D(x \ll 1, t) \sim t/\sqrt{x}$, which can be seen on the plot as horizontal lines at small x. This is also a feature of the simplified analytic solution D_0. The turbulent cascade involves a constant flux of energy propagating in time to $x \to 0$. The turbulence appears below $x < x_c(t)$, and the energy increases with time due to the time dependent emission phase space (see also $\omega_c(t)$ in figure 3).

Above $x > x_c(t)$ the HO kernel switches from $\mathcal{K} \sim 1/z^{3/2}$ to $1/z^3$ and thus the energy distribution starts going as $D(x,t) \sim 1/x^2$. This is especially visible in the middle panel of figure 10.

It is interesting to note that due to the running of $\hat{q}$ (see eq. (3.43)), deviations from the pure $D \sim 1/\sqrt{x}$ are expected in the HO solution. A consequence of this can be seen as a small kink on the dotted curves (most visible in the left panel at $x \approx 10^{-4}$). Below the kink, $\hat{q}(\omega) \to \hat{q}_0$ is used.

Improved opacity expansion. The IOE from section 3.4 makes it possible to extend the HO description to include rare hard scatterings, covering both red and green regions in figure 3. Here, the evolution equation is solved using the IOE kernel, which is

$$\mathcal{K}(z,E,t) \approx 2\left.\frac{\mathrm{d}^2 I^{\mathrm{IOE}}}{\mathrm{d}z\mathrm{d}t}\right|_E . \tag{5.2}$$

This is valid from early to late times, for energies above the Bethe-Heitler regime $x \gg x_{\mathrm{BH}}$.

The energy distribution obtained using the IOE (HO+NHO is used in the numerical implementation from eqs. (C.27)–(C.34)) is shown in figure 10 with black dashed lines. In the $x \ll x_c(t)$ region the HO dominates and $D(x,t)$ qualitatively does not change by including hard emissions. The turbulent tail $D \sim 1/\sqrt{x}$ is still present. The offset between the HO and HO+NHO results originates from the effective contribution of the NHO term to $\hat{q} \approx \hat{q}_0 \ln \frac{Q_x^2}{\mu_*^2}(1 + 1.016 \ln^{-1} \frac{Q_x^2}{\mu_*^2})$.

In the region $x > x_c(t)$, the distribution function changes from the purely HO result, due to the inclusion of the NHO corrections. The splitting function here behaves as $\mathcal{K} \sim 1/z^2$, resulting in the distribution going as $D(x) \sim 1/x$, which is visible in the green region. Our analytic result from section 4.3, given by $D_0 + D_1$, includes a single hard emission in addition to the HO cascade. Comparing the analytic results in figure 9 with the numerical ones in figure 10 it is evident that the analytical result succeeds in capturing qualitatively the behavior induced by using the IOE kernel.

JHEP02(2023)156

Several models are based mainly on the $N = 1$ term, see refs. [82–84], and also refs. [85, 86]. However, these models miss the soft scatterings that are present in a large medium with $L > \lambda$. Soft scatterings are crucial to include to achieve an accurate description of the energy distribution for $x < x_c(t)$. We would like to also emphasize that by numerically solving eq. (4.8), we also resum hard emissions and thus we consider the possibility of emitting arbitrarily many of them.

Full solution. Finally, we present the energy distribution using our new picture. It is given by solving the evolution equation (4.8) using the full kernel

$$\mathcal{K}(z,E,t) = 2\frac{\mathrm{d}^2 I^{\mathrm{Full}}}{\mathrm{d}z\mathrm{d}t} = 2\begin{cases} \frac{\mathrm{d}^2 I^{\mathrm{ROE}}}{\mathrm{d}z\mathrm{d}t}\,, & \text{if} \quad z(1-z)E < \min(\bar{\omega}_c(t), \omega_{\mathrm{BH}})\,, \\ \frac{\mathrm{d}^2 I^{\mathrm{IOE}}}{\mathrm{d}z\mathrm{d}t}\,, & \text{otherwise}\,. \end{cases} \tag{5.3}$$

This covers the full emission phase space in (z, t, E) and was derived in detail in section 3. In practice we used $N_r = 1$ and HO+NHO terms, derived from eqs. (3.31), (3.34), (3.58) and (3.61). This provides an excellent approximation to the true result, as can be seen when comparing to the numerical evaluation of the kernel in figure 7.

The resulting energy distribution is shown as full black lines in figure 10. One can see that the results obtained by solely using the IOE (dashed lines) agrees very well with the full kernel for $x \gg x_{\mathrm{BH}}$, but they start to differ when $x \lesssim x_{\mathrm{BH}}$. It is clear that all the difference between these two curves comes from Bethe-Heitler emissions. An increasing tail appears at low x, which becomes more and more important for later times, due to the logarithmic soft limit of the ROE kernel. In this region the ROE kernel is $\mathcal{K} \sim \ln(z)/z$, resulting in the energy distribution going as $D \sim \ln \frac{1}{x}$, which is similar to the DGLAP evolution in vacuum.

A new, interesting bump also appears in the full solution close to x_{BH}, which warrants an explanation. First, the energy flux that brings quanta from $x = 1$ to $x = 0$ is not the same on the two sides of x_{BH}. If the energy transport is more efficient from the right (red region), that will result in a slowing of the flux when going from higher to lower x. The bump can then be understood as a sediment of energy building up around x_{BH}. Secondly, the emission kernel defined in eq. (5.3) is not smooth around the transition point $\min(\bar{\omega}_c(t), \omega_{\mathrm{BH}})$. This discontinuity introduces additional uncertainties in the true behavior of the energy distribution around x_{BH}. This transition would smoothen by including more orders of the expansion, which should be studied in the future. This uncertainty could also be connected with the observed bump. In our numerical implementation we introduced a smoothing function to minimize this uncertainty, but it is still present. In addition, other effects like $2 \to 2$ scattering process and thermal masses are also important here [52], which we do not presently discuss. A thorough study of this region is needed where all these effects are included, which motivates future work.

Bethe-Heitler emissions have a soft divergence and thus an IR regulator ω_{min} has to be introduced. A similar regulator was introduced in the numerical differential equation solver. For more details, see appendix D.

JHEP02(2023)156

6 Conclusions and outlook

In this paper, we studied emissions induced by elastic scatterings on the quark gluon plasma. We considered the interplay of the relevant length scales of the problem, which are the propagation length, the mean free path between scatterings, and the formation time. From these considerations, we derived the emergent energy scales that separate the induced emission spectrum into regimes governed by different scattering processes. The emerging hierarchy of scales and related processes is illustrated in figure 3.

We presented a new theoretical framework consisting of different perturbative expansions, namely the opacity expansion (OE), the improved OE (IOE), and the resummed OE (ROE), which we derived rigorously. Together these are suitable to describe the induced emission spectrum in different regimes. We showed that at least one of the expansions is valid in every phase space (ω, t) point, where ω is the energy of the emitted gluon and t is the propagation time (length) in the medium.[19] While these expansions are formally simply reorganizations of the multiple scattering series, it is also important to note that none of the expansions *when truncated at any fixed order* is valid everywhere. Relying on the multiple approaches to obtain the spectrum, our composite framework can account for an arbitrary number of medium interactions that can be soft or hard. Most importantly, it is systematically improvable and can, in principle, describe the emission spectrum (and rate) up to arbitrary precision, which has only been achieved numerically before [48, 49, 51].

In the current work, we have elucidated the convergence properties up to second order in the studied resummations. Identifying the expansion structure in the different regimes opens for the possibility of studying the accuracy of resummations in the medium. Finally, our new description provides a quick and efficient way to evaluate the induced emission spectrum (and rate), which is an essential ingredient of medium-induced cascades and jet quenching study. Our implementation is available in an online repository [78].

In order to tackle multiple emissions, we also studied the in-medium energy distribution $D(x, t)$ within our new formalism. This is an essential ingredient of jet quenching phenomenology as it describes how the energy of a leading particle gets distributed within a cascade. Having first identified the conditions to generate multiple emissions in the medium, we demonstrated the separation between early, rare emissions — generated mainly by a single momentum exchange with the medium — and a following cascade of soft splittings. We also showed that emissions are formed independently up to suppressed terms, justifying a posteriori the formulation of the cascade via a rate equation. We developed new analytic tools to combine the resummation of multiple soft with rare hard emissions, and thus we showed how different scattering processes appear on the level of the energy distribution. Finally, using numerical evaluation based on the previously derived full splitting kernels, we showed more rigorously the transition effects between different scattering regions. We identified the importance of the time dependent phase space separation, multiple Bethe-Heitler emissions, and the running of $\hat{q}$, to mention some.

Even though this work can immediately be applied to jet quenching phenomenology one should also tackle other challenges, such as vacuum emissions and coherence effects, which

[19] For the general case valid beyond the strictly soft limit ω refers to the reduced energy of the three-body evolution $\omega = z(1-z)E$, where E is the initial energy of the emitter and z is the momentum sharing fraction.

JHEP02(2023)156

we have neglected here and left for future studies. We also acknowledge the importance of soft emissions for which one should study thermal masses, energy changing $2 \to 2$ scatterings, and thermalization [52, 87] to correctly describe the very infrared regime close to the thermal scale and below. Moreover, other non-perturbative effects such as expanding, inhomogeneous medium are also important, in which direction our framework is extendable [23, 28, 88, 89].

Acknowledgments

We thank Paul Caucal, Alexandre Falcão, Aleksas Mazeliauskas, and Sören Schlichting for helpful discussions. We also thank Carlota Andres, Liliana Apolinário, Fabio Dominguez, and Marcos Gonzalez Martinez for providing us with the numerical results for the spectrum computed in ref. [51]. The work is supported by a Starting Grant from Trond Mohn Foundation (BFS2018REK01) and the University of Bergen.

A All order formulas for medium-induced spectrum in the soft limit

Here we present some general formulas for the different expansions, that aided us in calculating the spectrum in section 3. In all cases the full spectrum is given as the sum of all the terms $\omega\frac{\mathrm{d}I}{\mathrm{d}\omega} = \sum_n \omega\frac{\mathrm{d}I^{N=n}}{\mathrm{d}\omega}$. The expansion therefore only converges if the terms decrease sufficiently fast order by order $\omega\frac{\mathrm{d}I^{N=n+1}}{\mathrm{d}\omega} < \omega\frac{\mathrm{d}I^{N=n}}{\mathrm{d}\omega}$.

A.1 Opacity expansion

Combining the formulas for the opacity expansion eq. (3.6) and the spectrum eq. (3.12) one can write the formula for the n^{th} term of the opacity expansion as[20]

$$\omega\frac{\mathrm{d}I^{N=n}}{\mathrm{d}\omega} = (-1)^{n-1}\frac{4\alpha_s C_R}{\omega}\int_{\boldsymbol{p}_n,\dots,\boldsymbol{p}_1}\Sigma(\boldsymbol{p}_n^2)\frac{\boldsymbol{p}_n\cdot\boldsymbol{p}_1}{\boldsymbol{p}_n^2}v(\boldsymbol{p}_n-\boldsymbol{p}_{n-1})\dots v(\boldsymbol{p}_2-\boldsymbol{p}_1) \tag{A.1}$$
$$\times\mathrm{Re}\,i\int_0^L \mathrm{d}t_n\int_0^{t_n}\mathrm{d}t_0\int_{t_0}^{t_n}\mathrm{d}t_{n-1}\int_{t_0}^{t_{n-1}}\mathrm{d}t_{n-2}\cdots\int_{t_0}^{t_2}\mathrm{d}t_1\,\mathrm{e}^{-i\frac{\boldsymbol{p}_n^2}{2\omega}(t_n-t_{n-1})}\dots\mathrm{e}^{-i\frac{\boldsymbol{p}_1^2}{2\omega}(t_1-t_0)},$$

where $v(\boldsymbol{p},s)$ is given in eq. (3.7) and $\Sigma(\boldsymbol{p}^2)$ in given in eq. (3.13) in the GW model. It is useful to go to unitless integration variables, by defining $\sqrt{L/2\omega}\boldsymbol{p}_k \to \boldsymbol{p}_k$ and $\frac{t_k}{L} \to t_k$,

$$\omega\frac{\mathrm{d}I^{N=n}}{\mathrm{d}\omega} = (-1)^{n-1}8\pi\bar{\alpha}\left(\frac{L}{\lambda}\right)^n\frac{\bar{\omega}_c}{\omega}\int_{\boldsymbol{p}_n,\dots,\boldsymbol{p}_1}\tilde{\Sigma}(\boldsymbol{p}_n^2)\frac{\boldsymbol{p}_n\cdot\boldsymbol{p}_1}{\boldsymbol{p}_n^2}\tilde{v}(\boldsymbol{p}_n-\boldsymbol{p}_{n-1})\dots\tilde{v}(\boldsymbol{p}_2-\boldsymbol{p}_1)$$
$$\times\mathrm{Re}\,i\int_0^1 \mathrm{d}t_n\int_0^{t_n}\mathrm{d}t_0\int_{t_0}^{t_n}\mathrm{d}t_{n-1}\int_{t_0}^{t_{n-1}}\mathrm{d}t_{n-2}\cdots\int_{t_0}^{t_2}\mathrm{d}t_1\,\mathrm{e}^{-i\boldsymbol{p}_n^2(t_n-t_{n-1})}\dots\mathrm{e}^{-i\boldsymbol{p}_1^2(t_1-t_0)}. \tag{A.2}$$

Here we have used the unitless function $\tilde{v}(\boldsymbol{p}) = (2\pi)^2\delta(\boldsymbol{p}) - \frac{\bar{\omega}_c}{\omega}\tilde{\sigma}(\boldsymbol{p})$, where in the GW model

$$\tilde{\sigma}(\boldsymbol{p}) = \frac{4\pi}{\left(\boldsymbol{p}^2+\frac{\bar{\omega}_c}{\omega}\right)^2},$$
$$\tilde{\Sigma}(\boldsymbol{p}^2) = \frac{1}{\boldsymbol{p}^2+\frac{\bar{\omega}_c}{\omega}}. \tag{A.3}$$

[20] The gluon color factors appearing in v are not trivial at n^{th} order and thus we refer the reader to appendix C for further discussion.

JHEP02(2023)156

The momentum and time integrals only depend on the unitless combination $\bar{\omega}_c/\omega$. This means that the spectrum will take the form $\omega\frac{\mathrm{d}I^{N=n}}{\mathrm{d}\omega} = \bar{\alpha}\left(\frac{L}{\lambda}\right)^n h_n(\frac{\omega}{\bar{\omega}_c})$, for some function h_n. This would naively imply convergence for $L/\lambda < 1$, but the exact for of the function h_n must also be taken into account. In section 3.2 we showed for $N = 1, 2$ at low energy $\omega \ll \bar{\omega}_c$ that $h_n(\frac{\omega}{\bar{\omega}_c})$ is a finite function, meaning the series converges when $L/\lambda < 1$. However, at high energy $\omega \gg \bar{\omega}_c$ the function takes the form $h_n(\frac{\bar{\omega}_c}{\omega}) \sim (\frac{\bar{\omega}_c}{\omega})^n \tilde{h}_n(\frac{\bar{\omega}_c}{\omega})$, where $\tilde{h}_n$ is finite, implying convergence when $\frac{L}{\lambda}\frac{\bar{\omega}_c}{\omega} < 1$. Proving this is true to all orders is deferred to future work.

One can also derive a corresponding formula for the rate $\omega\frac{\mathrm{d}I^{N=n}}{\mathrm{d}\omega\mathrm{d}L}$, by taking the length derivative of eq. (A.2)

$$\begin{aligned}
&\omega\frac{\mathrm{d}I^{N=n}}{\mathrm{d}\omega\mathrm{d}L} \\
&= (-1)^{n-1} 8\pi\bar{\alpha}\frac{1}{L}\left(\frac{L}{\lambda}\right)^n \frac{\bar{\omega}_c}{\omega}\int_{\boldsymbol{p}_n,\ldots,\boldsymbol{p}_1} \tilde{\Sigma}(\boldsymbol{p}_n^2)\frac{\boldsymbol{p}_n\cdot\boldsymbol{p}_1}{\boldsymbol{p}_n^2}\tilde{v}(\boldsymbol{p}_n-\boldsymbol{p}_{n-1})\ldots\tilde{v}(\boldsymbol{p}_2-\boldsymbol{p}_1) \\
&\quad\times\mathrm{Re}\, i\int_0^1 \mathrm{d}t_0\int_{t_0}^1 \mathrm{d}t_{n-1}\int_{t_0}^{t_{n-1}}\mathrm{d}t_{n-2}\cdots\int_{t_0}^{t_2}\mathrm{d}t_1\, \mathrm{e}^{-i\boldsymbol{p}_n^2(1-t_{n-1})}\mathrm{e}^{-i\boldsymbol{p}_{n-1}^2(t_{n-1}-t_{n-2})}\ldots\mathrm{e}^{-i\boldsymbol{p}_1^2(t_1-t_0)}.
\end{aligned} \tag{A.4}$$

A.2 Resummed opacity expansion

One can also derive all order formulas for the resummed opacity expansion, by starting with eq. (3.26) and following the same procedure,

$$\begin{aligned}
&\omega\frac{\mathrm{d}I^{N_r=n}}{\mathrm{d}\omega} \\
&= \frac{4\alpha_s C_R}{\omega}\int_{\boldsymbol{p}_n,\ldots,\boldsymbol{p}_1}\Sigma(\boldsymbol{p}_n^2)\frac{\boldsymbol{p}_n\cdot\boldsymbol{p}_1}{\boldsymbol{p}_n^2}\sigma(\boldsymbol{p}_n-\boldsymbol{p}_{n-1})\ldots\sigma(\boldsymbol{p}_2-\boldsymbol{p}_1) \\
&\quad\times\mathrm{Re}\, i\int_0^L \mathrm{d}t_n\int_0^{t_n}\mathrm{d}t_0\int_{t_0}^{t_n}\mathrm{d}t_{n-1}\int_{t_0}^{t_{n-1}}\mathrm{d}t_{n-2}\cdots\int_{t_0}^{t_2}\mathrm{d}t_1\,\Delta(t_n,t_0)\mathrm{e}^{-i\frac{\boldsymbol{p}_n^2}{2\omega}(t_n-t_{n-1})}\ldots\mathrm{e}^{-i\frac{\boldsymbol{p}_1^2}{2\omega}(t_1-t_0)}.
\end{aligned} \tag{A.5}$$

The ROE and OE expansions are equivalent at infinite order at low opacity $L/\lambda \ll 1$, but at finite order the terms are mixed up. The terms containing a delta function in the potential v in the OE are included in the Sudakov factor Δ in the ROE. One can see this by taking $\Delta \to 1$ and $v \to -\sigma$ in eq. (A.5) and eq. (A.1) respectively, in which case the expansions become exactly the same. This is however not a good approximation at any order.

After changing to unitless variables we arrive at

$$\begin{aligned}
&\omega\frac{\mathrm{d}I^{N_r=n}}{\mathrm{d}\omega} \\
&= 8\pi\bar{\alpha}\left(\frac{L}{\lambda}\right)^n\left(\frac{\bar{\omega}_c}{\omega}\right)^n\int_{\boldsymbol{p}_n,\ldots,\boldsymbol{p}_1}\tilde{\Sigma}(\boldsymbol{p}_n^2)\frac{\boldsymbol{p}_n\cdot\boldsymbol{p}_1}{\boldsymbol{p}_n^2}\tilde{\sigma}(\boldsymbol{p}_n-\boldsymbol{p}_{n-1})\ldots\tilde{\sigma}(\boldsymbol{p}_2-\boldsymbol{p}_1) \\
&\quad\times\mathrm{Re}\, i\int_0^1 \mathrm{d}t_n\int_0^{t_n}\mathrm{d}t_0\int_{t_0}^{t_n}\mathrm{d}t_{n-1}\int_{t_0}^{t_{n-1}}\mathrm{d}t_{n-2}\cdots\int_{t_0}^{t_2}\mathrm{d}t_1\,\mathrm{e}^{-\frac{L}{\lambda}(t_n-t_0)}\mathrm{e}^{-i\boldsymbol{p}_n^2(t_n-t_{n-1})}\ldots\mathrm{e}^{-i\boldsymbol{p}_1^2(t_1-t_0)},
\end{aligned} \tag{A.6}$$

where we have used that in the static medium $\Delta(t_n, t_0) = \exp(-\frac{t_n-t_0}{\lambda})$. Again we use the unitless functions $\tilde{\Sigma}(\boldsymbol{p}^2)$ and $\tilde{\sigma}(\boldsymbol{p})$ defined in eq. (A.3) for the GW model. In this case the

JHEP02(2023)156

time integrals can be done analytically

$$
\begin{aligned}
&i\int_0^1 \mathrm{d}t_n \int_0^{t_n} \mathrm{d}t_0 \int_{t_0}^{t_n} \mathrm{d}t_{n-1} \int_{t_0}^{t_{n-1}} \mathrm{d}t_{n-2} \cdots \int_{t_0}^{t_2} \mathrm{d}t_1\, \mathrm{e}^{-\frac{L}{\lambda}(t_n-t_0)} \mathrm{e}^{-i\boldsymbol{p}_n^2(t_n-t_{n-1})} \ldots \mathrm{e}^{-i\boldsymbol{p}_1^2(t_1-t_0)} \\
&= i^n \int_0^1 \mathrm{d}t_n \int_0^{t_n} \mathrm{d}t_0 \sum_{k=1}^{n} \frac{\mathrm{e}^{-i\left(\boldsymbol{p}_k^2 - i\frac{L}{\lambda}\right)(t_n-t_0)}}{\prod_{l\neq k}^{n}(\boldsymbol{p}_k^2-\boldsymbol{p}_l^2)} = \left(\cos\frac{n\pi}{2} + i\sin\frac{n\pi}{2}\right) \sum_{k=1}^{n} \frac{T\left(\boldsymbol{p}_k^2 - i\frac{L}{\lambda}\right)}{\prod_{l\neq k}^{n}(\boldsymbol{p}_k^2-\boldsymbol{p}_l^2)},
\end{aligned}
\tag{A.7}
$$

where we have defined the function

$$
T(x) = \int_0^1 \mathrm{d}t_n \int_0^{t_n} \mathrm{d}t_0\, \mathrm{e}^{-ix(t_n-t_0)} = \frac{1-ix-\mathrm{e}^{-x}}{x^2}. \tag{A.8}
$$

The reason this simplification was not possible in the OE is that the formula for the time integrals is only valid if $\boldsymbol{p}_k \neq \boldsymbol{p}_l$, while the OE has terms containing $\delta(\boldsymbol{p}_k - \boldsymbol{p}_{k-1})$. However, in the ROE there are no such deltas, so the formula is valid. In the end the spectrum is

$$
\begin{aligned}
\omega\frac{\mathrm{d}I^{N_r=n}}{\mathrm{d}\omega} &= 8\pi\bar{\alpha}\left(\frac{L}{\lambda}\right)^n \left(\frac{\bar{\omega}_c}{\omega}\right)^n \int_{\boldsymbol{p}_n,\ldots,\boldsymbol{p}_1} \tilde{\Sigma}(\boldsymbol{p}_n^2)\frac{\boldsymbol{p}_n\cdot\boldsymbol{p}_1}{\boldsymbol{p}_n^2}\tilde{\sigma}(\boldsymbol{p}_n-\boldsymbol{p}_{n-1})\ldots\tilde{\sigma}(\boldsymbol{p}_2-\boldsymbol{p}_1) \\
&\quad\times\sum_{k=1}^{n}\frac{\cos\frac{n\pi}{2}\mathrm{Re}\,T(\boldsymbol{p}_k^2-i\chi) - \sin\frac{n\pi}{2}\mathrm{Im}\,T(\boldsymbol{p}_k^2-i\chi)}{\prod_{l\neq k}^{n}(\boldsymbol{p}_k^2-\boldsymbol{p}_l^2)}.
\end{aligned}
\tag{A.9}
$$

The remaining momentum integrals will in the end be some function of $\bar{\omega}_c/\omega$ and opacity $\chi = L/\lambda$. As a reference, the real and imaginary parts of $T(\boldsymbol{p}^2 - i\chi)$ are

$$
\begin{aligned}
\mathrm{Re}\,T(\boldsymbol{p}^2-i\chi) &= \frac{\chi(\boldsymbol{p}^4+\chi^2)+\boldsymbol{p}^4-\chi^2-\mathrm{e}^{-\chi}\left((\boldsymbol{p}^4-\chi^2)\cos\boldsymbol{p}^2+2\chi\boldsymbol{p}^2\sin\boldsymbol{p}^2\right)}{(\boldsymbol{p}^4+\chi^2)^2} \\
-\mathrm{Im}\,T(\boldsymbol{p}^2-i\chi) &= \frac{\boldsymbol{p}^2(\boldsymbol{p}^4+\chi^2)-2\chi\boldsymbol{p}^2-\mathrm{e}^{-\chi}\left((\boldsymbol{p}^4-\chi^2)\sin\boldsymbol{p}^2-2\chi\boldsymbol{p}^2\cos\boldsymbol{p}^2\right)}{(\boldsymbol{p}^4+\chi^2)^2}.
\end{aligned}
\tag{A.10}
$$

It is also possible to look at the limits of this function when the opacity χ is low or high. In the low opacity case $\chi \ll 1$ we have

$$
\begin{aligned}
\omega\frac{\mathrm{d}I^{N_r=n}}{\mathrm{d}\omega} &\simeq 8\pi\bar{\alpha}\left(\frac{L}{\lambda}\right)^n \left(\frac{\bar{\omega}_c}{\omega}\right)^n \int_{\boldsymbol{p}_n,\ldots,\boldsymbol{p}_1} \tilde{\Sigma}(\boldsymbol{p}_n^2)\frac{\boldsymbol{p}_n\cdot\boldsymbol{p}_1}{\boldsymbol{p}_n^2}\tilde{\sigma}(\boldsymbol{p}_n-\boldsymbol{p}_{n-1})\ldots\tilde{\sigma}(\boldsymbol{p}_2-\boldsymbol{p}_1) \\
&\quad\times\sum_{k=1}^{n}\frac{\cos\frac{n\pi}{2}(1-\cos\boldsymbol{p}_k^2)+\sin\frac{n\pi}{2}(\boldsymbol{p}_k^2-\sin\boldsymbol{p}_k^2)}{\boldsymbol{p}_k^4\prod_{l\neq k}^{n}(\boldsymbol{p}_k^2-\boldsymbol{p}_l^2)}.
\end{aligned}
\tag{A.11}
$$

In this case the momentum integrals will give a function of $\bar{\omega}_c/\omega$. As mentioned in section 3.3 is it preferable to use the OE (A.2) at low opacity, as the order of opacity is mixed up in the ROE.

In the high opacity case $\chi \gg 1$ you get

$$
\begin{aligned}
\omega\frac{\mathrm{d}I^{N_r=n}}{\mathrm{d}\omega} &\simeq 8\pi\bar{\alpha}\frac{L}{\lambda}\left(\frac{\omega_{\mathrm{BH}}}{\omega}\right)^n \int_{\boldsymbol{p}_n,\ldots,\boldsymbol{p}_1} \tilde{\Sigma}(\boldsymbol{p}_n^2,\omega_{\mathrm{BH}})\frac{\boldsymbol{p}_n\cdot\boldsymbol{p}_1}{\boldsymbol{p}_n^2}\tilde{\sigma}(\boldsymbol{p}_n-\boldsymbol{p}_{n-1},\omega_{\mathrm{BH}})\ldots\tilde{\sigma}(\boldsymbol{p}_2-\boldsymbol{p}_1,\omega_{\mathrm{BH}}) \\
&\quad\times\sum_{k=1}^{n}\frac{\cos\frac{n\pi}{2}+\sin\frac{n\pi}{2}\boldsymbol{p}_k^2}{(1+\boldsymbol{p}_k^4)\prod_{l\neq k}^{n}(\boldsymbol{p}_k^2-\boldsymbol{p}_l^2)}.
\end{aligned}
\tag{A.12}
$$

JHEP02(2023)156

Here we see the emergence of the scale $\omega_{\rm BH} = \frac{1}{2}\mu^2\lambda$, which takes the place of $\bar{\omega}_c$ in the functions $\tilde{\Sigma}$ and $\tilde{\sigma}$, as emphasized in the above formula. The momentum integrals then only become some function of $\omega_{\rm BH}/\omega$. Notice that this scales as $\sim L/\lambda$ at every order, meaning it can converge also when $L/\lambda > 1$.

In section 3.4 we calculated the high opacity limit to orders $N = 1, 2$. There it was clear that for low $\omega \ll \omega_{\rm BH}$ the $N = 2$ limit is subleading compared to the $N = 1$ limit, while at high $\omega \gg \omega_{\rm BH}$ the $N = 2$ and $N = 1$ limits are of the same order. This seems to imply that the expansion converges at low ω, while it breaks down at high ω. Again, proving this for all orders is deferred to future work.

Again the calculation for the rate $\omega\frac{\mathrm{d}I^{N_r=n}}{\mathrm{d}\omega\mathrm{d}L}$ is similar

$$\begin{aligned}
&\omega\frac{\mathrm{d}I^{N_r=n}}{\mathrm{d}\omega\mathrm{d}L} \\
&= 8\pi\bar{\alpha}\frac{1}{L}\left(\frac{L}{\lambda}\right)^n\left(\frac{\bar{\omega}_c}{\omega}\right)^n\int_{\boldsymbol{p}_n,\ldots,\boldsymbol{p}_1}\tilde{\Sigma}(\boldsymbol{p}_n^2)\frac{\boldsymbol{p}_n\cdot\boldsymbol{p}_1}{\boldsymbol{p}_n^2}\tilde{\sigma}(\boldsymbol{p}_n-\boldsymbol{p}_{n-1})\ldots\tilde{\sigma}(\boldsymbol{p}_2-\boldsymbol{p}_1) \\
&\quad\times\sum_{k=1}^{n}\frac{\cos\frac{n\pi}{2}\left[\chi-\mathrm{e}^{-\chi}\left(\chi\cos\boldsymbol{p}_k^2-\boldsymbol{p}_k^2\sin\boldsymbol{p}_k^2\right)\right]+\sin\frac{n\pi}{2}\left[\boldsymbol{p}_k^2-\mathrm{e}^{-\chi}\left(\boldsymbol{p}_k^2\cos\boldsymbol{p}_k^2+\chi\sin\boldsymbol{p}_k^2\right)\right]}{\left(\boldsymbol{p}_k^4+\chi^2\right)\prod_{l\neq k}^{n}(\boldsymbol{p}_k^2-\boldsymbol{p}_l^2)}\,.
\end{aligned} \tag{A.13}$$

A.3 Improved opacity expansion

The spectrum for the improved opacity expansion can be written as the iterative equation

$$\begin{aligned}
\omega\frac{\mathrm{d}I^{\rm IOE}}{\mathrm{d}\omega} &= -\frac{2\alpha_s C_R}{\omega^2}\mathrm{Re}\int_0^\infty\mathrm{d}t_2\int_0^{t_2}\mathrm{d}t_1\int_{t_1}^{t_2}\mathrm{d}s\int\mathrm{d}^2\boldsymbol{z} \\
&\quad\times\partial_{\boldsymbol{x}}\cdot\partial_{\boldsymbol{y}}[\mathcal{K}_{\rm HO}(\boldsymbol{x},t_2;\boldsymbol{z},s)\delta v(\boldsymbol{z},s)\mathcal{K}(\boldsymbol{z},s;\boldsymbol{y},t_1)]_{\boldsymbol{x}=\boldsymbol{y}=0}\,,
\end{aligned} \tag{A.14}$$

where the zeroth order solution is the HO from eq. (3.57) has to be added. In the following we will make use of results from refs. [53, 54]. The harmonic oscillator propagator is

$$\mathcal{K}_{\rm HO}\left(\boldsymbol{x},t_2;\boldsymbol{y},t_1\right) = \frac{\omega}{2\pi i S\left(t_2,t_1\right)}\exp\left(\frac{i\omega}{2S\left(t_2,t_1\right)}\left[C\left(t_1,t_2\right)\boldsymbol{x}^2+C\left(t_2,t_1\right)\boldsymbol{y}^2-2\boldsymbol{x}\cdot\boldsymbol{y}\right]\right)\,. \tag{A.15}$$

The functions S and C are given implicitly by

$$\begin{aligned}
&\left[\frac{\mathrm{d}^2}{\mathrm{d}^2t}+\Omega^2(t)\right]S\left(t,t_0\right)=0, \qquad S\left(t_0,t_0\right)=0, \qquad \partial_t S\left(t,t_0\right)_{t=t_0}=1 \\
&\left[\frac{\mathrm{d}^2}{\mathrm{d}^2t}+\Omega^2(t)\right]C\left(t,t_0\right)=0, \qquad C\left(t_0,t_0\right)=1, \qquad \partial_t C\left(t,t_0\right)_{t=t_0}=0\,,
\end{aligned} \tag{A.16}$$

where the frequency $\Omega(t)$ is given by

$$\Omega(t) = \frac{1-i}{2}\sqrt{\frac{\hat{q}(t)}{\omega}} \tag{A.17}$$

To continue it is useful to apply the formulas

$$\begin{aligned}
\int_s^\infty\mathrm{d}t_2\partial_{\boldsymbol{x}}\mathcal{K}_{\rm HO}(\boldsymbol{x},t_2;\boldsymbol{z},s)|_{\boldsymbol{x}=0} &= \frac{-i\omega}{\pi}\frac{\boldsymbol{z}}{\boldsymbol{z}^2}\mathrm{e}^{i\frac{\omega}{2}\Omega(s)^2\frac{S(s,L)}{C(s,L)}\boldsymbol{z}^2} \\
\int_0^s\mathrm{d}t_1\partial_{\boldsymbol{y}}\mathcal{K}_{\rm HO}(\boldsymbol{z},s;\boldsymbol{y},t_1)|_{\boldsymbol{y}=0} &= \frac{-i\omega}{\pi}\frac{\boldsymbol{z}}{\boldsymbol{z}^2}\mathrm{e}^{-i\frac{\omega}{2}\frac{C(0,s)}{S(0,s)}\boldsymbol{z}^2}\,.
\end{aligned} \tag{A.18}$$

JHEP02(2023)156

Then the spectrum takes the following form

$$\omega\frac{\mathrm{d}I^{\mathrm{IOE}}}{\mathrm{d}\omega}=\frac{2\bar{\alpha}}{\omega}\mathrm{Re}\,i\int_0^L\mathrm{d}t_2\int_0^{t_2}\mathrm{d}t_1\int\mathrm{d}^2\boldsymbol{x}\,\mathrm{e}^{-i\frac{\omega\Omega}{2}\tan(\Omega(L-t_2))\boldsymbol{x}^2}\delta v(\boldsymbol{x})\frac{\boldsymbol{x}}{\boldsymbol{x}^2}\cdot\partial_{\boldsymbol{y}}\mathcal{K}(\boldsymbol{x},t_2;\boldsymbol{y},t_1)|_{\boldsymbol{y}=0}\,, \tag{A.19}$$

where we have used that for the brick medium $\hat{q}(t)=\Theta(L-t)\hat{q}$, the functions S and C are simply

$$S\left(t_2,t_1\right)=\frac{1}{\Omega}\sin\Omega\left(t_2-t_1\right),\quad\text{and}\quad C\left(t_2,t_1\right)=\cos\Omega\left(t_2-t_1\right). \tag{A.20}$$

One can derive a formula for the IOE at arbitrary order, which was also explored in ref. [55].

$$\begin{aligned}\omega\frac{\mathrm{d}I^{N_I=n}}{\mathrm{d}\omega}=(-1)^{n-1}\frac{2\bar{\alpha}}{\pi}\int\mathrm{d}^2\boldsymbol{x}_n\ldots\mathrm{d}^2\boldsymbol{x}_1\,\frac{\boldsymbol{x}_n\cdot\boldsymbol{x}_1}{\boldsymbol{x}_n^2\boldsymbol{x}_1^2}\delta v(\boldsymbol{x}_n)\ldots\delta v(\boldsymbol{x}_1)\\ \times\,\mathrm{Re}\int_0^L\mathrm{d}t_n\int_0^{t_n}\mathrm{d}t_{n-1}\cdots\int_0^{t_2}\mathrm{d}t_1\,\mathrm{e}^{i\frac{\omega\Omega}{2}\left[\cot(\Omega t_1)\boldsymbol{x}_1^2-\tan(\Omega(L-t_n))\boldsymbol{x}_n^2\right]}\\ \times\,\mathcal{K}_{\mathrm{HO}}(\boldsymbol{x}_n,t_n;\boldsymbol{x}_{n-1},t_{n-1})\ldots\mathcal{K}_{\mathrm{HO}}(\boldsymbol{x}_2,t_2;\boldsymbol{x}_1,t_1)\,.\end{aligned} \tag{A.21}$$

After changing to unitless variables by defining $\boldsymbol{u}_k=\sqrt{\mu^2\omega/(2\bar{\omega}_c)}\boldsymbol{x}_k$ and $s_k=\frac{t_k}{L}$,

$$\begin{aligned}\omega\frac{\mathrm{d}I^{N_I=n}}{\mathrm{d}\omega}=(-1)^{n-1}\frac{2\bar{\alpha}}{\pi}\left(\frac{L}{\lambda}\right)^n\left(\frac{\bar{\omega}_c}{\omega}\right)^n\int\mathrm{d}^2\boldsymbol{u}_n\ldots\mathrm{d}^2\boldsymbol{u}_1\,\frac{\boldsymbol{u}_n\cdot\boldsymbol{u}_1}{\boldsymbol{u}_n^2\boldsymbol{u}_1^2}\delta\tilde{v}(\boldsymbol{u}_n)\ldots\delta\tilde{v}(\boldsymbol{u}_1)\\ \times\,\mathrm{Re}\int_0^1\mathrm{d}s_n\int_0^{s_n}\mathrm{d}s_{n-1}\cdots\int_0^{s_2}\mathrm{d}s_1\,\mathrm{e}^{i\frac{\sigma}{2}\sqrt{\frac{\omega_c}{\omega}}\left[\cot(\sigma\sqrt{\frac{\omega_c}{\omega}}s_1)\boldsymbol{u}_1^2-\tan(\sigma\sqrt{\frac{\omega_c}{\omega}}(1-s_n))\boldsymbol{u}_n^2\right]}\\ \times\,\tilde{\mathcal{K}}_{\mathrm{HO}}(\boldsymbol{u}_n,s_n;\boldsymbol{u}_{n-1},s_{n-1})\ldots\tilde{\mathcal{K}}_{\mathrm{HO}}(\boldsymbol{u}_2,s_2;\boldsymbol{u}_1,s_1)\,.\end{aligned} \tag{A.22}$$

Here we have defined $\sigma=\frac{1-i}{\sqrt{2}}$ and the unitless functions

$$\begin{aligned}\delta\tilde{v}(\boldsymbol{u})&=\frac{1}{2}\boldsymbol{u}^2\ln\left(\frac{\omega}{\bar{\omega}_c}\frac{\mu^2}{2Q^2}\frac{1}{\boldsymbol{u}^2}\right),\\ \tilde{\mathcal{K}}_{\mathrm{HO}}(\boldsymbol{u}_2,s_2;\boldsymbol{u}_1,s_1)&=\frac{\sigma\sqrt{\frac{\omega_c}{\omega}}}{2\pi i\sin\left(\sigma\sqrt{\frac{\omega_c}{\omega}}(s_2-s_1)\right)}\\ &\quad\times\mathrm{e}^{\frac{i\sigma\sqrt{\frac{\omega_c}{\omega}}}{2\sin\left(\sigma\sqrt{\frac{\omega_c}{\omega}}(s_2-s_1)\right)}\left[\cos\left(\sigma\sqrt{\frac{\omega_c}{\omega}}(s_2-s_1)\right)(\boldsymbol{u}_2^2+\boldsymbol{u}_1^2)-2\boldsymbol{u}_2\cdot\boldsymbol{u}_1\right]}.\end{aligned} \tag{A.23}$$

As the integrals only depend on $\sqrt{\frac{\omega_c}{\omega}}$ and $\frac{\omega}{\bar{\omega}_c}\frac{\mu^2}{2Q^2}$ the IOE spectrum can be written as $\omega\frac{\mathrm{d}I^{N_I=n}}{\mathrm{d}\omega}=\frac{2\bar{\alpha}}{\pi}\left(\frac{L}{\lambda}\right)^n\left(\frac{\bar{\omega}_c}{\omega}\right)^n f_n\left(\sqrt{\frac{\omega_c}{\omega}},\frac{\omega}{\bar{\omega}_c}\frac{\mu^2}{2Q^2}\right)$ where the function f_n is given by the integrals. The soft limit $\omega\ll\omega_c$ of the IOE expansion was discussed in detail in [55], and also in section 3.4.

In the hard limit $\omega\gg\omega_c$, the spectrum becomes

$$\begin{aligned}\omega\frac{\mathrm{d}I^{N_I=n}}{\mathrm{d}\omega}\simeq(-1)^{n-1}\frac{2\bar{\alpha}}{\pi}\left(\frac{L}{\lambda}\right)^n\left(\frac{\bar{\omega}_c}{\omega}\right)^n\int\mathrm{d}^2\boldsymbol{u}_n\ldots\mathrm{d}^2\boldsymbol{u}_1\,\frac{\boldsymbol{u}_n\cdot\boldsymbol{u}_1}{\boldsymbol{u}_n^2\boldsymbol{u}_1^2}\delta\tilde{v}(\boldsymbol{u}_n)\ldots\delta\tilde{v}(\boldsymbol{u}_1)\\ \times\,\mathrm{Re}\int_0^1\mathrm{d}s_n\int_0^{s_n}\mathrm{d}s_{n-1}\cdots\int_0^{s_2}\mathrm{d}s_1\,\mathrm{e}^{i\frac{\boldsymbol{u}_1^2}{2s_1}}\\ \times\,\tilde{\mathcal{K}}_0(\boldsymbol{u}_n,s_n;\boldsymbol{u}_{n-1},s_{n-1})\ldots\tilde{\mathcal{K}}_0(\boldsymbol{u}_2,s_2;\boldsymbol{u}_1,s_1),\end{aligned} \tag{A.24}$$

JHEP02(2023)156

where the BDMPS propagator has gone to the vacuum propagator

$$\tilde{\mathcal{K}}_0(\boldsymbol{u}_2, s_2; \boldsymbol{u}_1, s_1) = \frac{1}{2\pi i(s_2 - s_1)} \mathrm{e}^{i\frac{i(\boldsymbol{u}_2 - \boldsymbol{u}_1)^2}{2(s_2 - s_1)}} \,. \tag{A.25}$$

Notice that there is no remaining dependence on the BDMPS scale ω_c. In section 3.4 we calculated the hard limit to first order, and the resulting expression (3.61) is the same as the OE limit (3.16) for high ω. Whether this correspondence is also true at higher orders is an interesting question that will be explored in future work.

B General formulas for the emission rate in the soft limit

In this appendix, we gather the formulas relevant for computing the emission rate in the soft limit. The rate can be written as

$$\omega \frac{\mathrm{d}I}{\mathrm{d}\omega \mathrm{d}t} = \frac{4\bar{\alpha}\pi}{\omega} \mathrm{Re}\, i \int_0^t \mathrm{d}t_1 \int_{\boldsymbol{p},\boldsymbol{p}_0} \Sigma(\boldsymbol{p}^2, t) \frac{\boldsymbol{p}\cdot\boldsymbol{p}_0}{\boldsymbol{p}^2} \mathcal{K}(\boldsymbol{p}, t; \boldsymbol{p}_0, t_1) \,, \tag{B.1}$$

in momentum-space representation of the three-point function, and

$$\omega \frac{\mathrm{d}I}{\mathrm{d}\omega \mathrm{d}t} = \frac{2\bar{\alpha}}{\omega} \mathrm{Re}\, i \int_0^t \mathrm{d}t_1 \int_{\boldsymbol{z}} v(\boldsymbol{z}, t) \frac{\boldsymbol{z}}{\boldsymbol{z}^2} \cdot \boldsymbol{\partial}_{\boldsymbol{y}} \, \mathcal{K}(\boldsymbol{z}, t; \boldsymbol{y}, t_1)|_{\boldsymbol{y}=0} \,, \tag{B.2}$$

in coordinate-space representation. Equation (B.1) can be employed directly to derive expressions for the rate in the OE and ROE, by simply inserting the expansions (3.6) and (3.26). We will not attempt at deriving higher-order corrections to these rates here, since they can be also be found for a medium with constant density by taking the appropriate derivative with respect to length on the expression for the spectrum.

For the IOE, the harmonic oscillator spectrum is directly calculable, following the decomposition in eqs. (3.49) and (3.50). For the IOE rates, we find

$$\omega \frac{\mathrm{d}I^{\mathrm{HO}}}{\mathrm{d}\omega \mathrm{d}t} = \frac{\bar{\alpha}}{2\omega} \mathrm{Re}\, i \int_0^t \mathrm{d}t_1 \int_{\boldsymbol{z}} \hat{q}(t) \boldsymbol{z} \cdot \boldsymbol{\partial}_{\boldsymbol{y}} \, \mathcal{K}_{\mathrm{HO}}(\boldsymbol{z}, t; \boldsymbol{y}, t_1)|_{\boldsymbol{y}=0} \,, \tag{B.3}$$

$$\omega \frac{\mathrm{d}I^{\mathrm{IOE}}}{\mathrm{d}\omega \mathrm{d}t} = \frac{2\bar{\alpha}}{\omega} \mathrm{Re}\, i \int_0^t \mathrm{d}t_1 \int_{\boldsymbol{z}} \delta v(\boldsymbol{z}, t) \frac{\boldsymbol{z}}{\boldsymbol{z}^2} \cdot \boldsymbol{\partial}_{\boldsymbol{y}} \, \mathcal{K}(\boldsymbol{z}, t; \boldsymbol{y}, t_1)|_{\boldsymbol{y}=0} \,, \tag{B.4}$$

where the three-point correlator $\mathcal{K}(\boldsymbol{z}; \boldsymbol{y})$ is found from iterating (3.45). As a cross-check, for a medium with constant density we obtain

$$\omega \frac{\mathrm{d}I^{\mathrm{HO}}}{\mathrm{d}\omega \mathrm{d}t} = \bar{\alpha} x \, \mathrm{Re}\,(i-1) \tan\left[\frac{1-i}{2} xt\right] = \bar{\alpha} x \frac{\sinh(xt) - \sin(xt)}{\cosh(xt) + \cos(xt)} \,, \tag{B.5}$$

where $x \equiv \sqrt{\hat{q}/\omega}$, for the harmonic oscillator term.

C Medium-induced spectrum and rate with finite-z corrections

The process we study is a parton of energy E splitting into two partons with energy zE and $(1-z)E$. In the main text, we refer to the emitted energy zE as ω. However, we stress that this definition is only true in the soft limit. In the more general case we refer to ω as the

JHEP02(2023)156

reduced energy of the three-body evolution, that is $\omega = z(1-z)E$. This quantity is only equal to the emitted energy when $z \to 0$, in which case it reduces to $\omega \simeq zE$. This appendix accounts for how our framework generalizes when considering all the finite-z contributions, which means we strictly use the full definition $\omega = z(1-z)E$.

Reference [57] already has an implementation of the OE ($N = 1$) and IOE (HO+NHO) medium-induced emission spectrum in the strictly soft limit ($\omega \ll E$) for a homogeneous brick. We improve on this by keeping finite-z terms, including the rates, and by including the OE and ROE expansions. The resulting code is available online [78].

The starting equation for keeping finite-z corrections can be found in ref. [54],

$$\frac{\mathrm{d}I_{ba}^{\mathrm{med}}}{\mathrm{d}z} = \frac{\alpha_s}{\omega^2} P_{ba}(z) \mathrm{Re} \int_0^\infty \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_1 \, \boldsymbol{\partial}_{\boldsymbol{x}} \cdot \boldsymbol{\partial}_{\boldsymbol{y}} \left[\mathcal{K}_{ba}(\boldsymbol{x}, t_2; \boldsymbol{y}, t_1) - \mathcal{K}_0(\boldsymbol{x}, t_2; \boldsymbol{y}, t_1) \right]_{\boldsymbol{x}=\boldsymbol{y}=0} , \tag{C.1}$$

where the parent parton a carrying energy E splits into partons b and c, carrying energy zE and $(1-z)E$, respectively. It is the finite-z analog of eq. (3.1). The Altarelli-Parisi splitting functions are

$$\begin{aligned} P_{gq}(z) &= C_F \frac{1+(1-z)^2}{z} , & P_{qq}(z) &= P_{gq}(1-z) , \\ P_{gg}(z) &= C_A \frac{[1+z(1-z)]^2}{z(1-z)} , & P_{qg}(z) &= N_f T_F \left[z^2 + (1-z)^2 \right] , \end{aligned} \tag{C.2}$$

which are valid for $0 < z < 1$. In the soft limit $z \ll 1$, for quarks the splitting function reduces to $P_q(z) \approx \frac{2C_F}{z}$, while for gluons (where $1 - z \ll 1$ also has to be included) $P_g(z) \approx \frac{C_A}{z(1-z)} \approx \frac{2C_A}{z}$, where in the last step the $1 - z$ contribution has been folded to z with the additional factor of 2. The three-point correlator $\mathcal{K}(\boldsymbol{x}; \boldsymbol{y})$ satisfies the following Schrödinger-like equation

$$\left[i\frac{\partial}{\partial t} + \frac{\boldsymbol{\partial}_{\boldsymbol{x}}^2}{2\omega} + i v_{ba}(\boldsymbol{x}, t) \right] \mathcal{K}_{ba}(\boldsymbol{x}, t; \boldsymbol{y}, t_0) = i\delta(t - t_0)\delta(\boldsymbol{x} - \boldsymbol{y}) , \tag{C.3}$$

where the potential v_{ba} describes the splitting induced by partons scattering with the medium,

$$v_{ba}(\boldsymbol{x}, t) = \frac{C_{cba}}{2N_c} v(\boldsymbol{x}, t) + \frac{C_{acb}}{2N_c} v(z\boldsymbol{x}, t) + \frac{C_{bac}}{2N_c} v((1-z)\boldsymbol{x}, t) , \tag{C.4}$$

where $C_{ijk} \equiv C_i + C_j - C_k$ and C_i is the Casimir operator squared for particle i and $v(\boldsymbol{x}, t)$ is defined in eq. (3.3). In the soft limit, $v_g(\boldsymbol{x}, t) \approx v_q(\boldsymbol{x}, t) \approx \frac{C_{b,c}}{N_c} v(\boldsymbol{x}, t)$, where the soft emission is always a gluon (C_b or $C_c = N_c$). Surprisingly, this shows that the potential is sensitive to the emitted gluon's and not the emitter's color in the soft limit because v is proportional with N_c by definition. This was observed previously in the opacity expansion and was explained heuristically in [36]. For quarks in the $z \to 1$ limit, the potential goes to $v_q(\boldsymbol{x}, t) \approx \frac{C_F}{N_c} v(\boldsymbol{x}, t)$, where the color factor compensates the N_c in v.

In momentum space, we find

$$v_{ba}(\boldsymbol{q}, t) = \frac{C_{cba}}{2N_c} v(\boldsymbol{q}, t) + \frac{C_{acb}}{2N_c} \frac{1}{z^2} v\left(\frac{\boldsymbol{q}}{z}, t\right) + \frac{C_{bac}}{2N_c} \frac{1}{(1-z)^2} v\left(\frac{\boldsymbol{q}}{1-z}, t\right) , \tag{C.5}$$

JHEP02(2023)156

where $v(\boldsymbol{q},t) = (2\pi)^2\delta(\boldsymbol{q})\Sigma(t) - \sigma(\boldsymbol{q},t)$, and $\Sigma(t) = \Sigma(0,t) = \int_{\boldsymbol{q}}\sigma(\boldsymbol{q},t)$. In the soft limit, $v_g(\boldsymbol{q},t) \approx v_q(\boldsymbol{q},t) \approx v(\boldsymbol{q},t)$, which only becomes apparent once the explicit form of $\sigma(\boldsymbol{q})$ is used in $v(\frac{\boldsymbol{q}}{z})$. It is at this point worth extending the definition of the interaction potential to also include an argument defining the screening mass, i.e.

$$\sigma(\boldsymbol{q},t,\mu) \equiv \frac{4\pi\hat{q}_0(t)}{(\boldsymbol{q}^2+\mu^2)^2}\,, \tag{C.6}$$

for the GW model (see eq. (3.4)), and similarly for the inverse mean free path $\Sigma(\boldsymbol{p}^2,t) \to \Sigma(\boldsymbol{p}^2,t,\mu)$, where $\Sigma(\boldsymbol{p}^2,t,\mu) = \int_{\boldsymbol{q}}\sigma(\boldsymbol{q},t,\mu)\Theta(\boldsymbol{q}^2-\boldsymbol{p}^2)$. Then, $\frac{1}{z^2}\sigma(\frac{\boldsymbol{p}}{z},t,\mu) = z^2\sigma(\boldsymbol{p},t,z\mu)$.

Further simplifications can be made following the discussion in section 3.1. For the spectrum in momentum space representation, generalizing eq. (3.12) to finite-z, we arrive at

$$\frac{\mathrm{d}I_{ba}}{\mathrm{d}z} = \frac{2\alpha_s}{\omega}P_{ba}(z)\,\mathrm{Re}\,i\int_0^L \mathrm{d}t_2\int_0^{t_2}\mathrm{d}t_1\int_{\boldsymbol{p},\boldsymbol{p}_0}\Sigma_{ba}(\boldsymbol{p}^2,t_2)\frac{\boldsymbol{p}\cdot\boldsymbol{p}_0}{\boldsymbol{p}^2}\mathcal{K}_{ba}(\boldsymbol{p},t_2;\boldsymbol{p}_0,t_1)\,, \tag{C.7}$$

where

$$\Sigma_{ba}(\boldsymbol{q}^2,t) = \frac{C_{cba}}{2N_c}\Sigma(\boldsymbol{q}^2,t,\mu)+\frac{C_{acb}}{2N_c}z^2\Sigma(\boldsymbol{q}^2,t,z\mu)+\frac{C_{bac}}{2N_c}(1-z)^2\Sigma(\boldsymbol{q}^2,t,(1-z)\mu)\,, \tag{C.8}$$

and the three-point function in momentum-space representation is found through the implicit equation

$$\begin{aligned}\mathcal{K}_{ba}(\boldsymbol{p},t;\boldsymbol{p}_0,t_0) &= (2\pi)^2\delta(\boldsymbol{p}-\boldsymbol{p}_0)\mathcal{K}_0(\boldsymbol{p};t-t_0)\\ &\quad-\int_{t_0}^t\mathrm{d}s\int_{\boldsymbol{q}}\mathcal{K}_0(\boldsymbol{p};t-s)v_{ba}(\boldsymbol{q},s)\mathcal{K}_{ba}(\boldsymbol{p}-\boldsymbol{q},s;\boldsymbol{p}_0,t_0)\,.\end{aligned} \tag{C.9}$$

Then, analogously to the derivations in appendix B, the rate at finite-z reads

$$\frac{\mathrm{d}I_{ba}}{\mathrm{d}z\mathrm{d}t} = \frac{2\alpha_s}{\omega}P_{ba}(z)\mathrm{Re}\,i\int_0^t\mathrm{d}t_1\int_{\boldsymbol{p},\boldsymbol{p}_0}\Sigma_{ba}(\boldsymbol{p}^2,t)\frac{\boldsymbol{p}\cdot\boldsymbol{p}_0}{\boldsymbol{p}^2}\mathcal{K}_{ba}(\boldsymbol{p},t;\boldsymbol{p}_0,t_1)\,. \tag{C.10}$$

Similar manipulations in coordinate-space representation will be done directly in the IOE section below.

To simplify the expressions below we also introduce the shorthand that accounts for the recurring combinations of color and z factors, see e.g. in eq. (C.8). Hence, we have

$$\sum_{p=1}^{3}\mathcal{C}_p z_p^2 f(z_p^2 x) = \frac{C_{cba}}{2N_c}f(x)+\frac{C_{acb}}{2N_c}z^2 f(z^2x)+\frac{C_{bac}}{2N_c}(1-z)^2 f((1-z)^2x)\,, \tag{C.11}$$

that runs over the three cyclic permutations of $\{a,b,c\}$. Here $\mathcal{C}_p = \left[\frac{C_{cba}}{2N_c},\frac{C_{acb}}{2N_c},\frac{C_{bac}}{2N_c}\right]$ and $z_p = [1,z,(1-z)]$, with p running from 1 to 3. In the soft limit this expression simply becomes $\sum_p \mathcal{C}_p z_p^2 f(z_p^2 x) \to f(x)$.

C.1 Opacity expansion

From eq. (C.7), and following section 3.2, we find the $N=1$ contribution of the OE at finite-z to be,

$$\frac{\mathrm{d}I_{ba}^{N=1}}{\mathrm{d}z} = \frac{2\alpha_s}{\pi}\frac{P_{ba}(z)}{z(1-z)}\frac{L}{\lambda}\frac{\mu^2 L}{2E}\sum_p \mathcal{C}_p z_p^2\mathcal{I}_{N=1}(z_p^2 y)\,, \tag{C.12}$$

JHEP02(2023)156

where $\lambda = \mu^2/\hat{q}_0$ is the mean free path, $y = \frac{\bar{\omega}_c}{\omega} = \frac{\mu^2 L}{2z(1-z)E}$, and the relevant integral is[21]

$$\mathcal{I}_{N=1}(y) = \int_0^\infty \frac{\mathrm{d}u}{u^2}\frac{u-\sin u}{u+y} = \begin{cases} \frac{\pi}{4}\,, & \text{for } y \ll 1\,, \\ \frac{1}{y}\left(\gamma_E + \ln y\right)\,, & \text{for } y \gg 1\,. \end{cases} \tag{C.13}$$

We have also extracted the asymptotic behaviors for future convenience. The soft limit of a hard emission ($\bar{\omega}_c \ll zE \ll E$) reduces to $\frac{\mathrm{d}I}{\mathrm{d}z} = \bar{\alpha}\frac{\pi}{2}\frac{L}{\lambda}\frac{\bar{\omega}_c}{z^2 E}$, which is in agreement with eq. (3.16), with $\omega \to zE$. One could also use $P_g(z) \approx \frac{C_A}{z(1-z)}$ and keep $z(1-z)$, and then an extra 1/2 factor will appear to not double count both contributions $z, 1-z \ll 1$. The rate can be found directly from eq. (C.10) and is given by

$$\frac{\mathrm{d}I_{ba}^{N=1}}{\mathrm{d}z\mathrm{d}t} = \frac{2\alpha_s}{\pi}\frac{P_{ba}(z)}{z(1-z)}\frac{1}{\lambda}\frac{\mu^2 t}{2E}\sum_p \mathcal{C}_p z_p^2\,\tilde{\mathcal{I}}_{N=1}(z_p^2 y)\,, \tag{C.14}$$

where

$$\tilde{\mathcal{I}}_{N=1}(y) = \int_0^\infty \frac{\mathrm{d}u}{u}\frac{1-\cos u}{u+y} = \begin{cases} \frac{\pi}{2}\,, & \text{for } y \ll 1\,, \\ \frac{1}{y}\left(1 + \gamma_E + \ln y\right)\,, & \text{for } y \gg 1\,. \end{cases} \tag{C.15}$$

C.2 Resummed opacity expansion

The finite-z potential in eq. (C.4) introduces an additional complication for the ROE. In section 3.3, we separated and resummed the zero momentum exchange mode, while expanding in real scatterings. In eq. (C.4), however, this separation is more complicated because additional zero modes appear in the real terms in the $z, 1-z \to 0$ limits. Therefore, we make sure to explicitly subtract the zero mode terms in real scatterings

$$v_{ba}(\boldsymbol{p}, t) = (2\pi)^2 \sum_p \mathcal{C}_p \left[(1 - f(z_p))\delta(\boldsymbol{p})\Sigma(t) - \int_{\boldsymbol{q}} (\delta(\boldsymbol{p} - z_p\boldsymbol{q}) - f(z_p)\delta(\boldsymbol{p}))\sigma(\boldsymbol{q})\right]\,. \tag{C.16}$$

We introduced $f(z)$ arbitrary function, that goes to 1 in the soft limit $z \to 0$ (and $1-z \to 0$). In this paper, we make the choice $f(z) = 1 - z^2$. The resummed opacity expansion involves the Sudakov factor of the no elastic scattering probability,

$$\Delta_{ba}(t_2, t_1) = \exp\left[-\int_{t_1}^{t_2} \mathrm{d}s\,\hat{\Sigma}(s, z)\right]\,, \tag{C.17}$$

here $\hat{\Sigma}(s,z) = [\frac{C_{cba}}{2N_c} + z^2\frac{C_{acb}}{2N_c} + (1-z)^2\frac{C_{bac}}{2N_c}]\Sigma(s)$. In the soft limit $\hat{\Sigma} \to \Sigma$, and thus the Sudakov goes to eq. (3.25). The expansion reads

$$\begin{aligned} \mathcal{K}_{ba}(\boldsymbol{p}, t; \boldsymbol{p}_0, t_0) &= (2\pi)^2\delta(\boldsymbol{p} - \boldsymbol{p}_0)\Delta_{ba}(t_2, t_1)\mathcal{K}_0(\boldsymbol{p}; t_2 - t_1) \\ &\quad - \int_{t_1}^{t_2}\mathrm{d}s\int_q \Delta_{ba}(t_2, s)\mathcal{K}_0(\boldsymbol{p}; t_2 - s)\hat{\sigma}_{ba}(\boldsymbol{q}, s)\mathcal{K}_{ba}(\boldsymbol{p} - \boldsymbol{q}, s; \boldsymbol{p}_0, t_0)\,, \end{aligned} \tag{C.18}$$

where $\hat{\sigma}_{ba}(\boldsymbol{q}, s) = \sum_p \mathcal{C}_p \frac{1}{z_p^2}\sigma(\frac{\boldsymbol{q}}{z_p}, t) - \sum_p \mathcal{C}_p f(z_p)\Sigma(s)$.

[21] The integrals are available with trigonometric integral functions

$$\int_0^\infty \frac{\mathrm{d}u}{u^2}\frac{u-\sin u}{u+y} = \frac{1}{y^2}\left[y(\gamma_E - 1 + \ln y) + \pi\sin^2\frac{y}{2} - \mathrm{Ci}(y)\sin y + \mathrm{Si}(y)\cos y\right]\,,$$
$$\int_0^\infty \frac{\mathrm{d}u}{u}\frac{1-\cos u}{u+y} = \frac{1}{2y}\left[2(\gamma_E + \ln(y)) - 2\cos(y)\mathrm{Ci}(y) + \sin(y)(\pi - 2\mathrm{Si}(y))\right]\,,$$

where Euler Gamma γ_E, $\mathrm{Ci}(z) = -\int_z^\infty \mathrm{d}t/t\cos t$, and $\mathrm{Si}(z) = \int_0^t \mathrm{d}t/t\sin t$.

JHEP02(2023)156

The first order ($N_r = 1$) can be read directly off from eq. (C.7), and reads

$$\frac{\mathrm{d}I_{ba}^{N_r=1}}{\mathrm{d}z} = \frac{2\alpha_s}{\omega} P_{ba}(z)\,\mathrm{Re}\,i\int_0^L \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_1 \int_{\boldsymbol{p}} \Sigma_{ba}(\boldsymbol{p}^2, t_2)\mathcal{K}_0(\boldsymbol{p}; t_2 - t_1)\Delta_{ba}(t_2, t_1)\,,$$
$$= \frac{2\alpha_s}{\pi}\frac{P_{ba}(z)}{z(1-z)}\frac{L}{\lambda}\frac{\mu^2 L}{2E}\sum_p \mathcal{C}_p z_p^2 \mathcal{I}_{N_r=1}(z_p^2 y)\,, \tag{C.19}$$

where $y = \frac{\mu^2 L}{2\omega}$ and, defining $\hat{\chi} \equiv \hat{\Sigma} L$,

$$\mathcal{I}_{N_r=1}(y) = -\int_0^\infty \frac{\mathrm{d}u}{u+y}\,\mathrm{Im}T(u - i\hat{\chi})\,, \tag{C.20}$$

and $T(u) = (1 - iu - \mathrm{e}^{-iu})/u^2$. The imaginary part can also be written explicitly, as

$$-\,\mathrm{Im}\,T(u - i\hat{\chi}) = \frac{u[u^2 + \hat{\chi}(\hat{\chi} - 2)] + [2u\hat{\chi}\cos u - (u^2 - \hat{\chi}^2)\sin u]\,\mathrm{e}^{-\hat{\chi}}}{[u^2 + \hat{\chi}^2]^2}\,. \tag{C.21}$$

In the big medium limit ($\hat{\chi} \gg 1$), our formula reproduces eq. (3.34), $\frac{\mathrm{d}I}{\mathrm{d}z} = \frac{\bar{\alpha}}{z}\frac{L}{\lambda}\ln\frac{\mu^2\lambda}{2zE}$.

The rate follows directly from eq. (C.10), and reads

$$\frac{\mathrm{d}I_{ba}^{N_r=1}}{\mathrm{d}z\mathrm{d}t} = \frac{2\alpha_s}{\pi}\frac{P_{ba}(z)}{z(1-z)}\frac{1}{\lambda}\frac{\mu^2 t}{2E}\sum_p \mathcal{C}_p z_p^2 \tilde{\mathcal{I}}_{N_r=1}(z_p^2 y)\,, \tag{C.22}$$

with the relevant integral being,[22]

$$\tilde{\mathcal{I}}_{N_r=1}(y) = -\int_0^\infty \frac{\mathrm{d}u}{u+y}\,\mathrm{Im}\tilde{T}(u - i\hat{\chi})\,, \tag{C.23}$$

with $\tilde{T}(y) = (-i + i\mathrm{e}^{-iu})/u$. The imaginary part of this function is

$$-\,\mathrm{Im}\,\tilde{T}(u - i\hat{\chi}) = \frac{u - \mathrm{e}^{-\hat{\chi}}(u\cos u + \hat{\chi}\sin u)}{u^2 + \hat{\chi}^2}\,. \tag{C.24}$$

C.3 Improved opacity expansion

The IOE is similar to that we used in section 3.4, one takes a perturbative expansion in $\mu|\boldsymbol{x}| \ll 1$ of eq. (3.3) in eq. (C.4). By including the color and z-dependence of the splitting,

[22] The integral is analytical using $\mathrm{Ei}(x) = -\int_{-x}^\infty \mathrm{d}t\mathrm{e}^{-t}/t$,

$$\int_0^\infty \mathrm{d}u \frac{u(\mathrm{e}^\chi - \cos u) - \chi\sin u}{(u+y)(u^2+\chi^2)} = \frac{\mathrm{e}^{-\chi}}{2(\chi^2+y^2)}\Big[\pi x\mathrm{e}^\chi - \pi\chi(\cos y - \sin y) - 2y(\cos(y)\mathrm{Ci}(y) + \sin(y)\mathrm{Si}(y))$$
$$+2\chi(\cos(y)\mathrm{Si}(y) - \sin(y)\mathrm{Ci}(y)) + 2y\mathrm{e}^\chi\left(\ln\frac{y}{\chi} + \mathrm{Ei}(-x)\right)\Big]\,.$$

JHEP02(2023)156

an effective jet quenching parameter can be defined as[23]

$$\hat{q}_{ba}(z,t) \equiv \hat{q}_0(t)\left[\frac{C_{cba}}{2C_A}\ln\frac{Q^2}{\mu_*^2} + \frac{C_{acb}}{2C_A}z^2\ln\frac{Q^2}{z^2\mu_*^2} + \frac{C_{bac}}{2C_A}(1-z)^2\ln\frac{Q^2}{(1-z)^2\mu_*^2}\right],$$
$$= \hat{q}_0(t)\sum_p \mathcal{C}_p z_p^2 \ln\frac{Q^2}{z_p^2\mu_*^2}\,. \tag{C.25}$$

The harmonic oscillator potential is then $v_{ba}^{\rm HO}(\boldsymbol{x},t) = \frac{1}{4}\hat{q}_{ba}(z,t)\boldsymbol{x}^2$, which in the soft limit recovers the expression below eq. (3.42), $\hat{q}_{ba} \to \hat{q}$.

Harmonic oscillator. The harmonic oscillator spectrum is given by

$$\frac{\mathrm{d}I_{ba}^{\rm HO}}{\mathrm{d}z} = \frac{\alpha_s}{\pi}P_{ba}(z)\ln\left[\frac{1}{2}\left(\cos\left(\sqrt{\frac{2\omega_c}{\omega}}\right) + \cosh\left(\sqrt{\frac{2\omega_c}{\omega}}\right)\right)\right], \tag{C.26}$$

where $\omega_c = \frac{1}{2}\hat{q}_{ba}(z)L^2$. In the soft limit, $z\frac{\mathrm{d}I}{\mathrm{d}z} \approx 2\bar{\alpha}\sqrt{\hat{q}/(2z^3)}$, that reproduces the formula from eq. (3.58). The time-differential rate that appears in the evolution equation is

$$\frac{\mathrm{d}I_{ba}^{(\rm HO)}}{\mathrm{d}z\mathrm{d}L} = \frac{\alpha_s}{\pi}P_{ba}(z)\frac{1}{L}\sqrt{\frac{2\omega_c}{\omega}}\frac{\sinh\left(\sqrt{\frac{2\omega_c}{\omega}}\right) - \sin\left(\sqrt{\frac{2\omega_c}{\omega}}\right)}{\cos\left(\sqrt{\frac{2\omega_c}{\omega}}\right) + \cosh\left(\sqrt{\frac{2\omega_c}{\omega}}\right)}\,. \tag{C.27}$$

Next-to harmonic oscillator. The NHO spectrum is given by

$$\frac{\mathrm{d}I_{ba}^{\rm NHO}}{\mathrm{d}z} = \frac{\alpha_s}{\omega^2}P_{ba}(z)\mathrm{Re}\int_0^\infty \mathrm{d}t_2\int_0^{t_2}\mathrm{d}t_1\int_{\boldsymbol{z}}\int_0^L \mathrm{d}s\,\partial_{\boldsymbol{x}}\partial_{\boldsymbol{y}}\mathcal{K}_{ba}^{\rm HO}(\boldsymbol{x},t_2;\boldsymbol{z},s)\delta v_{ba}(z,s)$$
$$\times\,\mathcal{K}_{ba}^{\rm HO}(\boldsymbol{z},s;\boldsymbol{y},t_1)|_{\boldsymbol{x}=\boldsymbol{y}=0} \tag{C.28}$$
$$= \frac{\alpha_s}{\pi^2}P_{ba}(z)\mathrm{Re}\int_0^L \mathrm{d}s\int_{\boldsymbol{u}}\frac{1}{\boldsymbol{u}^2}\delta v_{ba}(\boldsymbol{u},s)\mathrm{e}^{-k^2(s)\boldsymbol{u}^2}\,, \tag{C.29}$$

where we have defined

$$k^2(s) = i\frac{\omega\Omega}{2}\left[\cot(\Omega s) - \tan(\Omega(L-s))\right], \tag{C.30}$$
$$\delta v_{ba}(\boldsymbol{x},t) = \frac{\hat{q}_0}{4}\boldsymbol{x}^2 C_{ba}(z)\ln\frac{1}{\boldsymbol{x}^2Q^2}\,, \tag{C.31}$$

and we have used $\Omega = \sqrt{\hat{q}_{ba}/(2i\omega)}$, $C_{ba}(z) = \frac{C_{cba}}{2C_A} + \frac{C_{acb}}{2C_A}z^2 + \frac{C_{bac}}{2C_A}(1-z)^2$ and some sub-leading terms have already been included in $\hat{q}_{ba}$. The integral over the transverse position can be done

$$\int_{\boldsymbol{u}}\frac{1}{\boldsymbol{u}^2}\delta v(\boldsymbol{u},s)\mathrm{e}^{-k^2(s)\boldsymbol{u}^2} = \frac{\pi}{4}\hat{q}_0\frac{1}{-k^2(s)}\left(\gamma_{\rm E} + \ln\frac{-k^2(s)}{Q^2}\right). \tag{C.32}$$

[23] Our definition includes sub-leading $\sim z^2\ln z^2$ terms to the HO term and thus these terms get resummed. This should make the IOE expansion converge faster. The leading form without these terms would look like

$$\hat{q}_{ba}(z,t) = \hat{q}_0(t)\left[\frac{C_{cba}}{2C_A} + \frac{C_{acb}}{2C_A}z^2 + \frac{C_{bac}}{2C_A}(1-z)^2\right]\ln\frac{Q^2}{\mu_*^2}\,.$$

JHEP02(2023)156

Collecting all terms, we get

$$\frac{\mathrm{d}I^{\mathrm{NHO}}_{ba}}{\mathrm{d}z} = \frac{\alpha_s}{2\pi} P_{ba}(z)\hat{q}_0 C_{ba}(z)\mathrm{Re}\int_0^L \frac{\mathrm{d}s}{-k^2(s)}\left(\gamma_{\mathrm{E}} + \ln\frac{-k^2(s)}{Q^2}\right) \tag{C.33}$$

$$\approx \begin{cases} \frac{\alpha_s}{\pi} P_{ba}(z) C_{ba}(z) \frac{\hat{q}_0}{\hat{q}_{ba}} \sqrt{\frac{\omega_c}{2\omega}} \left\{ \sqrt{\frac{\omega}{2\omega_c}} \left(\frac{\pi^2}{12} \tanh\left(\sqrt{\frac{\omega_c}{2\omega}}\right) - 2\ln 2 \right) \right. \\ \left. +1 + \tanh\left(\sqrt{\frac{\omega_c}{2\omega}}\right) \left[\gamma_E - 1 + \frac{\pi}{4} + \ln\left(\frac{\sqrt{\omega\hat{q}_{ba}}}{\sqrt{2}Q^2}\right)\right]\right\}, & \text{for } \omega \ll \omega_c\,, \\ \frac{\alpha_s}{2} P_{ba}(z) C_{ba}(z) \frac{\hat{q}_0 L^2}{2\omega} \left[1 + \frac{2}{3\pi}\frac{\hat{q}_{ba}L^2}{2\omega}\left(2\gamma_E - \frac{7}{12} + \ln\frac{\omega}{2LQ^2}\right)\right], & \text{for } \omega \gg \omega_c\,. \end{cases}$$

The rate can be given explicitly in the soft and hard limit

$$\frac{\mathrm{d}I^{\mathrm{NHO}}_{ba}}{\mathrm{d}z\mathrm{d}L} = \frac{\partial}{\partial L}\frac{\mathrm{d}I^{\mathrm{NHO}}_{ba}}{\mathrm{d}z} \tag{C.34}$$

$$\approx \begin{cases} \frac{\alpha_s}{\pi} P_{ba}(z) C_{ba}(z) \frac{\hat{q}_0}{\hat{q}_{ba}} \frac{1}{2L} \sqrt{\frac{\omega_c}{2\omega}} \mathrm{sech}^2\left(\sqrt{\frac{\omega_c}{2\omega}}\right) \left\{ 1 + \frac{\pi^2}{12} + \cosh\left(\sqrt{\frac{2\omega_c}{\omega}}\right) \right. \\ \left. + \left(\sqrt{\frac{2\omega_c}{\omega}} + \sinh\left(\sqrt{\frac{2\omega_c}{\omega}}\right)\right)\left(\gamma_E - 1 + \frac{\pi}{4} + \ln\left(\frac{\sqrt{\omega\hat{q}_{ba}}}{\sqrt{2}Q^2}\right)\right)\right\}, & \text{for } \omega \ll \omega_c\,, \\ \frac{\alpha_s}{2} P_{ba}(z) C_{ba}(z) \frac{\hat{q}_0 L}{\omega} \left[1 + \frac{2}{3\pi}\frac{\hat{q}_{ba}L^2}{2\omega}\left(4\gamma_E - \frac{5}{3} + 2\ln\left(\frac{\omega}{2LQ^2}\right)\right)\right], & \text{for } \omega \gg \omega_c\,. \end{cases}$$

The approximated formulas capture the exact formulas up to a few percent deviances and are therefore suitable for numeric implementation, as they do not contain any integrals.

Matching scale. We already introduced the ω dependence of $Q^2(\omega)$, that relied on the soft limit ($\omega \ll \omega_c$) of the spectrum (where the finite-z correction disappears), and thus we use the same definition as in eq. (3.44). In the numerical implementation, we set $\hat{q} = \max(\hat{q}_0, \hat{q}(z))$ and $Q^2 = \max(\sqrt{e}\mu_*^2, Q^2(z))$. This will only become relevant if $L < \lambda$ and $\bar{\omega}_c < \omega < \omega_{\mathrm{BH}}$ which is a small corner of the phase space, where instead of the IOE one should use the OE. We showed numerically that using the IOE with the frozen matching scale or using the OE for $L < \lambda$ does not matter, however, the latter would need the introduction of a new smoothing between OE and IOE at $L = \lambda$ that complicates the implementation (see also appendix D).

To summarize this section, the full emission phase space is covered by using a similar formula that was presented in section 3.5

$$\frac{\mathrm{d}I}{\mathrm{d}z} = \begin{cases} \frac{\mathrm{d}I^{\mathrm{ROE}}}{\mathrm{d}z}\,, & \omega < \omega_{\mathrm{tr}}\,, \\ \frac{\mathrm{d}I^{\mathrm{IOE}}}{\mathrm{d}z}\,, & \omega > \omega_{\mathrm{tr}}\,, \end{cases} \tag{C.35}$$

where $\omega_{\mathrm{tr}} = \min(\omega_{\mathrm{BH}}, \bar{\omega}_c)$ and $\omega = z(1-z)E$. The condition on ω comes from the limits calculated in this section. Importantly, the conditions are the same as the ones derived in section 3.

We would like emphasize the $z, 1-z$ symmetry presented in the ω condition. The gluon spectrum is trivially symmetric in $z, 1-z$ as the emitted particles' kinematics is equivalent. The quark spectrum, on the other hand, is strongly asymmetric in z. One can still use the symmetric condition on ω as we saw in the limiting formulas in this section.

JHEP02(2023)156

D Numerical implementation of the evolution equation

The evolution equation is given in eq. (4.8) and can be rewritten by introducing the variable $\xi = \frac{x}{z}$ ($\xi = xz$) in the gain (loss) term

$$\partial_t D(x,t) = \int_x^1 \mathrm{d}\xi\, f(x,\xi,t) D(\xi,t) - D(x,t) \int_0^1 \mathrm{d}\xi\, f(\xi,x,t)\,, \tag{D.1}$$

$$f(x,\xi,t) = \frac{x}{\xi^2} \left.\frac{\mathrm{d}^2 I}{\mathrm{d}z\mathrm{d}t}\right|_{E\mapsto\xi E} \left(z \mapsto \frac{x}{\xi}\right) . \tag{D.2}$$

We implemented eq. (C.14) (as $N = 1$), eq. (C.22) (as $N_r = 1$), and eqs. (C.27) and (C.34) (HO+NHO) in f as

$$\left.\frac{\mathrm{d}^2 I^{\mathrm{med}}}{\mathrm{d}z\mathrm{d}t}\right|_E = (1-S) \left.\frac{\mathrm{d}^2 I^{N_r=1}}{\mathrm{d}z\mathrm{d}t}\right|_E + S \left.\frac{\mathrm{d}^2 I^{\mathrm{IOE}}}{\mathrm{d}z\mathrm{d}t}\right|_E , \tag{D.3}$$

where $\omega_{\mathrm{tr}} = \min(\omega_{\mathrm{BH}}, \bar{\omega}_c)$ and $\omega = z(1-z)E$. The upper formula is not smooth for any finite order of truncation in the transition between the IOE or ROE, and therefore in some cases, we used the switching function $S = \cos\left[\frac{\pi}{2}(1-\alpha)\right]$, with $\alpha = \frac{2\omega-\omega_{\mathrm{tr}}}{3\omega_{\mathrm{rt}}}$ if $\frac{1}{2}\omega_{\mathrm{tr}} < \omega < 2\omega_{\mathrm{tr}}$ to smoothing the transition. The uncertainty introduced by this procedure is smaller than the next higher-order contribution. One can study the matching uncertainty around the BH region by varying ω_{tr} with a factor of 2. To study the matching condition of the IOE, Q can also be varied by a factor of 2 in eq. (3.44) as it was done in ref. [57]. In figure 11 we show the deviation from the numeric solution including both of these variations. Other than the band, figure 11 is equivalent to figure 7, we only use the soft limit ($z \ll 1$). There is a further uncertainty coming from going to one higher order ($N_r = 1 \mapsto 2$ and NHO $\mapsto$ NNHO), that we leave for future studies. We expect this uncertainty, however, to extend the error band in figure 11 up to the numeric solution.

The integrals can then be divided into

$$\mathrm{Gain} = G_{\xi\to x} + G_{\mathrm{reg}} + G_{>} = \left[\int_x^{x+\epsilon} + \int_{x+\epsilon}^{1-\delta} + \int_{1-\delta}^{1}\right] \mathrm{d}\xi\, f(x,\xi,t) D(\xi,t)\,, \tag{D.4}$$

$$\mathrm{Loss} = L_{<} + L_{\mathrm{reg}} + L_{\xi\to x} = -\left[\int_0^{\delta} + \int_{\delta}^{x-\epsilon} + \int_{x-\epsilon}^{x}\right] \mathrm{d}\xi\, f(\xi,x,t) D(x,t)\,. \tag{D.5}$$

All divergences are present in the $\xi \to x$ terms, which cancel exactly and thus the trapezoid rule is used

$$G_{\xi\to x} + L_{\xi\to x} \approx \frac{\epsilon}{2}\left[f(x,x+\epsilon)D(x+\epsilon,\tau) - f(x-\epsilon,x)D(x,\tau)\right] , \tag{D.6}$$

which contributes to the regular part of the integrals. In our implementation $\epsilon = 10^{-6}$, and thus we have $x > \epsilon$. Similarly to vacuum physics, the ϵ cut was necessary to introduce because the soft divergence in the Bethe-Heitler region has to be regulated. The $G_{\mathrm{reg}}, L_{<}, L_{\mathrm{reg}}$ are simple integrals and can be done numerically on a grid. So can $G_{>}$, however, we neglect this latter contribution by using the fact $\lim_{x\to 1} D \to 0$ (the kernel is soft divergent and thus it moves quanta towards $x < 1$).

JHEP02(2023)156

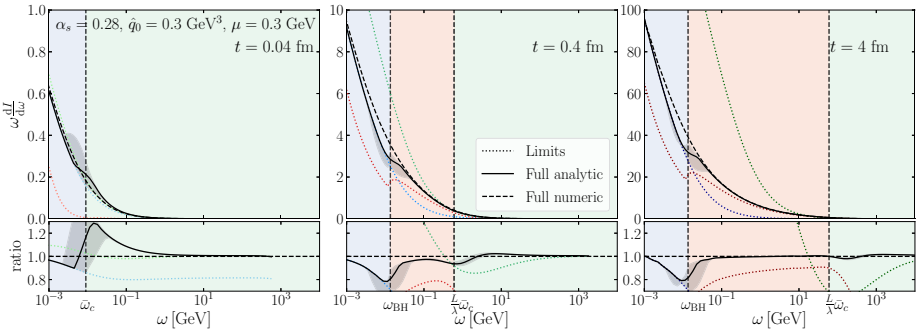


Figure 11. Same as figure 7, on a semilog scale.

E HTL potential

In this appendix, we investigate what happens if employ the HTL potential in the multiple-scattering series. The potential now reads

$$\sigma(\boldsymbol{q}) = N_c n(t) \frac{\mathrm{d}^2\sigma_{\rm el}}{\mathrm{d}^2\boldsymbol{q}} = \frac{4\pi\,\hat{q}_0}{\boldsymbol{q}^2(\boldsymbol{q}^2 + m_D^2)}\,, \tag{E.1}$$

where m_D is the Debye screening mass in a thermal medium with temperature T and $\hat{q}_0 = 4\pi\alpha_s m_D^2 T$ [65]. In this case, we find that

$$\Sigma(\boldsymbol{p}^2) = \frac{\hat{q}_0}{m_D^2} \ln\left(\frac{\boldsymbol{p}^2 + m_D^2}{\boldsymbol{p}^2}\right)\,, \tag{E.2}$$

where we have a logarithmic divergence as $\boldsymbol{p}^2 \to 0$. This implies that the mean free path has to be regulated by an IR cut-off so that $\lambda \sim \Sigma^{-1}(p^2_{\rm min})$.

In this case, the OE remains unmodified because the divergence at $p^2_{\rm min} \to 0$ cancels order by order between the real and virtual contributions. The ROE, however, has to be treated with care when truncated at a finite order since we have to introduce an explicit IR cut-off in order to define the elastic Sudakov factor. Hence, the modified Sudakov reads

$$\Delta(t, t_0) = \mathrm{e}^{-\Sigma_{\rm reg}(t - t_0)}\,, \tag{E.3}$$

in a static medium, where $\Sigma_{\rm reg} \equiv \Sigma(p^2_{\rm min})$ with $p^2_{\rm min}$ and unknown IR regulator. Clearly, an all-order resummation of the ROE series would remove the spurious IR dependence.

The role of the medium potential in the IOE was clearly elucidated in [55], where the information about the scattering potential is fully contained in the definition of μ_*^2. We refer the reader to this paper for an exhaustive discussion.

In order to clarify what modifications arise in the OE and ROE, we compute the respective first-order terms of the expansions, i.e. $N = 1$ and $N_r = 1$, explicitly here. In the OE, we find

$$\omega \frac{\mathrm{d}I^{N=1}}{\mathrm{d}\omega} = 2\bar{\alpha}\,\frac{\hat{q}_0 L}{m_D^2} \int_0^\infty \mathrm{d}u\, \ln\left(\frac{u+y}{u}\right) \left[-\mathrm{Im}\, T(u)\right]\,, \tag{E.4}$$

JHEP02(2023)156

where now we defined $y = \bar{\omega}_c/\omega$ with $\bar{\omega}_c = m_D^2 L/2$. The prefactor contains $\hat{q}_0/m_D^2$, which is similar to the inverse mean free path (in fact, in the GW model it would be exactly equal to λ^{-1}), with a missing logarithmic factor. Using (3.30), we can immediately extract the limiting behaviors, which yield

$$\omega \frac{\mathrm{d}I^{N=1}}{\mathrm{d}\omega} = 2\bar{\alpha}\,\frac{\hat{q}_0 L}{m_D^2} \begin{cases} \ln(\frac{\bar{\omega}_c}{\omega})\left[-1+\gamma_E+\ln\left(\frac{\bar{\omega}_c}{\omega}\right)\right] & \text{for } \omega \ll \bar{\omega}_c \\ \frac{\pi}{4}\frac{\bar{\omega}_c}{\omega} & \text{for } \omega \gg \bar{\omega}_c\,. \end{cases} \tag{E.5}$$

As expected, in this limit there is no sensitivity to the mean free path *per se*. However, compared to the $N = 1$ result in (3.16), there is an additional logarithmic enhancement $\sim \ln \bar{\omega}_c/\omega$ in the soft limit. Comparing to eq. (E.2), we see that this logarithm can be absorbed into the prefactor $\sim \hat{q}_0/m_D^2$ to recreate an effective, regularized mean free path in the HTL theory, i.e.

$$\lambda_{\mathrm{reg}}^{-1}\Big|_{L\ll\lambda} = \frac{\hat{q}_0}{m_D^2}\ln\left(\frac{\bar{\omega}_c}{\omega}\right)\,. \tag{E.6}$$

This regularization follows also from the discussion in refs. [36, 46]. With this modification, the soft limit for low medium opacity is equivalent in the GW and HTL theory.

In the hard limit, $\omega \gg \bar{\omega}_c$, there is no additional logarithmic enhancement in the HTL compared to the GW theory, see (3.16), and the mean free path is simply $\lambda_0^{-1} = \hat{q}_0/m_D^2$. Apart from this subtlety, the expressions are again equivalent.

Let us now turn to the ROE resummation which is valid in dilute (small) media or in the soft limit for dense or large media. At first order $N_r = 1$, we find now

$$\omega \frac{\mathrm{d}I^{N_r=1}}{\mathrm{d}\omega} = 2\bar{\alpha}\,\frac{\hat{q}_0 L}{m_D^2}\int_0^\infty \mathrm{d}u\,\ln\left(\frac{u+y}{u}\right)\left[-\mathrm{Im}\,T(u-i\chi_{\mathrm{reg}})\right], \tag{E.7}$$

where the opacity is $\chi_{\mathrm{reg}} = \Sigma_{\mathrm{reg}} L$ and $\Sigma_{\mathrm{reg}} = \int_{\boldsymbol{q}} \sigma(\boldsymbol{q})\Theta(\boldsymbol{q}^2 - q_{\mathrm{min}}^2)$ is the regularized inverse mean free path.[24] As discussed at length in section 3.3, at small opacities the ROE is equivalent order by order to the OE. This was discussed in the paragraphs above. For large opacities, $\chi_{\mathrm{reg}} \gg 1$, we use (3.32) to solve the integral analytically. As discussed below eq. (3.32), the expressions permits a transmutation of the relevant scale from $\bar{\omega}_c$ to ω_{BH}, where now $\omega_{\mathrm{BH}} = m_D^2/(2\Sigma_{\mathrm{reg}})$. It is then straightforward to extract the following limiting behavior,

$$\omega \frac{\mathrm{d}I^{N_r=1}}{\mathrm{d}\omega} = 2\bar{\alpha}\,\frac{\hat{q}_0 L}{m_D^2} \begin{cases} \frac{1}{24}\left[5\pi^2 + 12\ln^2\left(\frac{\omega_{\mathrm{BH}}}{\omega}\right)\right] & \text{for } \omega \ll \omega_{\mathrm{BH}} \\ \frac{\pi}{2}\frac{\omega_{\mathrm{BH}}}{\omega} & \text{for } \omega \gg \omega_{\mathrm{BH}}\,. \end{cases} \tag{E.8}$$

The soft and hard limits have again subtly different characteristics. In the former case, we again observe a double-logarithmic enhancement, similar to the soft limit in dilute media (E.5) and stronger than the single-logarithmic behavior in the GW model, see eq. (3.34). We could again absorb one of these factors in an effective mean free path, by defining

$$\lambda_{\mathrm{reg}}^{-1}\Big|_{L\gg\lambda} = \frac{\hat{q}_0}{m_D^2}\ln\left(\frac{\omega_{\mathrm{BH}}}{\omega}\right)\,, \tag{E.9}$$

[24] For consistency, compared to the "regularized expansion in eq. (3.26), one should also include an IR regulator $\sim q_{\mathrm{min}}^2$ in the lower limit of the integral. We will neglect this subtlety for now.

JHEP02(2023)156

which demonstrates once more the transmutation of relevant scales. Note that the spurious IR regulator $q^2_{\rm min}$ appears now on the level of $\sim \ln\ln(q^2_{\rm min})$. In the hard limit, the result is again equivalent to the GW model, see (3.34), by identifying the "bare" mean free path and rescaling the Bethe-Heitler energy $\omega_{\rm BH}$. We recall that in this limit, the ROE opacity breaks down and should be replaced by the IOE resummation.

JHEP02(2023)156

References

[1] R.K. Ellis, W.J. Stirling and B.R. Webber, *QCD and collider physics*, Cambridge monographs on particle physics, nuclear physics, and cosmology **8**, Cambridge University Press (2011) [DOI] [INSPIRE].

[2] Y.L. Dokshitzer, V.A. Khoze, A.H. Mueller and S.I. Troian, *Basics of perturbative QCD*, Editions Frontieres (1991) [INSPIRE].

[3] A.J. Larkoski, I. Moult and B. Nachman, *Jet Substructure at the Large Hadron Collider: A Review of Recent Advances in Theory and Machine Learning*, *Phys. Rept.* **841** (2020) 1 [arXiv:1709.04464] [INSPIRE].

[4] S. Marzani, G. Soyez and M. Spannowsky, *Looking inside jets: an introduction to jet substructure and boosted-object phenomenology*, Lecture Notes in Physics **958**, Springer (2019) [DOI] [arXiv:1901.10342] [INSPIRE].

[5] M. Dasgupta, F.A. Dreyer, K. Hamilton, P.F. Monni, G.P. Salam and G. Soyez, *Parton showers beyond leading logarithmic accuracy*, *Phys. Rev. Lett.* **125** (2020) 052002 [arXiv:2002.11114] [INSPIRE].

[6] M. Gyulassy and M. Plumer, *Jet Quenching in Dense Matter*, *Phys. Lett. B* **243** (1990) 432 [INSPIRE].

[7] M. Gyulassy, I. Vitev, X.-N. Wang and B.-W. Zhang, *Jet quenching and radiative energy loss in dense nuclear matter*, nucl-th/0302077 [INSPIRE].

[8] S. Peigne and A.V. Smilga, *Energy losses in a hot plasma revisited*, *Phys. Usp.* **52** (2009) 659 [arXiv:0810.5702] [INSPIRE].

[9] Y. Mehtar-Tani, J.G. Milhano and K. Tywoniuk, *Jet physics in heavy-ion collisions*, *Int. J. Mod. Phys. A* **28** (2013) 1340013 [arXiv:1302.2579] [INSPIRE].

[10] J. Ghiglieri and D. Teaney, *Parton energy loss and momentum broadening at NLO in high temperature QCD plasmas*, *Int. J. Mod. Phys. E* **24** (2015) 1530013 [arXiv:1502.03730] [INSPIRE].

[11] H.A. Andrews et al., *Novel tools and observables for jet physics in heavy-ion collisions*, *J. Phys. G* **47** (2020) 065102 [arXiv:1808.03689] [INSPIRE].

[12] N. Armesto and E. Scomparin, *Heavy-ion collisions at the Large Hadron Collider: a review of the results from Run 1*, *Eur. Phys. J. Plus* **131** (2016) 52 [arXiv:1511.02151] [INSPIRE].

[13] M. Connors, C. Nattrass, R. Reed and S. Salur, *Jet measurements in heavy ion physics*, *Rev. Mod. Phys.* **90** (2018) 025005 [arXiv:1705.01974] [INSPIRE].

[14] R. Kogler et al., *Jet Substructure at the Large Hadron Collider: Experimental Review*, *Rev. Mod. Phys.* **91** (2019) 045003 [arXiv:1803.06991] [INSPIRE].

[15] L. Cunqueiro and A.M. Sickles, *Studying the QGP with Jets at the LHC and RHIC*, *Prog. Part. Nucl. Phys.* **124** (2022) 103940 [arXiv:2110.14490] [INSPIRE].

[16] L. Apolinário, Y.-J. Lee and M. Winn, *Heavy quarks and jets as probes of the QGP*, *Prog. Part. Nucl. Phys.* **127** (2022) 103990 [arXiv:2203.16352] [INSPIRE].

[17] Y.-T. Chien and I. Vitev, *Towards the understanding of jet shapes and cross sections in heavy ion collisions using soft-collinear effective theory*, *JHEP* **05** (2016) 023 [arXiv:1509.07257] [INSPIRE].

[18] Y. Mehtar-Tani and K. Tywoniuk, *Groomed jets in heavy-ion collisions: sensitivity to medium-induced bremsstrahlung*, *JHEP* **04** (2017) 125 [arXiv:1610.08930] [INSPIRE].

[19] R. Kunnawalkam Elayavalli and K.C. Zapp, *Medium response in JEWEL and its impact on jet shape observables in heavy ion collisions*, *JHEP* **07** (2017) 141 [arXiv:1707.01539] [INSPIRE].

[20] P. Caucal, E. Iancu and G. Soyez, *Deciphering the z_g distribution in ultrarelativistic heavy ion collisions*, *JHEP* **10** (2019) 273 [arXiv:1907.04866] [INSPIRE].

[21] J. Casalderrey-Solana, G. Milhano, D. Pablos and K. Rajagopal, *Modification of Jet Substructure in Heavy Ion Collisions as a Probe of the Resolution Length of Quark-Gluon Plasma*, *JHEP* **01** (2020) 044 [arXiv:1907.11248] [INSPIRE].

[22] F. Ringer, B.-W. Xiao and F. Yuan, *Can we observe jet P_T-broadening in heavy-ion collisions at the LHC?*, *Phys. Lett. B* **808** (2020) 135634 [arXiv:1907.12541] [INSPIRE].

[23] P. Caucal, E. Iancu, A.H. Mueller and G. Soyez, *Nuclear modification factors for jet fragmentation*, *JHEP* **10** (2020) 204 [arXiv:2005.05852] [INSPIRE].

[24] P. Caucal, A. Soto-Ontoso and A. Takacs, *Dynamically groomed jet radius in heavy-ion collisions*, *Phys. Rev. D* **105** (2022) 114046 [arXiv:2111.14768] [INSPIRE].

[25] Y. Mehtar-Tani, D. Pablos and K. Tywoniuk, *Cone-Size Dependence of Jet Suppression in Heavy-Ion Collisions*, *Phys. Rev. Lett.* **127** (2021) 252301 [arXiv:2101.01742] [INSPIRE].

[26] A. Takacs and K. Tywoniuk, *Quenching effects in the cumulative jet spectrum*, *JHEP* **10** (2021) 038 [arXiv:2103.14676] [INSPIRE].

[27] M. Attems, J. Brewer, G.M. Innocenti, A. Mazeliauskas, S. Park, W. van der Schee et al., *The medium-modified $g \to c\bar{c}$ splitting function in the BDMPS-Z formalism*, *JHEP* **01** (2023) 080 [arXiv:2203.11241] [INSPIRE].

[28] J. Casalderrey-Solana, Y. Mehtar-Tani, C.A. Salgado and K. Tywoniuk, *New picture of jet quenching dictated by color coherence*, *Phys. Lett. B* **725** (2013) 357 [arXiv:1210.7765] [INSPIRE].

[29] P. Caucal, E. Iancu, A.H. Mueller and G. Soyez, *Vacuum-like jet fragmentation in a dense QCD medium*, *Phys. Rev. Lett.* **120** (2018) 232001 [arXiv:1801.09703] [INSPIRE].

[30] V. Vaidya, *Effective Field Theory for jet substructure in heavy ion collisions*, *JHEP* **11** (2021) 064 [arXiv:2010.00028] [INSPIRE].

[31] Y. Mehtar-Tani and K. Tywoniuk, *Sudakov suppression of jets in QCD media*, *Phys. Rev. D* **98** (2018) 051501 [arXiv:1707.07361] [INSPIRE].

[32] R. Baier, Y.L. Dokshitzer, A.H. Mueller, S. Peigne and D. Schiff, *Radiative energy loss of high-energy quarks and gluons in a finite volume quark-gluon plasma*, *Nucl. Phys. B* **483** (1997) 291 [hep-ph/9607355] [INSPIRE].

JHEP02(2023)156

[33] B.G. Zakharov, *Radiative energy loss of high-energy quarks in finite size nuclear matter and quark-gluon plasma*, *JETP Lett.* **65** (1997) 615 [hep-ph/9704255] [INSPIRE].

[34] P.B. Arnold, G.D. Moore and L.G. Yaffe, *Photon and gluon emission in relativistic plasmas*, *JHEP* **06** (2002) 030 [hep-ph/0204343] [INSPIRE].

[35] M. Gyulassy, P. Levai and I. Vitev, *Jet quenching in thin quark gluon plasmas I: Formalism*, *Nucl. Phys. B* **571** (2000) 197 [hep-ph/9907461] [INSPIRE].

[36] U.A. Wiedemann, *Gluon radiation off hard quarks in a nuclear environment: Opacity expansion*, *Nucl. Phys. B* **588** (2000) 303 [hep-ph/0005129] [INSPIRE].

[37] M. Djordjevic and M. Gyulassy, *Heavy quark radiative energy loss in QCD matter*, *Nucl. Phys. A* **733** (2004) 265 [nucl-th/0310076] [INSPIRE].

[38] M.D. Sievert and I. Vitev, *Quark branching in QCD matter to any order in opacity beyond the soft gluon emission limit*, *Phys. Rev. D* **98** (2018) 094010 [arXiv:1807.03799] [INSPIRE].

[39] C.A. Salgado and U.A. Wiedemann, *Calculating quenching weights*, *Phys. Rev. D* **68** (2003) 014008 [hep-ph/0302184] [INSPIRE].

[40] N. Armesto, C.A. Salgado and U.A. Wiedemann, *Medium induced gluon radiation off massive quarks fills the dead cone*, *Phys. Rev. D* **69** (2004) 114003 [hep-ph/0312106] [INSPIRE].

[41] P.B. Arnold, *Simple Formula for High-Energy Gluon Bremsstrahlung in a Finite, Expanding Medium*, *Phys. Rev. D* **79** (2009) 065025 [arXiv:0808.2767] [INSPIRE].

[42] P.B. Arnold, *High-energy gluon bremsstrahlung in a finite medium: harmonic oscillator versus single scattering approximation*, *Phys. Rev. D* **80** (2009) 025004 [arXiv:0903.1081] [INSPIRE].

[43] J.-P. Blaizot, F. Domínguez, E. Iancu and Y. Mehtar-Tani, *Medium-induced gluon branching*, *JHEP* **01** (2013) 143 [arXiv:1209.4585] [INSPIRE].

[44] A. Kurkela and U.A. Wiedemann, *Picturing perturbative parton cascades in QCD matter*, *Phys. Lett. B* **740** (2015) 172 [arXiv:1407.0293] [INSPIRE].

[45] F. Domínguez, J.G. Milhano, C.A. Salgado, K. Tywoniuk and V. Vila, *Mapping collinear in-medium parton splittings*, *Eur. Phys. J. C* **80** (2020) 11 [arXiv:1907.03653] [INSPIRE].

[46] C. Andres, F. Domínguez and M. Gonzalez Martinez, *From soft to hard radiation: the role of multiple scatterings in medium-induced gluon emissions*, *JHEP* **03** (2021) 102 [arXiv:2011.06522] [INSPIRE].

[47] B.G. Zakharov, *Radiative parton energy loss and jet quenching in high-energy heavy-ion collisions*, *JETP Lett.* **80** (2004) 617 [hep-ph/0410321] [INSPIRE].

[48] S. Caron-Huot and C. Gale, *Finite-size effects on the radiative energy loss of a fast parton in hot and dense strongly interacting matter*, *Phys. Rev. C* **82** (2010) 064902 [arXiv:1006.2379] [INSPIRE].

[49] X. Feal and R. Vazquez, *Intensity of gluon bremsstrahlung in a finite plasma*, *Phys. Rev. D* **98** (2018) 074029 [arXiv:1811.01591] [INSPIRE].

[50] W. Ke, Y. Xu and S.A. Bass, *Modified Boltzmann approach for modeling the splitting vertices induced by the hot QCD medium in the deep Landau-Pomeranchuk-Migdal region*, *Phys. Rev. C* **100** (2019) 064911 [arXiv:1810.08177] [INSPIRE].

[51] C. Andres, L. Apolinário and F. Domínguez, *Medium-induced gluon radiation with full resummation of multiple scatterings for realistic parton-medium interactions*, *JHEP* **07** (2020) 114 [arXiv:2002.01517] [INSPIRE].

JHEP02(2023)156

[52] S. Schlichting and I. Soudi, *Medium-induced fragmentation and equilibration of highly energetic partons*, *JHEP* **07** (2021) 077 [arXiv:2008.04928] [INSPIRE].

[53] Y. Mehtar-Tani, *Gluon bremsstrahlung in finite media beyond multiple soft scattering approximation*, *JHEP* **07** (2019) 057 [arXiv:1903.00506] [INSPIRE].

[54] Y. Mehtar-Tani and K. Tywoniuk, *Improved opacity expansion for medium-induced parton splitting*, *JHEP* **06** (2020) 187 [arXiv:1910.02032] [INSPIRE].

[55] J.a. Barata and Y. Mehtar-Tani, *Improved opacity expansion at NNLO for medium induced gluon radiation*, *JHEP* **10** (2020) 176 [arXiv:2004.02323] [INSPIRE].

[56] J.a. Barata, Y. Mehtar-Tani, A. Soto-Ontoso and K. Tywoniuk, *Revisiting transverse momentum broadening in dense QCD media*, *Phys. Rev. D* **104** (2021) 054047 [arXiv:2009.13667] [INSPIRE].

[57] J.a. Barata, Y. Mehtar-Tani, A. Soto-Ontoso and K. Tywoniuk, *Medium-induced radiative kernel with the Improved Opacity Expansion*, *JHEP* **09** (2021) 153 [arXiv:2106.07402] [INSPIRE].

[58] P.B. Arnold, G.D. Moore and L.G. Yaffe, *Effective kinetic theory for high temperature gauge theories*, *JHEP* **01** (2003) 030 [hep-ph/0209353] [INSPIRE].

[59] J.-P. Blaizot, F. Domínguez, E. Iancu and Y. Mehtar-Tani, *Probabilistic picture for medium-induced jet evolution*, *JHEP* **06** (2014) 075 [arXiv:1311.5823] [INSPIRE].

[60] Y. Mehtar-Tani, C.A. Salgado and K. Tywoniuk, *Anti-angular ordering of gluon radiation in QCD media*, *Phys. Rev. Lett.* **106** (2011) 122002 [arXiv:1009.2965] [INSPIRE].

[61] Y. Mehtar-Tani, C.A. Salgado and K. Tywoniuk, *The Radiation pattern of a QCD antenna in a dense medium*, *JHEP* **10** (2012) 197 [arXiv:1205.5739] [INSPIRE].

[62] S. Schlichting and I. Soudi, *Splitting rates in QCD plasmas from a nonperturbative determination of the momentum broadening kernel $C(q\perp)$*, *Phys. Rev. D* **105** (2022) 076002 [arXiv:2111.13731] [INSPIRE].

[63] G.D. Moore, S. Schlichting, N. Schlusser and I. Soudi, *Non-perturbative determination of collisional broadening and medium induced radiation in QCD plasmas*, *JHEP* **10** (2021) 059 [arXiv:2105.01679] [INSPIRE].

[64] M. Gyulassy and X.-n. Wang, *Multiple collisions and induced gluon Bremsstrahlung in QCD*, *Nucl. Phys. B* **420** (1994) 583 [nucl-th/9306003] [INSPIRE].

[65] P. Aurenche, F. Gelis and H. Zaraket, *A Simple sum rule for the thermal gluon spectral function and applications*, *JHEP* **05** (2002) 043 [hep-ph/0204146] [INSPIRE].

[66] M. Gyulassy, P. Levai and I. Vitev, *Reaction operator approach to nonAbelian energy loss*, *Nucl. Phys. B* **594** (2001) 371 [nucl-th/0006010] [INSPIRE].

[67] M. Gyulassy, P. Levai and I. Vitev, *NonAbelian energy loss at finite opacity*, *Phys. Rev. Lett.* **85** (2000) 5535 [nucl-th/0005032] [INSPIRE].

[68] C. Andres, F. Domínguez and M. Gonzalez Martinez, private communication (2022).

[69] B.G. Zakharov, *Fully quantum treatment of the Landau-Pomeranchuk-Migdal effect in QED and QCD*, *JETP Lett.* **63** (1996) 952 [hep-ph/9607440] [INSPIRE].

[70] P. Arnold and S. Iqbal, *The LPM effect in sequential bremsstrahlung*, *JHEP* **04** (2015) 070 [arXiv:1501.04964] [INSPIRE].

[71] P. Arnold, T. Gorda and S. Iqbal, *The LPM effect in sequential bremsstrahlung: nearly complete results for QCD*, *JHEP* **11** (2020) 053 [arXiv:2007.15018] [INSPIRE].

JHEP02(2023)156

[72] P. Arnold, T. Gorda and S. Iqbal, *The LPM effect in sequential bremsstrahlung: analytic results for sub-leading (single) logarithms*, *JHEP* **04** (2022) 085 [arXiv:2112.05161] [INSPIRE].

[73] A. Banfi, F.A. Dreyer and P.F. Monni, *Higher-order non-global logarithms from jet calculus*, *JHEP* **03** (2022) 135 [arXiv:2111.02413] [INSPIRE].

[74] J.a. Barata, F. Domínguez, C.A. Salgado and V. Vila, *A modified in-medium evolution equation with color coherence*, *JHEP* **05** (2021) 148 [arXiv:2101.12135] [INSPIRE].

[75] Y. Mehtar-Tani and S. Schlichting, *Universal quark to gluon ratio in medium-induced parton cascade*, *JHEP* **09** (2018) 144 [arXiv:1807.06181] [INSPIRE].

[76] J.-P. Blaizot, E. Iancu and Y. Mehtar-Tani, *Medium-induced QCD cascade: democratic branching and wave turbulence*, *Phys. Rev. Lett.* **111** (2013) 052001 [arXiv:1301.6102] [INSPIRE].

[77] J.-P. Blaizot and Y. Mehtar-Tani, *Energy flow along the medium-induced parton cascade*, *Annals Phys.* **368** (2016) 148 [arXiv:1501.03443] [INSPIRE].

[78] *Medium-induced kernels*, https://github.com/adam-takacs/kernels.git.

[79] B. Schenke, C. Gale and S. Jeon, *MARTINI: An Event generator for relativistic heavy-ion collisions*, *Phys. Rev. C* **80** (2009) 054913 [arXiv:0909.2037] [INSPIRE].

[80] K. Kutak, W. Płaczek and R. Straka, *Solutions of evolution equations for medium-induced QCD cascades*, *Eur. Phys. J. C* **79** (2019) 317 [arXiv:1811.06390] [INSPIRE].

[81] E. Blanco, K. Kutak, W. Płaczek, M. Rohrmoser and R. Straka, *Medium induced QCD cascades: broadening and rescattering during branching*, *JHEP* **04** (2021) 014 [arXiv:2009.03876] [INSPIRE].

[82] A. Idilbi and A. Majumder, *Extending Soft-Collinear-Effective-Theory to describe hard jets in dense QCD media*, *Phys. Rev. D* **80** (2009) 054022 [arXiv:0808.1087] [INSPIRE].

[83] G. Ovanesyan and I. Vitev, *An effective theory for jet propagation in dense QCD matter: jet broadening and medium-induced bremsstrahlung*, *JHEP* **06** (2011) 080 [arXiv:1103.1074] [INSPIRE].

[84] A. Majumder, *Incorporating Space-Time Within Medium-Modified Jet Event Generators*, *Phys. Rev. C* **88** (2013) 014909 [arXiv:1301.5323] [INSPIRE].

[85] X.-f. Guo and X.-N. Wang, *Multiple scattering, parton energy loss and modified fragmentation functions in deeply inelastic e A scattering*, *Phys. Rev. Lett.* **85** (2000) 3591 [hep-ph/0005044] [INSPIRE].

[86] X.-N. Wang and Y. Zhu, *Medium Modification of γ-jets in High-energy Heavy-ion Collisions*, *Phys. Rev. Lett.* **111** (2013) 062301 [arXiv:1302.5874] [INSPIRE].

[87] J. Ghiglieri, G.D. Moore and D. Teaney, *Jet-Medium Interactions at NLO in a Weakly-Coupled Quark-Gluon Plasma*, *JHEP* **03** (2016) 095 [arXiv:1509.07773] [INSPIRE].

[88] S.P. Adhya, C.A. Salgado, M. Spousta and K. Tywoniuk, *Medium-induced cascade in expanding media*, *JHEP* **07** (2020) 150 [arXiv:1911.12193] [INSPIRE].

[89] S.P. Adhya, C.A. Salgado, M. Spousta and K. Tywoniuk, *Multi-partonic medium induced cascades in expanding media*, *Eur. Phys. J. C* **82** (2022) 20 [arXiv:2106.02592] [INSPIRE].

JHEP02(2023)156

Graphic design: Communication Division, UiB / Print: Skipnes Kommunikasjon AS

uib.no

ISBN: 9788230868492 (print)
9788230844595 (PDF)