TV-Stokes And Its Variants For Image Processing

Bin Wu

Thesis for the degree of Philosophiae Doctor (PhD)
University of Bergen, Norway
2021

UNIVERSITY OF BERGEN

TV-Stokes And Its Variants For Image Processing

Bin Wu

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

Date of defense: 30.03.2021

Year: 2021

Title: TV-Stokes And Its Variants For Image Processing

Name: Bin Wu

Print: Skipnes Kommunikasjon / University of Bergen

Preface

This thesis is submitted as a partial fulfillment of the requirements for the degree of Philosophiae Doctor (PhD) at the Universitetet i Bergen. The thesis consists of two parts, Part I and Part II. The Part I reviews the mathematical background for the papers in the part II. The Part II collects the contributed papers covering several topics in image processing, with main focus on the methods based on variational methods and partial differential equations, as well as the underlying geometric relating to the mathematical models.

Acknowledgements

I would like to give my utmost thanks to my supervisors, Talal Rahman and Xue-Cheng Tai. Without your continuous support, inspiration, and guidance, this work would not be accomplished. I also would like to thank my collaborators, Alexander Malyshev and Leszek Marcinkowski, for all the helps and contributions to the papers.

I appreciate valuable communications with Jan Lellmann, Carola Bibiane-Schönlieb, Ke Yin, Andreas Langer, and Jie Qiu.

I am also grateful for my friends and colleagues at the Høgskulen på Vestlandet and the Universitetet i Bergen, especially all the PhD students. Thank you, Erik Eikeland, Murugesan Rasukkannu, and Alexander Selvikvåg Lundervold for your helps.

Last but not least, I would like to thank my family. Without your company and encouragement, I could not go through this journey.

Abstract

The total variational minimization with a Stokes constraint, also known as the TV-Stokes model, has been considered as one of the most successful models in image processing, especially in image restoration and sparse-data-based 3D surface reconstruction. This thesis studies the TV-Stokes model and its existing variants, proposes new and more effective variants of the model and their algorithms applied to some of the most interesting image processing problems.

We first review some of the variational models that already exist, in particular the TV-Stokes model and its variants. Common techniques like the augmented Lagrangian- and the dual formulation, are also introduced. We then present our models as new variants of the TV-Stokes.

The main focus of the work has been on the sparse surface reconstruction of 3D surfaces. A model (WTR) with a vector fidelity, that is the gradient vector fidelity, has been proposed, applying it to both 3D cartoon design and height map reconstruction. The model employs the second order total variation minimization, where the curl-free condition is satisfied automatically. Because the model couples both the height and the gradient vector representing the surface in the same minimization, it constructs the surface correctly. A variant of this model is then introduced, which includes a vector matching term. This matching term gives the model capability to accurately represent the shape of a geometry in the reconstruction. Experiments show a significant improvement over the state-of-the-art models, such as the TV model, higher order TV models, and the anisotropic third order regularization model, when applied to some general applications.

In another work, the thesis generalizes the TV-Stokes model from two dimensions to an arbitrary number of dimensions, introducing a convenient form for the constraint in order it to be extended to higher dimensions.

The thesis explores also the idea of feature accumulation through iterative regularization in another work, introducing a Richardson like iteration for the TV-Stokes. This is then followed by a more general model, a combined model, based on the modified variant of the TV-stokes. The resulting model is found to be equivalent to the well known TGV model.

The thesis introduces some interesting numerical strategies for the solution of the TV-Stokes model and its variants. Higher order PDEs are turned into inhomogeneous modified Helmholtz equations through transformations. These equations are then solved using the preconditioned conjugate gradients method or the fast Fourier transformation. The thesis proposes a simple but quite general approach to finding closed form solutions to a general L^1 minimization problem, and applies it to design algorithms for our models.

List of Papers (In chronological order)

Paper A **Bin Wu**, Talal Rahman, and Xue-Cheng Tai, *Sparse-Data Based 3D Surface Reconstruction for Cartoon and Map*, Imaging, Vision and Learning Based on Optimization and PDEs, Springer, Cham, (2016): 47.

Paper B **Bin Wu**, Xue-Cheng Tai, and Talal Rahman, *Sparse-Data Based 3D Surface Reconstruction with Vector Matching*. Submitted. arXiv:2009.12994.

Paper C **Bin Wu**, Xue-Cheng Tai, and Talal Rahman, *Multidimensional TV-Stokes for Image Processing*. Submitted. arXiv:2009.11971.

Paper D **Bin Wu**, Leszek Marcinkowski, Xue-Cheng Tai, and Talal Rahman, *Iterative Regularization Algorithms for Image Denoising with the TV-Stokes Model*. Submitted. arXiv:2009.11976.

Paper E **Bin Wu**, Xue-Cheng Tai, and Talal Rahman, *Alternating Minimization for a Single Step TV-Stokes Model for Image Denoising*. Submitted. arXiv:2009.11973.

Contents

Part I

Summary

Chapter 1

Background

I hate TV. I hate it as much as peanuts.
But I can't stop eating peanuts.

Orson Welles (1956)

In this chapter, we present a brief review of methods based on the total variation (TV) minimization as well as their applications to image processing. The focus of this presentation is on some representative numerical approaches to solve the associated models.

With the world getting more and more digital, our daily lives become more and more dependent on digital image processing, finding its applications in almost everywhere, cf. [1, 8, 14, 32, 33]. There are different ways to handle image processing problem, cf. [8], we only consider the variational approach in this thesis.

1.1 Variational Image Processing

Image processing tasks are often defined as inverse problems, cf. [1, 8, 14, 32, 33]. These problems are often ill-posed, cf. [12, 21]. Consider the task of image restoration: Consider a noisy image $I_0 : \Omega \mapsto \mathbb{R}$, where Ω is a bounded open subset of $\mathbb{R}^2$, the problem is to find a decomposition such that $I_0 = I + v$, where I is the true signal and v is the noise. Finding such a decomposition can be formulated as an optimization problem. The simplest of all is the problem of least squares, in other words, is to find the minimizer, in the L^2 space, i.e.

$$I = \arg\min_{v} \|I - I_0\|^2_{(\Omega)}.$$

The notation $\|\cdot\|$, in this chapter, refers to the L^2 norm. This model, however, only works when we know the structure of u otherwise there is only a trivial solution $u = f$. It is obvious that, without sufficient priori, to find such a decomposition is an ill-posed inverse problem, cf. e.g., [1, 8]. A regularizer, which provides somehow the priori, is thus necessary. The Tikhonov regularizer is among the first and most common regularizers used for this problem. In general, we define a regularizer as

$$J_p(u) := |\nabla u|^p_{(\Omega)}.$$

The Tikhonov regularizer is the case where $p = 2$. For models with this regularizer, it is difficult to preserve edges (discontinuity along lines) while smoothing noise. Rudin, Osher, and Fatemi proposed a model, commonly known as the ROF model, with $p = 1$ for the regularizer, cf. [31]. This model, as compared to the Tikhonov regulatizer, preserves irregularities better, and has been widely used in diverse practical problems. The model has inspired and generated numerous variants, cf. [1, 2, 4, 6–11, 17, 18, 23, 30, 33, 36, 39–41, 43, 44]. For the sake of simplicity, we denote $J_1(\cdot)$ for $p = 1$, as $J(\cdot)$, which is also known as the total variation regularizer.

The ROF model is known to preserve edges, however, it suffers a staircase effect which makes the restored image patternized. Higher order models have been developed to overcome this problem, e.g. the Lysaker-Osher-Tai (LOT) model, cf. [25], and the TV-Stokes model, cf. [30].

The ROF model for denoising has the form as follows.

$$\min_I \left\{ \alpha |\nabla I|_{(\Omega)} + \theta \|I - I^o\|^2_{(\Omega)} \right\}. \tag{1.1}$$

Here, α and θ are non-negative scalar parameters; I^o is the known data, as for instance representing the observed image, and I is a scalar-valued function, both defined on Ω. The notations $|\cdot|$ and $\|\cdot\|$ represent L^1 norm and L^2 norm, respectively.

To solve the ROF model, one derives the optimality condition, i.e., the Euler-Lagrange equation (ELE). By the simple calculus of variational, the ELE, a partial differential equation (PDE), is derived as follows.

$$-\frac{\alpha}{2\theta} \nabla \cdot \frac{\nabla I}{|\nabla I|} = I^o - I. \tag{1.2}$$

The difficulty in solving the ROF model is associated with the non-differentiability of the L^1 term. There are several approaches that avoid the problem, we list two of them here.

1.2 Solution techniques for the simplest TV minimization

1.2.1 The Augmented Lagrangian Method

The idea of the augmented Lagrangian method used here in solving the above ROF model is turning the nondifferentiable PDE to a set of solvable equations, by introducing an auxiliary variable, that is $\mathbf{P} := \nabla I$. The primarily unconstrained minimization problem is thus turned into a constrained minimization problem of the following form.

$$\min_{I,\mathbf{P}} \left\{ \alpha |\mathbf{P}|_{(\Omega)} + \theta \|I - I^o\|^2_{(\Omega)} \right\}, \tag{1.3}$$

subject to $\mathbf{P} := \nabla I$. Adding the condition to object functional with one more augmented quadratic term, we obtained the following equivalent saddle point problem.

$$\min_{I,\mathbf{P}} \max_{\mathbf{\Lambda}} \left\{ \alpha|\mathbf{P}| + \theta\|I - I^o\|^2 + \mathbf{\Lambda}\cdot\Big(\mathbf{P} - \nabla I\Big) + \frac{c_{\mathbf{P}}}{2}\|\mathbf{P} - \nabla I\|^2 \right\}, \tag{1.4}$$

where $\mathbf{\Lambda}$ is the Lagrange multiplier and $c_{\mathbf{P}}$ is a non-negative parameter. This saddle point problem is solved alternatingly with respect to its three subproblems, cf. the follows.

$$\begin{aligned} &\min_{I} \left\{ \theta\|I - I^o\|^2 - \mathbf{\Lambda}\cdot\nabla I + \frac{c_{\mathbf{P}}}{2}\|\mathbf{P} - \nabla I\|^2 \right\}; \\ &\min_{\mathbf{P}} \left\{ \alpha|\mathbf{P}| + \mathbf{\Lambda}\cdot\mathbf{P} + \frac{c_{\mathbf{P}}}{2}\|\mathbf{P} - \nabla I\|^2 \right\}; \\ &\max_{\mathbf{\Lambda}} \left\{ \mathbf{\Lambda}\cdot\Big(\mathbf{P} - \nabla I\Big) \right\}. \end{aligned} \tag{1.5}$$

The first subproblem concerning I is solvable via the fast Fourier transform. The second subproblem about $\mathbf{P}$ has a closed form solution. A simple steepest ascent method is usually applied on the last maximizing problem with step length $c_{\mathbf{P}}$.

1.2.2 The Dual Method

The idea of the dual method is to turn a nondifferentiable problem into a differentiable one in the weak sense, by using the dual formulation of the total variation term, i.e.

$$|\nabla I|_{(\Omega)} = \max_{\|\mathbf{p}\|_\infty \le 1} \langle I, \nabla\cdot\mathbf{p}\rangle_{(\Omega)} \tag{1.6}$$

where Ω is an open bounded domain, $\mathbf{p} = (p_1, p_2)$ is a dual variable subject to $\|\mathbf{p}\|_\infty \le 1$. The ROF model, refer to (1.1), reduces to the equivalent min-max formulation

$$\min_{I} \max_{\|\mathbf{p}\|_\infty \le 1} \alpha\langle I, \nabla\cdot\mathbf{p}\rangle_{(\Omega)} + \theta\|I - I^o\|^2_{(\Omega)}. \tag{1.7}$$

The generalized mini-max theorem justifies the equality

$$\max_{\|\mathbf{p}\|_\infty \le 1} \min_{I} \alpha\langle I, \nabla\cdot\mathbf{p}\rangle_{(\Omega)} + \theta\|I - I^o\|^2_{(\Omega)}. \tag{1.8}$$

In spite of a possible nonuniqueness of $\mathbf{p}$, the solution I obtained by means of (1.8) is unique. The unique solution is obtained as follows

$$I = I^o - \frac{\alpha}{2\theta}\nabla\cdot\mathbf{p}. \tag{1.9}$$

Under the condition (1.9), we arrive at the minimum distance problem

$$\min_{\|\mathbf{p}\|_\infty \le 1} \left\| I^o - \frac{\alpha}{2\theta}\nabla\cdot\mathbf{p} \right\|^2. \tag{1.10}$$

As in [5] the Karush-Kuhn-Tucker conditions for (1.10) are given by the equation

$$\nabla\left(\frac{2\theta}{\alpha}I^o-\nabla\cdot\mathbf{p}\right)+\left|\nabla\left(\frac{2\theta}{\alpha}I^o-\nabla\cdot\mathbf{p}\right)\right|\cdot\mathbf{p}=0, \tag{1.11}$$

where the dot signifies the entry wise product. Solution of (1.10) is approximated by the semi-implicit iteration

$$\begin{aligned}&\mathbf{p}^0=\mathbf{0},\\&\mathbf{p}^{k+1}=\frac{\mathbf{p}^k-\tau\nabla\left(\frac{2\theta}{\alpha}I^o-\nabla\cdot\mathbf{p}\right)}{1+\tau\left|\nabla\left(\frac{2\theta}{\alpha}I^o-\nabla\cdot\mathbf{p}\right)\right|}.\end{aligned} \tag{1.12}$$

1.3 The classical TV-Stokes Model [19, 30]

In this section, we present the classical TV-Stokes model, which was originally applied to image denoising and -inpainting; cf., [9–11, 20, 26, 30, 35, 42].

Definition 1 (Denoising)**.** *In image processing, denoising refers to the process of removing noise from an image.*

Definition 2 (Inpainting)**.** *In image processing, inpainting refers to the process of filling-in missing information into an image.*

The classical TV-Stokes model is a two-step model, one minimizing in each step, the first step consists of smoothing of the tangent vector field under the condition of divergence-free, the second step consists of reconstructing the image incorporating both an intensity constraint and a vector constraint from the previous step. The model is formulated in two steps as follows,

$$\min_{\substack{\boldsymbol{\tau}\\ \nabla\cdot\boldsymbol{\tau}=0}}\left\{\alpha_1|\nabla\boldsymbol{\tau}|_{(\Omega)}+\theta_1\|\boldsymbol{\tau}-\nabla^\perp I^o\|^2_{(\Gamma)}\right\}, \tag{1.13}$$

$$\min_{I}\left\{\alpha_2|\nabla I|_{(\Omega)}-\beta_2\frac{\boldsymbol{\tau}^\perp}{|\boldsymbol{\tau}^\perp|}\cdot\nabla I_{(\Omega)}+\theta_2\|I-I^o\|^2_{(\Sigma)}\right\}. \tag{1.14}$$

Here, $\alpha_1,\alpha_2,\theta_1,\theta_2,\beta_2\in\mathbb{R}^+\cup\{0\}$ are non-negative scalar parameters; $I^o\in\mathbb{R}(\Sigma)$ is the known data, e.g., the observed image or intensity; $I\in BV(\mathbb{R},\Omega)$ and $\boldsymbol{\tau}\in BV(\mathbb{R}^2,\Omega)$ are scalar- and vector-valued functions, respectively. Γ and Σ are two domains affiliated with the known information, e.g., intensities and vectors. Ω is the computing domain where our main interest is on. Γ and Σ are normally the subsets of Ω, $\Gamma,\Sigma\subset\Omega\subset\mathbb{R}$. The notations $|\cdot|$ and $\|\cdot\|$ represent L^1 norm and L^2 norm, respectively, if there are no specific subscripts. Note that $\nabla^\perp I^o$ for some cases, e.g., sparse-data-based reconstructions, is not calculable. Instead of calculating the value of ∇I^o from I^o, it is often given manually for such cases.

The TV-Stokes [35] was originally applied to inpainting problem, inspired by the similarity that isophote lines of an image have with the streamlines in a fluid flow, cf. e.g. [3]. In this model, the information propagates into the inpainting domain along isophote lines guided by the tangent vectors $\boldsymbol{\tau}$, by solving the minimization problem (1.14). These vectors are obtained via the optimization step of (1.13), imposing a divergence-free condition, which is also known as the incompressibility condition in the Stokes equation.

1.4 Variants of the TV-Stokes - I

In this section, we present variants of the TV-Stokes model that existed before this thesis; the modified-TV-Stokes model [17, 23], the TV-Curl-free model [17], the modified-TV-Curl-free model [16], the TV-H^1-Curl-free model [15].

1.4.1 The Modified-TV-Stokes Model [23]

The model was applied to both the problem of inpainting and the problem of denoising problem, and contains the same smoothing of the tangent vectors step as that of the classical TV-Stokes model. The difference between the classical TV-Stokes model and the modified-TV-Stokes model lies in the second step. Instead of using the orientation matching term $(\beta_2 \frac{\boldsymbol{\tau}^\perp}{|\boldsymbol{\tau}^\perp|} \cdot \nabla I_{(\Omega)})$, a regularizer containing vector information involves, that is $\alpha_2 \left|\nabla I - \boldsymbol{\tau}^\perp\right|_{(\Omega)}$. The model refers to (1.15) and (1.16).

$$\min_{\substack{\boldsymbol{\tau} \\ \nabla \cdot \boldsymbol{\tau} = 0}} \left\{ \alpha_1 |\nabla \boldsymbol{\tau}|_{(\Omega)} + \theta_1 \|\boldsymbol{\tau} - \nabla^\perp I^o\|^2_{(\Gamma)} \right\}, \tag{1.15}$$

$$\min_{I} \left\{ \alpha_2 |\nabla I - \boldsymbol{\tau}^\perp|_{(\Omega)} + \theta_2 \|I - I^o\|^2_{(\Sigma)} \right\}. \tag{1.16}$$

Here, we use the same notations as those used in the previous section.

1.4.2 The Generalized TV-Stokes Model [17]

A generalization of the classical TV-Stokes and the modified TV-Stokes has been proposed in [17]. Similar to the modified-TV-Stokes model, instead of using the orientation matching term $\beta_2 \frac{\mathbf{n}}{|\mathbf{n}|} \cdot \nabla I_{(\Omega)}$, a regularizer containing vector information involves, that is $\alpha_2 \left|\nabla I - \mathbf{n}\right|_{(\Omega)}$, cf. [16]. The model refers to (1.17) and (1.18).

$$\min_{\substack{\mathbf{n} \\ \nabla \times \mathbf{n} = 0}} \left\{ \alpha_1 |\nabla \mathbf{n}|^q_{(\Omega)} + \theta_1 \|\mathbf{n} - \nabla I^o\|^p_{(\Gamma)} \right\}, \tag{1.17}$$

$$\min_{I} \left\{ \alpha_2 |\nabla I - \mathbf{n}|^\beta_{(\Omega)} + \theta_2 \|I - I^o\|^\alpha_{(\Sigma)} \right\}. \tag{1.18}$$

Here, we use the same notations as those used in the previous statement. Switching superscripts q and β from 1 to 2, the model varies from TV to H^1 regularization, while changing the other superscripts p and α, the L^1 or the L^2 fidelity follows.

1.4.3 The TV-H^1-Curl-Free Model [15]

Definition 3 (Sparse Reconstruction)**.** *In image processing, sparse reconstruction is the process of reconstructing 3D surfaces from sparse or limited information, e.g., a small number of level lines, point elevations, and vectors.*

The model is initially for the stroke-based surface reconstruction, cf. [15], where only sparse data are available. It cooperates the H^1 term and the TV term in the same model, cf. (1.19) and (1.20).

$$\min_{\substack{\mathbf{n}\\ \nabla\times\mathbf{n}=0}} \left\{ (1-\alpha_1)|\nabla\mathbf{n}|_{(\Omega)} + \alpha_1\|\nabla\mathbf{n}\|^2_{(\Omega)} + \theta_1|\mathbf{n}-\mathbf{n}^o|_{(\Gamma)} \right\}, \tag{1.19}$$

$$\min_{I} \left\{ (1-\alpha_2)|\nabla I-\mathbf{n}|_{(\Omega)} + \alpha_2\|\nabla I-\mathbf{n}\|^2_{(\Omega)} + \theta_2|I-I^o|_{(\Sigma)} \right\}. \tag{1.20}$$

Here, $\alpha_1, \alpha_2 \in [0,1]$ are convex scalar parameters; $\mathbf{n}^o \in \mathbb{R}^2(\Gamma)$ is the known data, e.g., the given vector. The model uses L^1 fidelities.

Chapter 2

Variants of the TV-Stokes - II

In this chapter, we present the new variants of the TV-Stokes model, providing with some alternative approaches to solve these models in addition to the traditional methods via a periodic assumption. We also discuss the relations of these models with Meyer's theory and the TGV model.

2.1 Variants of the TV-Stokes Model

In this section, we present the variants of the TV-Stokes model those contributed by this thesis; these are the multi-dimensional TV-Stokes model, the WTR models [38], and the combined modified TV-Stokes model, respectively.

2.1.1 The Multi-Dimensional TV-Stokes Model, ref. Paper C

To study imageries with more than two dimensions, e.g., Magnetic Resonance Imaging (MRI) data, it is necessary to extend the TV-Stokes model to multiple dimensions. Assume that $\hat{I} \in \mathbb{R}$ is a scalar function. The gradient of $\hat{I}$, in 2D, satisfies the curl-free condition, and its orthogonal vector, i.e., the tangent vector, satisfies the divergence-free condition. The essential of the TV-Stokes model and the TV-Curl-free model is the existence of some scalar function associating to the normal vectors or the tangent vectors. We extend this idea to arbitrary dimensions. The model refers to (2.1) and (2.2).

$$\min_{\substack{\mathbf{n}\\ \mathbf{n}=\Pi\mathbf{n}}} \left\{ \alpha_1 |\nabla \mathbf{n}|_{(\Omega)} + \theta_1 \|\mathbf{n} - \nabla I^o\|^2_{(\Gamma)} \right\}, \tag{2.1}$$

$$\min_{I} \left\{ \alpha_2 |\nabla I|_{(\Omega)} - \beta_2 \nabla I \cdot \frac{\mathbf{n}}{|\mathbf{n}|}_{(\Omega)} + \theta_2 \|I - I^o\|^2_{(\Sigma)} \right\}, \tag{2.2}$$

such that $\Omega \subset \mathbb{R}^{N_1 \times \cdots \times N_d}$ and $\mathbf{n} \in \mathbb{R}^{N_1 \times \cdots \times N_d \times d}$. Here, d is the dimension of the considered problem. Instead of finding the high dimensional curl-free or the high dimensional divergence-free conditions, the associated scalar function constraint applies on the first step (2.1). The operator Π is an orthogonal projector defined as $\Pi(\mathbf{n}) = \nabla(\nabla \cdot \nabla)^{\dagger} \nabla \cdot \mathbf{n}$, where $(\nabla \cdot \nabla)^{\dagger}$ is the Moore-Penrose pseudoinverse.

2.1.2 The WTR Model - Version 1, ref. Paper A

The decoupled models, e.g., TV-H^1-Curl-free model, result somehow less precision in elevations for cases the smoothed vector fields inconsistent with the given elevation constraints. The potential application of the TV-H^1-Curl-free model is locked to shape reconstructions, e.g., 3D cartoon design. The visually correctness and operational simpleness are the focus, so that the precision of elevation is secondary in its concerns. However, for cases, such as maps' reconstruction, elevation is the key part of concerns. The model is then need to be adapted. To improve the precision of the TV-H^1-Curl-free model in sparse reconstructions, we have proposed the first version of WTR model incorporating the first-order and the second-order total variations, cf. [38]. The model refers to (2.3).

$$\min_{I} \left\{ \alpha_1 |\nabla\nabla I|_{(\Omega)} + \alpha_2 |\nabla I|_{(\Omega)} + \theta_1 |\nabla I - \mathbf{n}^o|_{(\Gamma)} + \theta_2 |I - I^o|_{(\Sigma)} \right\}. \tag{2.3}$$

Since I itself is a scalar function, the curl-free condition is not necessary.

2.1.3 The WTR Model - Version 2, ref. Paper B

In the absence of extra regularization along specific directions and appropriate geometric constraints, most of isotropic TV models, including TV-Stokes and its variants, underperform in the balance of the smoothness and the directional structure preservation. To better preserve the irregularity along the level lines, and meanwhile keep the smoothness of reconstructed surface across the level lines, we have proposed another version of WTR model, involving a normal vector matching term. The model refers to (2.4).

$$\min_{I} \left\{ \alpha_1 |\nabla\nabla I|_{(\Omega)} + \alpha_2 |\nabla I|_{(\Omega)} - \theta_1 \nabla I \cdot \mathbf{n}^o_{(\Gamma)} + \theta_2 \|I - I^o\|^2_{(\Sigma)} \right\}. \tag{2.4}$$

2.1.4 The One-Step TV-Stokes Model, ref. Paper E

Inspired by the WTR model, it wise to look at co-optimization models incorporating normals and elevations. In the WTR model, the scalar function I takes the similar role of the scalar function $\hat{I}$ in the TV-Stokes model or the TV-Curl-free model. This equality assumption is not always necessary. We have proposed a new model combining the two steps of the Modified-TV-Stokes model. For convenience, we adopt the curl-free form, cf. (2.5).

$$\min_{\substack{\mathbf{n},I \\ \Pi\mathbf{n}=\mathbf{n}}} \left\{ \alpha_1 |\nabla\mathbf{n}|_{(\Omega)} + \alpha_2 |\nabla I - \mathbf{n}|_{(\Omega)} + \theta_1 \|\mathbf{n} - \nabla I^o\|^2_{(\Omega)} + \theta_2 \|I - I^o\|^2_{(\Omega)} \right\}. \tag{2.5}$$

This work has theoretical importance.

2.1.5 Iterative Regularization for TV-Stokes Model, ref. Paper D

The pioneering work on iterative regularization from Osher and his coworkers, cf. [29], encourages our thinking of different type of iterative algorithms. The idea is to accumulate the features and structures from the decomposed noisy part for each iteration. Distinguishing to the iterative regularization algorithms proposed for the classic Rudin-Osher-Fatemi (ROF) model, cf. [29], Richardson like iterative algorithms are applied on the TV-Stokes model, where the two-minimization-step model involves a vector smoothing in one of them. The model refers to paper D.

2.2 Relation to Meyer's Theory

The denoising problem is to find a decomposition $I^o = I + v$, where $I^o : \Omega \mapsto \mathbb{R}$ is the observed image; Ω is a bounded open subset of $\mathbb{R}^2$; I and v are the signal and the noise, respectively. The ROF model, cf. [31], is one of the most successful and popular models decomposing the observed image in such a way. The model is of the form as follows.

$$\min_I \left\{ \alpha_2 |\nabla I|_{(\Omega)} + \theta_2 \|I - I^o\|^2_{(\Omega)} \right\}. \tag{2.6}$$

Here, the notations are the same as the ones defined for TV-Stokes model, cf. (1.14). The Euler-Lagrange equation of the ROF model is the following.

$$-\frac{\alpha_2}{2\theta_2} \nabla \cdot \frac{\nabla I}{|\nabla I|} = I^o - I. \tag{2.7}$$

It implies that the noise decomposed from the ROF model is of the form as $-\frac{\alpha_2}{2\theta_2} \nabla \cdot \frac{\nabla I}{|\nabla I|}$. In practice, this decomposition is not ideal, that is the noisy part v often contains some fine structures from signal. Meyer has pointed out this v, the high oscillatory part, is the element of the dual space of $BV(\Omega)$, cf. [27, 29].

Remark 1 (G space and G-norm, cf. [27])**.** *Let G denote the Banach space consisting of all generalized functions $v(x,y)$ which can be written as $v(x,y) = \partial_x g_1(x,y) + \partial_y g_2(x,y)$, $g_1, g_2 \in L^\infty(\mathbb{R}^2)$. The norm $\|v\|_G$ is defined as the lower bound of all L^∞ norms of the functions $|\mathbf{g}|$ where $\mathbf{g} = (g1, g2)$, $|\mathbf{g}(x,y)| = \sqrt{g_1(x,y)^2 + g_2(x,y)^2}$.*

In Meyer's theory, the elements of the dual space G are considered as textures of image including noises. Meyer suggested to minimize the G-norm of the textures in cooperating with the total variation for cartoon part, i.e., the decomposed signal part. The model refers to (2.8)

$$\min_I \left\{ \alpha_2 |\nabla I| + \theta_2 \|I - I^o\|_G \right\}. \tag{2.8}$$

. This model is hard to solve via the approaches with Euler-Lagrange equation. One effective way to solve is via the second-order cone program, cf. [13]. A more

solvable model has been proposed by Vese and Osher, cf. [37]. It approximates Meyer's model with a L^p. The model refers to (2.9).

$$\min_{I,\mathbf{g}} \left\{ \alpha_2 |\nabla I| + \theta_2 \|I^o - I - \nabla \cdot \mathbf{g}\|^2 + \gamma_1 \|\mathbf{g}\|_p \right\}. \tag{2.9}$$

Here, α, θ, and γ are positive scalars; $p \geq 1$. It approaches to Meyer's model while $\theta, \gamma \mapsto \infty$. Approximate this model as a two-step model as follows.

$$\min_{\mathbf{g}} \left\{ \gamma_1 \|\mathbf{g}\|_p + \theta_2 \|I^o - \nabla \cdot \mathbf{g}\|^2 \right\},$$
$$\min_{I} \left\{ \alpha_2 |\nabla I| - 2\theta_2 \langle \mathbf{g}, \nabla I \rangle + \theta_2 \|I^o - I\|^2 \right\}.$$

Comparing with the classic TV-Stokes model (the TV-Cur-Free model), the second steps are the same. The TV-Stokes model (the TV-Cur-Free model) uses a different way to approximate $\mathbf{g}$.

2.3 Relation to TGV

There is no evidence showing that the decomposed signal I and associated vector field $\boldsymbol{\tau}(\mathbf{n})$ should be uncorrelated. In contrast, many models show the correlation between these two variables, cf., e.g., [26, 29, 37]. The simplest model we can think of in cooperating these two variables is combining the two steps of the TV-Stokes (TV-Curl-Free) model. For the convenience in discussion, we adopt only the curl form named after TV-Stokes. The model refers to (2.5), showing as follows.

$$\min_{\substack{\mathbf{n},I \\ \Pi\mathbf{n}=\mathbf{n}}} \left\{ \alpha_1 |\nabla \mathbf{n}|_{(\Omega)} + \alpha_2 |\nabla I - \mathbf{n}|_{(\Omega)} + \theta_1 \|\mathbf{n} - \nabla I^o\|^2_{(\Omega)} + \theta_2 \|I - I^o\|^2_{(\Omega)} \right\}.$$

We may solve this model via an iterative way. For the first iteration, we solve the following two-step model.

$$\min_{\substack{\mathbf{n} \\ \Pi\mathbf{n}=\mathbf{n}}} \left\{ \alpha_1 |\nabla \mathbf{n}|_{(\Omega)} + \alpha_2 |\mathbf{n} - \nabla I^o|_{(\Omega)} + \theta_1 \|\mathbf{n} - \nabla I^o\|^2_{(\Omega)} \right\}.$$
$$\min_{I} \left\{ \alpha_2 |\nabla I - \mathbf{n}|_{(\Omega)} + \theta_2 \|I - I^o\|^2_{(\Omega)} \right\}.$$

Here, the initial I is given as $I = I^o$. The model for the first iteration is same as the modified TV-Curl-Free model.

If set the parameter $\theta_1 = 0$, the combined model is equivalent to the $TGV^2_{\alpha(\beta)}$ model, subject to $\beta = 1$. The TGV model refers to papers [4, 22] for example. The concrete form of the $TGV^2_{\alpha(\beta)}$ is as follows.

$$\min_{\mathbf{n}} \left\{ \alpha_1 |\frac{1}{2}(\nabla \mathbf{n} + \nabla \mathbf{n}^\top)|_{(\Omega)} + \alpha_2 |\nabla I - \beta \mathbf{n}|_{(\Omega)} \right\}. \tag{2.10}$$

Here, $\frac{1}{2}(\nabla \mathbf{n} + \nabla \mathbf{n}^\top)$ denotes the symmetrized derivative. We note that, with sufficient smoothness, the vector $\mathbf{n}$ in the combined model is automatically satisfies the symmetric condition. That is because the condition $\Pi \mathbf{n} = \mathbf{n}$ imposing a potential field.

We therefore say the combined model with the parameter $\theta_1 = 0$ is equivalent to the $TGV^2_{\alpha(\beta)}$ model. The modified TV-Curl-Free model is an approximation to $TGV^2_{\alpha(\beta)}$, corresponding to its first iteration.

2.4 Solution Techniques for TV-Stokes and its Variants

In this section, we present some of the new ideas used to solve the TV-Stokes model and its variants. These techniques provide alternative ways to solve the models via the augmented Lagrangian method.

2.4.1 Constrained minimization to unconstrained minimization

We start by introducing unconstrained equivalents of the two constrained minimization problems, namely the original TV-Stokes model and the TV-Curl-free model. These are the first steps used in the two models, respectively.

Lemma 1. *Let* $\boldsymbol{\tau} \in (L_2(\mathbb{R}^2, \Omega))^2$. *Define the operator* Π *such that* $\Pi(\boldsymbol{\tau}) := \boldsymbol{\tau} - \nabla(\nabla \cdot \nabla)^\dagger \nabla \cdot \boldsymbol{\tau}$. *The constrained problem (1.13), i.e.,*

$$\min_{\substack{\boldsymbol{\tau} \\ \nabla \cdot \boldsymbol{\tau} = 0}} \left\{ \alpha_1 |\nabla \boldsymbol{\tau}|_{(\Omega)} + \theta_1 \|\boldsymbol{\tau} - \nabla^\perp I^o\|^2_{(\Gamma)} \right\},$$

is equivalent to the following unconstrained problem

$$\min_{\boldsymbol{\tau}} \left\{ \alpha_1 |\nabla \Pi \boldsymbol{\tau}|_{(\Omega)} + \theta_1 \|\boldsymbol{\tau} - \nabla^\perp I^o\|^2_{(\Gamma)} \right\}. \tag{2.11}$$

Proof. Let $\mathbf{p}$ be the dual variable such that $\mathbf{p} \in C^1_c(\mathbb{R}^4, \Omega)$ and $|\mathbf{p}| \leq 1$, the minimization problem (1.13) takes the following form,

$$\min_{\boldsymbol{\tau}} \max_{\substack{\boldsymbol{\lambda}, \mathbf{p} \\ |\mathbf{p}| \leq 1}} \left\{ \alpha_1 \langle \boldsymbol{\tau}, \nabla \cdot \mathbf{p} \rangle + \hat{\theta}_1 \|\boldsymbol{\tau} - \nabla^\perp I^o\|^2 + \langle \boldsymbol{\lambda}, \nabla \cdot \boldsymbol{\tau} \rangle \right\}, \tag{2.12}$$

where $\boldsymbol{\lambda} \in \mathbb{R}$ is the Lagrange multiplier, $\hat{\theta}_1 = \theta_1$ on Γ otherwise $\hat{\theta} = 0$. Again, by the Minimax theorem, cf. [34], we first consider the problem as the minimization with respect to the vectors $\boldsymbol{\tau}$ and then the maximization with respect to the scalar $\boldsymbol{\lambda}$. The corresponding Euler-Lagrange equations are as follows,

$$\alpha_1 \nabla \cdot \mathbf{p} + 2\hat{\theta}_1 (\boldsymbol{\tau} - \nabla^\perp I^o) - \nabla \boldsymbol{\lambda} = 0, \tag{2.13}$$

and

$$\nabla \cdot \boldsymbol{\tau} = 0.$$

Applying the divergence on both sides of (2.13), we obtain the following relation, using the Moore-Penrose Pseudoinverse,

$$\boldsymbol{\lambda} = \alpha_1(\nabla\cdot\nabla)^{\dagger}\nabla\cdot(\nabla\cdot\mathbf{p}).$$

Consequently, the problem (2.12) takes the form as

$$\min_{\boldsymbol{\tau}}\max_{\substack{\mathbf{p}\\|\mathbf{p}|\leq 1}}\left\{\alpha_1\langle\boldsymbol{\tau},\nabla\cdot\mathbf{p}\rangle+\hat{\theta}_1\|\boldsymbol{\tau}-\nabla^{\perp}I^o\|^2+\alpha_1\langle(\nabla\cdot\nabla)^{\dagger}\nabla\cdot\nabla\cdot\mathbf{p},\nabla\cdot\boldsymbol{\tau}\rangle\right\},$$

or equivalently as follows, with the adjoint,

$$\min_{\boldsymbol{\tau}}\max_{\substack{\mathbf{p}\\|\mathbf{p}|\leq 1}}\left\{\alpha_1\langle\boldsymbol{\tau}-\nabla(\nabla\cdot\nabla)^{\dagger}\nabla\cdot\boldsymbol{\tau},\nabla\cdot\mathbf{p}\rangle+\hat{\theta}_1\|\boldsymbol{\tau}-\nabla^{\perp}I^o\|^2\right\}.$$

Rewriting it with the operator Π, we get the following,

$$\min_{\boldsymbol{\tau}}\max_{\substack{\mathbf{p}\\|\mathbf{p}|\leq 1}}\left\{\langle\Pi(\boldsymbol{\tau}),\nabla\cdot\mathbf{p}\rangle+\hat{\theta}_1\|\boldsymbol{\tau}-\nabla^{\perp}I^o\|^2\right\},$$

which is equivalent to the following primal problem

$$\min_{\boldsymbol{\tau}}\left\{\alpha_1|\nabla\Pi\boldsymbol{\tau}|_{(\Omega)}+\theta_1\|\boldsymbol{\tau}-\nabla^{\perp}I^o\|^2_{(\Gamma)}\right\}.$$

The proof thus follows. □

It is worth mentioning that the operator Π is exactly an orthogonal projection operator into the divergence free subspace, cf. [10]. As a consequence, for all $\boldsymbol{\tau}\in Y=\{\mathbf{m}:\nabla\cdot\mathbf{m}=0\}$, we have $\Pi(\boldsymbol{\tau})=\boldsymbol{\tau}$.

Lemma 2. *Let* $\mathbf{n}\in(L_2(\mathbb{R}^2,\Omega))^2$. *Define operator* Π *such that* $\Pi(\mathbf{n}):=\nabla(\nabla\cdot\nabla)^{\dagger}\nabla\cdot\mathbf{n}$. *The constrained problem, i.e.,*

$$\min_{\substack{\mathbf{n}\\\nabla\times\mathbf{n}=0}}\left\{\alpha_1|\nabla\mathbf{n}|_{(\Omega)}+\theta_1\|\mathbf{n}-\nabla I^o\|^2_{(\Gamma)}\right\},$$

is equivalent to the following unconstrained problem

$$\min_{\mathbf{n}}\left\{\alpha_1|\nabla\Pi\mathbf{n}|_{(\Omega)}+\theta_1\|\mathbf{n}-\nabla I^o\|^2_{(\Gamma)}\right\}. \tag{2.14}$$

Proof. Let $\mathbf{p}$ be the dual variable such that $\mathbf{p}\in(C^1_c(\mathbb{R}^4,\Omega))^2$ and $|\mathbf{p}|\leq 1$, the minimization problem is thus

$$\min_{\mathbf{n}}\max_{\substack{\boldsymbol{\lambda},\mathbf{p}\\|\mathbf{p}|\leq 1}}\left\{\alpha_1\langle\mathbf{n},\nabla\cdot\mathbf{p}\rangle+\hat{\theta}_1\|\mathbf{n}-\nabla I^o\|^2+\langle\boldsymbol{\lambda},\nabla\times\mathbf{n}\rangle\right\}, \tag{2.15}$$

where $\boldsymbol{\lambda} \in \mathbb{R}$ is the Lagrange multiplier, $\hat{\theta}_1 = \theta_1$ on Γ otherwise $\hat{\theta} = 0$. Again by the Minimax theorem, cf. [34], we first consider the minimization with respect to the vectors $\mathbf{n}$ and then the maximization with respect to $\boldsymbol{\lambda}$. The corresponding Euler-Lagrange equations are given as follows,

$$\alpha_1 \nabla \cdot \mathbf{p} + 2\hat{\theta}_1(\mathbf{n} - \nabla I^o) - \nabla^{\perp}\boldsymbol{\lambda} = 0, \tag{2.16}$$

and

$$\nabla \times \mathbf{n} = 0.$$

Taking curl on both sides of (2.16), we obtain the following relation, using the Moore-Penrose Pseudoinverse, i.e.

$$\boldsymbol{\lambda} = \alpha_1 (\nabla \times \nabla^{\perp})^{\dagger} \nabla \times (\nabla \cdot \mathbf{p}).$$

The problem (2.15) is thus

$$\min_{\mathbf{n}} \max_{\substack{\mathbf{p} \\ |\mathbf{p}| \leq 1}} \left\{ \alpha_1 \langle \mathbf{n}, \nabla \cdot \mathbf{p} \rangle + \hat{\theta}_1 \|\mathbf{n} - \nabla I^o\|^2 + \alpha_1 \langle (\nabla \times \nabla^{\perp})^{\dagger} \nabla \times (\nabla \cdot \mathbf{p}), \nabla \times \mathbf{n} \rangle \right\},$$

or equivalently

$$\min_{\mathbf{n}} \max_{\substack{\mathbf{p} \\ |\mathbf{p}| \leq 1}} \left\{ \alpha_1 \langle \mathbf{n} - \nabla^{\perp}(\nabla \times \nabla^{\perp})^{\dagger} \nabla \times \mathbf{n}, \nabla \cdot \mathbf{p} \rangle + \hat{\theta}_1 \|\mathbf{n} - \nabla I^o\|^2 \right\}.$$

We note that $\mathbf{n} - \nabla^{\perp}(\nabla \times \nabla^{\perp})^{\dagger} \nabla \times \mathbf{n} = \nabla(\nabla \cdot \nabla)^{\dagger} \nabla \cdot \mathbf{n}$ by the Helmholtz decomposition. Rewriting it using the Π operator, we get

$$\min_{\mathbf{n}} \max_{\substack{\mathbf{p} \\ |\mathbf{p}| \leq 1}} \left\{ \langle \Pi(\mathbf{n}), \nabla \cdot \mathbf{p} \rangle + \hat{\theta}_1 \|\mathbf{n} - \nabla I^o\|^2 \right\},$$

which is equivalent to the following primal problem

$$\min_{\mathbf{n}} \left\{ \alpha_1 |\nabla \Pi \mathbf{n}|_{(\Omega)} + \theta_1 \|\mathbf{n} - \nabla I^o\|^2_{(\Gamma)} \right\}.$$

The proof thus follows. □

Again, the Π operator here is exactly an orthogonal projection operator onto the curl free subspace, and hence, for all $\mathbf{n} \in Y$ such that $Y = \{\mathbf{m} : \nabla \times \mathbf{m} = 0\}$, we have $\Pi(\mathbf{n}) = \mathbf{n}$.

By using the equivalent formulations, we are able to reduce the number of auxiliary variables as well as the number of multipliers, otherwise needed for the constraint, in the augmented Lagrangian method for each model.

2.4.2 Modified Helmholtz Equations and their Solutions

All models we consider involve a number of high order partial differential equations which we need to solve. We provide effective ways to reformulate those equations so that we solve a non-homogeneous modified Helmholtz equation of the following form, i.e.

$$\triangle u(x,y) - \lambda u(x,y) = F(x,y), \tag{2.17}$$

with Neumann boundary conditions. Here u is the unknown, and λ is a scalar constant or a function.

When λ is a function, this equation is discretized using finite differences (implicit scheme) resulting in a linear system of equations, which is then solved using the Conjugate Gradients iteration, cf. [28]. Diagonal preconditioner has been the simplest and most effective preconditioner for the system.

In case λ is a positive scalar, a fast solver based on the discrete cosine transform, similar to the one used for the Poisson equation, cf. [11, 24] and [10], has been used and is given below. We start by introducing the discrete version of Laplace operator in (2.17) is the matrix

$$\begin{bmatrix} -1 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & 1 & -1 \end{bmatrix},$$

with the help of discrete cosine transformation matrix $C \in \mathbb{R}^{N\times N}$ and the singular value decomposition, we get the following decomposition,

$$-C^\top \begin{bmatrix} 0 & \\ & \Sigma^2 \end{bmatrix} C,$$

where $\Sigma = \mathrm{diag}(\sigma_1, \cdots, \sigma_{N1})$ is the diagonal matrix with its entries representing the singular values $\sigma_k = 2\sin(\pi k/2/N)$ for $k = 1, 2, \cdots, N-1$. Substituting the decomposition back into (2.17), we get

$$-uC^\top \begin{bmatrix} 0 & \\ & \Sigma_x^2 \end{bmatrix} C - C^\top \begin{bmatrix} 0 & \\ & \Sigma_y^2 \end{bmatrix} Cu - \lambda u = F.$$

A further transformation by using $\tilde{u} := CuC^\top$ and $\tilde{F} := CFC^\top$ results with

$$-\tilde{u} \begin{bmatrix} 0 & \\ & \Sigma_x^2 \end{bmatrix} - \begin{bmatrix} 0 & \\ & \Sigma_y^2 \end{bmatrix} \tilde{u} - \lambda \tilde{u} = \tilde{F}.$$

The solution to the above equation can be obtained by a direct entrywise division due to linearity of the equation and the non-singularity of the coefficient matrix (non zero λ), giving

$$\tilde{u} = \tilde{F}./M,$$

where M is the $N \times N$ coefficient matrix defined as

$$M = -\begin{bmatrix} 0 & \sigma_{1,x}^2 & \cdots & \sigma_{N-2,x}^2 & \sigma_{N-1,x}^2 \\ \sigma_{1,y}^2 & \sigma_{1,x}^2+\sigma_{1,y}^2 & \cdots & \sigma_{N-2,x}^2+\sigma_{1,y}^2 & \sigma_{N-1,x}^2+\sigma_{1,y}^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sigma_{N-2,y}^2 & \sigma_{1,x}^2+\sigma_{N-2,y}^2 & \cdots & \sigma_{N-2,x}^2+\sigma_{N-2,y}^2 & \sigma_{N-1,x}^2+\sigma_{N-2,y}^2 \\ \sigma_{N-1,y}^2 & \sigma_{1,x}^2+\sigma_{N-1,y}^2 & \cdots & \sigma_{N-2,x}^2+\sigma_{N-1,y}^2 & \sigma_{N-1,x}^2+\sigma_{N-1,y}^2 \end{bmatrix} - \lambda.$$

The solution to (2.17) is thus

$$u = C^\top((CFC^\top)./M)C.$$

Note that the discrete cosine transformations are easy to implement in MATLAB using the commands `dct2` for the forward transform ($C[\cdot]C^\top$) and `idct2` for the inverse transform ($C^\top[\cdot]C$).

2.4.3 Reformulating the PDEs into Modified Helmholtz Equations

Here we present the types of partial differential equations that have been used in our models, and their respective ways to reformulate them into the required non-homogeneous modified Helmholtz equation.

PDE Type A (λ is a scalar constant)

The first partial differential equation we consider has the following form,

$$\nabla(\nabla\cdot\mathbf{u}) - \lambda\mathbf{u} = \mathbf{F},$$

with the boundary condition

$$\nabla\cdot\mathbf{u} = f,$$

and λ a scalar. The solution of the above equation is thus

$$\mathbf{u} = \frac{1}{\lambda}(\nabla(\nabla\cdot\mathbf{u}) - \mathbf{F}).$$

The term $\nabla\cdot\mathbf{u}$ on the right-hand side can be obtained by solving the following standard form of the inhomogeneous modified Helmholtz equation for $\nabla\cdot\mathbf{u}$ with the Dirichlet boundary condition $\nabla\cdot\mathbf{u} = f$.

$$\nabla\cdot\nabla(\nabla\cdot\mathbf{u}) - \lambda\nabla\cdot\mathbf{u} = \nabla\cdot\mathbf{F}.$$

This equation has a fast solver based on the discrete sine transform.

PDE Type B (λ is a scalar constant)

The second partial differential equation we consider has the following form,

$$\nabla^{\perp}(\nabla \times \mathbf{u}) - \lambda \mathbf{u} = \mathbf{F},$$

with the boundary condition,

$$\nabla \times \mathbf{u} = f.$$

The solution of the above equation is thus

$$\mathbf{u} = \frac{1}{\lambda}(\nabla^{\perp}(\nabla \times \mathbf{u}) - \mathbf{F}).$$

The term $\nabla \times \mathbf{u}$ on the right-hand side is obtained by solving the following non-homogeneous modified Helmholtz equation for $\nabla \times \mathbf{u}$ with the Dirichlet boundary condition $\nabla \times \mathbf{u} = f$.

$$\nabla \cdot \nabla(\nabla \times \mathbf{u}) - \lambda \nabla \times \mathbf{u} = \nabla \times \mathbf{F}.$$

This equation has a fast solver based on the discrete sine transform. Note that $\nabla \times \nabla^{\perp} = \nabla \cdot \nabla$.

PDE Type C (λ is a function)

The third partial differential equation we consider has the following form,

$$\nabla \cdot \nabla \mathbf{u} - \lambda \mathbf{u} = \mathbf{F} - \xi \nabla(\nabla \cdot \mathbf{u}),$$

corresponding to $\nabla \cdot \mathbf{u} = f$, $\nabla u_1 \cdot \boldsymbol{\nu} = \mathbf{a} \cdot \boldsymbol{\nu}$, and $\nabla u_2 \cdot \boldsymbol{\nu} = \mathbf{b} \cdot \boldsymbol{\nu}$ at boundary. Here λ is a function; ξ is a scalar; $\mathbf{a}, \mathbf{b}$ and f are known vectors and scalar, respectively. $\boldsymbol{\nu}$ is the outward unit normal vector on the boundary of the domain. If the term $\nabla \cdot \mathbf{u}$ is known, this is a nonstandard form of the inhomogeneous modified Helmholtz equation with Neumann boundary conditions. It has a fast solver via the preconditioned conjugate gradient method.

To find the value of $\nabla \cdot \mathbf{u}$, we solve the following nonstandard form of the inhomogeneous modified Helmholtz equation for $\nabla \cdot \mathbf{u}$ with a Dirichlet boundary condition $\nabla \cdot \mathbf{u} = f$.

$$\nabla \cdot \nabla(\nabla \cdot \mathbf{u}) - \frac{\lambda}{1+\xi} \nabla \cdot \mathbf{u} = \frac{1}{1+\xi} \nabla \cdot \mathbf{F}.$$

PDE Type D (λ is a function)

The fourth partial differential equation we consider has the following form,

$$\nabla \cdot \nabla \Pi \mathbf{u} - \lambda \mathbf{u} = \mathbf{F}\ .$$

Boundary conditions are of the following form.

$$\nabla(\Pi \mathbf{u})_1 \cdot \boldsymbol{\nu} = \mathbf{a} \cdot \boldsymbol{\nu} \quad \text{and} \quad \nabla(\Pi \mathbf{u})_2 \cdot \boldsymbol{\nu} = \mathbf{b} \cdot \boldsymbol{\nu},$$

where λ is a function; $\mathbf{F}$, $\mathbf{a}$, and $\mathbf{b}$ are known vectors. $\boldsymbol{\nu}$ is the outward unit normal vector on the boundary of the domain. Π is an orthogonal projector defined as $\Pi(\mathbf{u}) = \mathbf{u} - \nabla(\nabla\cdot\nabla)\nabla\cdot\mathbf{u}$ or $\Pi(\mathbf{u}) = \nabla(\nabla\cdot\nabla)\nabla\cdot\mathbf{u}$. The terms $(\Pi\mathbf{u})_1$ and $(\Pi\mathbf{u})_2$ are two components of vector $\Pi\mathbf{u}$. Reform the equation as follows.

$$\nabla\cdot\nabla\Pi\mathbf{u} - \lambda\Pi\mathbf{u} = \Pi\mathbf{F}.$$

This is a nonstandard form of the inhomogeneous modified Helmholtz equation for $\Pi\mathbf{u}$ with Neumann boundary conditions. It has a fast solver via the preconditioned conjugate gradient method.

Using the obtained $\Pi\mathbf{u}$, $\mathbf{u}$ is simply the following.

$$\mathbf{u} = \begin{cases} \dfrac{1}{\lambda}(\nabla\cdot\nabla\Pi\mathbf{u} - \mathbf{F}) & (\lambda \neq 0) \\ \Pi\mathbf{u}, & (\lambda = 0). \end{cases}$$

PDE Type E (λ is a function)

The fifth partial differential equation we consider has the following form,

$$\nabla\cdot\nabla\mathbf{u} - \lambda\mathbf{u} = \mathbf{F} - \xi\nabla^{\perp}\nabla\times\mathbf{u},$$

corresponding to $\nabla\times\mathbf{u} = f$, $\nabla n_1\cdot\boldsymbol{\nu} = \mathbf{a}\cdot\boldsymbol{\nu}$, and $\nabla n_2\cdot\boldsymbol{\nu} = \mathbf{b}\cdot\boldsymbol{\nu}$ at boundary. Here λ is a function; ξ is a scalar; $\mathbf{a},\mathbf{b}$ and f are known vectors and scalar, respectively. $\boldsymbol{\nu}$ is the outward unit normal vector on the boundary of the domain. If the term $\nabla\times\mathbf{u}$ is known, this is a nonstandard form of the inhomogeneous modified Helmholtz equation with Neumann boundary conditions. It has a fast solver via the preconditioned conjugate gradient method.

To find the value of $\nabla\times\mathbf{u}$, we solve the following nonstandard form of the inhomogeneous modified Helmholtz equation for $\nabla\times\mathbf{u}$ with a Dirichlet boundary condition $\nabla\times\mathbf{u} = f$.

$$\nabla\cdot\nabla(\nabla\times\mathbf{u}) - \frac{\lambda}{1+\xi}\nabla\times\mathbf{u} = \frac{1}{1+\xi}\nabla\times\mathbf{F}.$$

2.4.4 Optimization Problem Containing One L^1 Term

In all our models we have subproblems with one L^1 term, which is the total variation regularization term. We have a simple way to derive solution to each of the optimization problems which contain only one L^1 term. The minimization problem has the form as follows.

$$\min_{\mathbf{u}}\left\{\alpha|\mathbf{u}| + \sum_{i=1}^{k}\beta_i\mathbf{a}_i\cdot\mathbf{u} + \sum_{j=1}^{m}\frac{\gamma_j}{2}\|\mathbf{u}-\mathbf{b}_j\|^2\right\}, \tag{2.18}$$

where $\mathbf{u}\in\mathbb{R}^2$ or $\mathbf{u}\in\mathbb{R}^4$ is a vector or a matrix; $k,m\in\mathbb{N}$ are the number of respective terms; $\alpha,\gamma\in\mathbb{R}^+$, and $\beta\in\mathbb{R}$ are the corresponding coefficients.

Definition 4 (Compatibility)**.** *If A and B are two matrices such that $A = \lambda B$ for some non-negative scalar λ, then we say that A is compatible with B. It is easy to see that $A/|A| = B/|B|$.*

We will find a closed form solution to this. The corresponding optimality condition, or the Euler-Lagrange equation, is as follows

$$\frac{\alpha}{\sum_{j=1}^{m}\gamma_j}\frac{\mathbf{u}}{|\mathbf{u}|}+\mathbf{u}=\sum_{j=1}^{m}\gamma_j\mathbf{b}_j-\sum_{i=1}^{k}\beta_i\mathbf{a}_i.$$

Since α and $\sum_{j=1}^{m}\gamma_j \neq 0$ are both positive numbers and $|\mathbf{u}|$ is a positive number, the vectors (matrices) $\mathbf{u}$ and $(\sum_{j=1}^{m}\gamma_j\mathbf{b}_j-\sum_{i=1}^{k}\beta_i\mathbf{a}_i)$ become compatible in the sense of Definition 4 because $\big(\alpha/(\sum_{j=1}^{m}\gamma_j)|\mathbf{u}|^{-1}+1\big)\mathbf{u}=\sum_{j=1}^{m}\gamma_j\mathbf{b}_j-\sum_{i=1}^{k}\beta_i\mathbf{a}_i$. According to which, we can replace the matrix $\mathbf{u}/|\mathbf{u}|$ with $(\sum_{j=1}^{m}\gamma_j\mathbf{b}_j-\sum_{i=1}^{k}\beta_i\mathbf{a}_i)/|\sum_{j=1}^{m}\gamma_j\mathbf{b}_j-\sum_{i=1}^{k}\beta_i\mathbf{a}_i|$ in the above equation. Moving the first term to the right hand side we get

$$\mathbf{u}=\left(1-\frac{\alpha}{\sum_{j=1}^{m}\gamma_j|\sum_{j=1}^{m}\gamma_j\mathbf{b}_j-\sum_{i=1}^{k}\beta_i\mathbf{a}_i|}\right)\left(\sum_{j=1}^{m}\gamma_j\mathbf{b}_j-\sum_{i=1}^{k}\beta_i\mathbf{a}_i\right).$$

Again since we have already seen that in the sense of Definition 4, $\mathbf{u}$ and $(\sum_{j=1}^{m}\gamma_j\mathbf{b}_j-\sum_{i=1}^{k}\beta_i\mathbf{a}_i)$ are compatible, the coefficient $(1-\alpha/(\sum_{j=1}^{m}\gamma_j|\sum_{j=1}^{m}\gamma_j\mathbf{b}_j-\sum_{i=1}^{k}\beta_i\mathbf{a}_i|))$ must also be non-negative. Hence

$$\mathbf{u}=\max\left\{0,1-\frac{\alpha}{\sum_{j=1}^{m}\gamma_j|\sum_{j=1}^{m}\gamma_j\mathbf{b}_j-\sum_{i=1}^{k}\beta_i\mathbf{a}_i|}\right\}\left(\sum_{j=1}^{m}\gamma_j\mathbf{b}_j-\sum_{i=1}^{k}\beta_i\mathbf{a}_i\right),$$

which is the closed form solution to the minimization problem (2.18).

Chapter 3

Introduction to Papers and Conclusion

3.1 Abstracts of the Papers

Paper A: *Sparse-Data Based 3D Surface Reconstruction for Cartoon and Map*

A model combining the first-order and the second-order regularizer for the purpose of 3D surface reconstruction based on 2D sparse data is proposed. The model includes a hybrid fidelity constraint which allows the initial conditions to be switched flexibly between vectors and elevations. A numerical algorithm based on the augmented Lagrangian method is also proposed. The numerical experiments are presented, showing its excellent performance both in designing cartoon characters, as well as in recovering oriented three dimensional maps from contours or points with elevation information.

Paper B: *Sparse-Data Based 3D Surface Reconstruction with Vector Matching*

Three dimensional surface reconstruction based on two dimensional sparse information, in the form of only a small number of level lines of the surface with moderately complex structures containing both structured and unstructured geometries, is considered in this paper. A new model has been proposed which is based on the idea of using normal vector matching combined with a first order and a second order total variation regularization. A fast algorithm based on the augmented Lagrangian is also proposed. Numerical experiments are provided showing the effectiveness of the model and the algorithm in reconstructing surfaces with much detail featuring complex structures from both synthetic and real world digital maps.

Paper C: *Multidimensional TV-Stokes for Image Processing*

A multidimensional TV-Stokes model is proposed based on a potential field assumption. Numerical algorithm is proposed using the Chambolle's semi-implicit dual formula. Numerical results in 3-dimensional case show an excellent performance in fine structures preserving.

Paper D: *Iterative Regularization with TV-Stokes for Image Denoising*

We propose separate iterative regularization algorithms for the TV-Stokes model to restore images from Gaussian noise. These are extensions of the iterative regularization algorithms proposed for the classic Rudin-Osher-Fatemi (ROF) model, which is a one-minimization-step model, to the TV-Stokes model, a two-minimization-step model involving a vector smoothing in one of them. The iterative regularization algorithms proposed are Richardson like, which give a significant improvement over the original method in the quality of the restored image. The paper accomplishes with the convergence analysis and numerical experiments.

Paper E: *Alternating Minimization for One-Step TV-Stokes Model for Image Denoising*

The paper presents a fully coupled TV-Stokes model, and propose an algorithm based on alternating minimization of the objective functional whose first iteration is exactly the modified TV-Stokes model proposed earlier. The model is a generalization of the second order Total Generalized Variation model. A convergence analysis is given.

3.2 Conclusion and Future Work

The original idea of TV-Stokes for image processing has been based on smoothing its vector field and then using it in the reconstruction of the image in two separate steps. This does not work well in complex problems, problems with large sparsity and irregularity. It has been necessary to have them somehow fully coupled in the models. Our work has been based on finding ways to couple the image intensity and its vector fields for such complex problems, and solving them effectively.

The thesis studies the TV-Stokes model and its variants systematically. New models, as variants of the TV-Stokes, and their effective solvers, have been proposed. The cut-in point of the thesis has been the work on the iterative regularization algorithms for the TV-Stokes for denoising (Paper D). The work considered the decomposed noisy part where some features or structures of the image still remained. The idea has been to get those features or structures back into the decomposed clean part through iteration. Mathematically it is a Richardson-like algorithm. Numerical experiments have shown more details being recovered as compared to the TV-Stokes. The work demonstrates the success of using the TV-stokes model as a feature extractor with the assistance of the smooth tangent or normal field. The next work (Paper E) has been an extension of the iterative regularization algorithm, where a fully coupled one step model, instead of the two-step TV-Stokes, has been used. The coupled single-step model has a natural consistency because both normal fields and elevations are used in the same functional, making it possible to accurately recover objects with shapes and heights. The first iteration of the model is

exactly the modified variant of the TV-Stokes [23]. It is also a generalisation of the second-order TGV model [4]. The work has strong theoretical importance. Image processing problems in practice may also be defined in three dimensions, it was therefore necessary to extend the original TV-Stokes model [30] to higher dimensions. Our next work (Paper C) has been on the multi-dimensional TV-Stokes model where a more general constraint, that is the potential function constraint, has been used replacing the divergence-free constraint. The new constraint is applicable to arbitrary dimensions, and is interesting for its simplicity.

The major work in this thesis focus on the sparse data reconstruction which finds applications everywhere from 3D cartoon design to height map reconstruction based on sketches and contours. The first (Paper A) in this direction, has been a new model extending the TV-H^1-Curl-free model [15], where we include both the vector constraint and the height constraint in one place. The curl-free condition is satisfied automatically in the proposed model as it is of second order. Results from experiments show encouraging performance with respect to the accuracy in the height and the reconstructed shapes. Experiments also demonstrates the method's handling of data with extreme sparsity. The next work (Paper B) on sparse reconstruction has been based on the use of vector matching condition for improved accuracy. In this work, a normal matching term has been used replacing the vector fidelity proposed in the previous work. Experiments on height map reconstruction, with non-smooth contours, valleys, and flat areas, etc., have shown excellent results, preserving the horizontal irregularity as well as vertical continuity. It outperforms state-of-the-art variational models. The thesis proposes some unique techniques to solve the subproblems based on the augmented Lagrangian. Subproblems are formulated in three types of subproblems, i.e., quadratic problem, problems with only one L^1 term, and problems that can be reformulated as non-homogeneous modified Helmholtz problems with either constant or variable coefficients. For these three types of subproblems, we either have closed form solutions, or fast iterative solvers.

The thesis has demonstrated the excellent performance of the TV-Stokes model and its variants on some known problems. They also have huge potential in other problems, which may be considered for future. Experiments have shown the effect of vector matching on geometric shapes, in particular geometries with corners etc., but they depend on the parameters for the matching terms. An obvious choice for future work will be finding algorithms for the parameters for different geometries. One may use supervising functional such as learning based techniques for the purpose. Another future work can be a generalization of the vector matching concept to higher dimensional tensors which is also highly promising to achieve a better representation of features in natural structures. A systematic study of the models, coupled and decoupled, are also interesting.

Part II

Scientific Contributions

Paper A

Sparse-Data Based 3D Surface Reconstruction for Cartoon and Map

Bin Wu, Talal Rahman and Xue-Cheng Tai *Imaging, Vision and Learning Based on Optimization and PDEs*, Springer, Cham, (2016): 47.

Paper B

Sparse-Data Based 3D Surface Reconstruction with Vector Matching

Bin Wu, Xue-Cheng Tai and Talal Rahman *Submitted. arXiv:2009.12994*

SPARSE-DATA BASED 3D SURFACE RECONSTRUCTION WITH VECTOR MATCHING

BIN WU*, XUE-CHENG TAI†, AND TALAL RAHMAN*

Abstract. Three dimensional surface reconstruction based on two dimensional sparse information in the form of only a small number of level lines of the surface with moderately complex structures, containing both structured and unstructured geometries, is considered in this paper. A new model has been proposed which is based on the idea of using normal vector matching combined with a first order and a second order total variation regularizers. A fast algorithm based on the augmented Lagrangian is also proposed. Numerical experiments are provided showing the effectiveness of the model and the algorithm in reconstructing surfaces with detailed features and complex structures for both synthetic and real world digital maps.

Key words. total variation regularization, surface reconstruction, augmented Lagrangian

AMS subject classifications. 65F05

1. Introduction. Reconstructing an image from its little information available is a challenging but interesting task for image processing, which has attracted much attention over the years, finding applications in many areas where, for instance, information are available only in the form of a small number of level lines or isolated points. Two of its well known applications are the sketch based design where three dimensional figures or structures are designed from sketches made by an artist or a computer program, and the surface reconstruction where three dimensional surface is reconstructed from its level lines, see for instance two recent publications [11, 15], for a review on the subject. In this paper, we are interested surfaces that are moderately complex, in the sense that they contain both structured and unstructured geometries, man made or natural, and feature both kinks (sharp twists) and creases (folds). We are particularly interested in a variational model that combines both the features of a sketch based design and that of a surface reconstruction, by incorporating both the height and the vector information from the level lines in a natural way so that a more precise design or reconstruction is possible.

Sketch based design has been a popular way of design in three dimensions, cf. e.g. [37, 13, 14, 22, 26], because it is both intuitive and effective, particularly in the games and cartoon design. In a sketch based design, the information is available either in the form of contour lines (level lines) with or without the height values, cf. e.g. [13], complex sketches with elevations, cf. e.g. [14], or structured annotations, cf. e.g. [9]. The class of algorithms presented in these papers are however limited in their capabilities, particularly when it comes to reconstructing structures with crease. Although there exists a work on artificially adding crease to the design, cf. e.g. [25], for large and complex images, such algorithms become less effective and computationally more expensive. A different class of models, based on the total variation minimization, turned out to be much more effective, cf. [11]. The model presented in these two papers is based on interpolating normal vectors under the curl-free constraint and then reconstructing the 3D surface from the obtained vector field. Inspired by the use

*Department of Computer Science, Electrical Engineering and Mathematical Sciences, Western Norway University of Applied Sciences, Inndalsveien 28, 5063 Bergen, Norway (bin.wu@hvl.no, talal.rahman@hvl.no).

†Department of Mathematics, University of Bergen, Allégaten 41, 5007 Bergen, Norway (xuecheng.tai@uib.no).

of surface gradients in surface reconstruction, cf. e.g. [1, 8, 27, 30, 38, 36, 28, 23], the model is an extension of the TV-Stokes model, cf. [29], to surface reconstruction. The TV-Stokes models are based on using the curl-free or the divergence free constraint in image processing, cf. e.g. [6, 31, 12, 16] for further works on TV-Stokes models. The model of [11] performs very well in preserving both edges and crease structures, however, it requires prior information on the vectors, e.g. the length of vectors [34, 11], which are not always available.

Three dimensional (3D) surface reconstruction from contours or isolated points with height values, has been the second most popular application area. Unlike the three dimensional design, the height values here are needed because the reconstructed surfaces are expected to be as close as possible to the ground truth, e.g., in digital elevation maps and data compression, cf. e.g. [15]. One approach to solve the problem is to use explicit parametrization of the given contours, with subsequent point wise matching and interpolation between the contours, cf. e.g. [20, 17, 19]. In some cases, an explicit parametrization may be difficult and expensive to compute, and a loss of continuity of slope across contours may become a challenge to deal with. An alternative way is to consider the surface as a function over the domain, and interpolate the function based on solving partial differential equations, as in the AMLE (Absolutely Minimizing Lipschitz Extension) model, cf. e.g. [2, 4]. Although AMLE interpolation is able to interpolate data given on isolated points and on level lines, it has the drawback that the level lines of its interpolants are smooth making it difficult to preserve kinks, and consequently creases on the surface, plus it cannot interpolate slopes of the surface. To overcome this, one has to rely on higher-order methods or regularizers, cf. e.g. [7, 18, 3, 21, 15].

Particularly interesting to our present work, is the model proposed in [15], a variational model, which uses a third order anisotropic regularizer whose anisotropy is based on an auxiliary vector field connecting adjacent level lines; it is the field of directions formed explicitly in a separate step, in which the normals of the level lines change the least. The model performs very well for surface reconstruction, particularly in preserving the geometry of the given level lines, and propagating it smoothly across the level lines. The model does not require any regularity of the level lines, however, it requires the level lines lying next to each other to be similar, so that points lying on them can be associated. This may not always be the case, in particular when there are only few level lines available.

Our aim is to recover the surface from the few level lines that are available, and somehow use the geometry of those level lines in a physically consistent way providing further precision to our reconstruction. In case of a 3D design, which we will not cover in this paper, these level lines will be the sketch lines drawn by an artist. We propose a simple one step variational model incorporating both the height and the vector information in the model, the vectors being the unit normal vectors, either calculated from the level lines itself or provided by the artist on the sketches. The model consists of a first-order and a second-order total variation (isotropic) regularizer under a fidelity constraint on the height, together with a vector (normal vector) matching term to account for the anisotropy in the model.

We do not impose any regularity on the level lines, nor do we assume any similarity between the level lines close to each other. In case of very sparse data this is very likely to happen. We require the geometry of the level lines, in particular, the non-differentiability along level lines, as well as the smoothness of the gradient across level lines be preserved in the interpolated surface. The model allows for adaptive adjustment of the normal vectors, consequently the shape of the surface, and an almost

perfect reconstruction even with a small number of level lines for complex surface with mixed geometries. We propose a fast algorithm for the numerical solution, which is based on the augmented Lagrangian method [10, 32, 35], featuring sub-problems with closed form solutions and fast iterative solution.

The paper is organized as follows. In section 2, we propose our model for surface reconstruction from a set of sketches or level lines, and individual points, in section 3, we present a fast algorithm for the numerical solution, based on the augmented Lagrangian, and in section 4, we present our numerical experiments on both synthetic data and real data for the verification. Finally, in section 5, we give our conclusion.

2. The proposed model. The problem we consider is an inverse problem to recover the two dimensional (2D) height map $I : \Omega \mapsto \mathbb{R}$ on a domain $\Omega \subset \mathbb{R}^2$ in two space dimensions, from known sparse data given in the form of level lines as $C_i = \{x : I(x) = I_{\Sigma i}\}$ corresponding to the given elevation $I_{\Sigma i}$, for $i = 1, \ldots, N$. The collection of all the given level lines is denoted by Σ, that is $C_i \in \Sigma$. The aim is to propose a model that is based on the total variation minimization and an explicit use of the geometry of the level lines in the functional. The use of this geometry, in one form or another, is essential for the surface reconstruction, e.g., for capturing kinks, specially when the level lines have coastal like structures. However, how well the kinks are represented depends on how well the geometry of the level lines has been incorporated into the model. There are models which include such information, cf. e.g. [11, 15]. However, the use of such information, in these papers, has been done somewhat indirectly, that is by first constructing a smooth vector field from the vector information available on the level lines, that is in a prepossessing step, and then using the new vector field in the reconstruction step. These models are quite powerful and effective, however, there are cases where these models fall short; we refer to the numerical section of this paper for one such case. We propose a more direct approach, where we introduce a vector matching term into our model, involving explicitly the vector information available on the level lines, as follows:

$$\min_I \left\{ g|\nabla(\nabla I)|_{F\ (\Omega)} + h|\nabla I|_{\ (\Omega)} + \alpha V(\nabla I, \mathbf{v}_\Gamma)_{\ (\Gamma)} + \theta |I - I_\Sigma|^2_{2\ (\Sigma)} \right\}, \tag{1}$$

where the first two terms are the second order and the first order total variational regularizer, respectively, the third term is the vector matching term, and the fourth term is the data fidelity term. $|\cdot|_{F\ (\Omega)}$, $|\cdot|_{\ (\Omega)}$, and $|\cdot|_{2\ (\Sigma)}$ are the Frobenius-, L^1-, and L^2-, norm, respectively, over the domain specified inside bracket. The vector matching term is the integral over Γ of function V. g, h, α and θ are the scalar parameters. Γ is the set of points where the unit vectors $\mathbf{v}_\Gamma$ are given and Σ is the set of points where the elevation/height values I_Σ are given, see FIG. 1 for an illustration. We note that the vector $\mathbf{v}_\Gamma$ is either given, or can be extracted from the given level lines.

The vector matching term. On a given level line the vector $\mathbf{v}_\Gamma$ can either be given as the unit tangent vector or be given as the unit normal vector. The idea is to match ∇I^*, that is the gradient of the reconstructed surface I^*, to this vector in some consistent way, e.g., ∇I^* should be orthogonal to $\mathbf{v}_\Gamma$ if $\mathbf{v}_\Gamma$ is the unit tangent vector and parallel otherwise.

In case of the unit tangent vector, the only way this matching can happen is through minimizing a norm of $\nabla I \cdot \mathbf{v}_\Gamma$, e.g.,

$$V(\nabla I, \mathbf{v}_\Gamma)_{\ (\Gamma)} := |\nabla I \cdot \mathbf{v}_\Gamma|^2_{2\ (\Gamma)}. \tag{2}$$

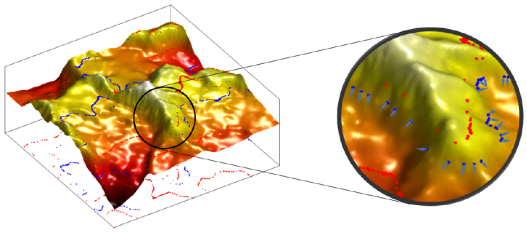

FIG. 1. *Given 3 dimensional surface on a 2 dimensional domain.* Γ *(the set of blue points) corresponds to the points where unit vectors are given;* Σ *(the set of red points) corresponds to the points where the elevation/height values are given. These are points that are either isolated or given on level lines.*

If we use this norm, the corresponding minimization problem becomes as follows,

$$\min_I \left\{ g|\nabla(\nabla I)|_{F\ (\Omega)} + h|\nabla I|_{\ (\Omega)} + \alpha|\nabla I \cdot \mathbf{v}_\Gamma|^2_{2\ (\Gamma)} + \theta|I - I_\Sigma|^2_{2\ (\Sigma)} \right\}. \tag{3}$$

Its minimum is attained when ∇I^* ($\nabla I^* \neq \mathbf{0}$) is perpendicular to the given vector $\mathbf{v}_\Gamma$, or $\nabla I^* = \mathbf{0}$. There are basically two shapes that the minimization generates. In the first case, when $\nabla I^* \perp \mathbf{v}_\Gamma$, $\nabla I^* \neq \mathbf{0}$, α can take any number, also when $\alpha = 0$, which is an isotropic case. In the second case, when $\nabla I^* = \mathbf{0}$, the minimization generates flattened structures.

In case of the unit normal vector, the only consistent way is minimizing the inner product $-\nabla I \cdot \mathbf{v}_{\Gamma\ (\Gamma)}$, that is

$$V(\nabla I, \mathbf{v}_\Gamma)_{\ (\Gamma)} := -\nabla I \cdot \mathbf{v}_{\Gamma\ (\Gamma)}. \tag{4}$$

With this, the corresponding minimization problem becomes as follows,

$$\min_I \left\{ g|\nabla(\nabla I)|_{F\ (\Omega)} + h|\nabla I|_{\ (\Omega)} - \alpha\nabla I \cdot \mathbf{v}_{\Gamma\ (\Gamma)} + \theta|I - I_\Sigma|^2_{2\ (\Sigma)} \right\}. \tag{5}$$

Its minimum is attained only when ∇I^* is pointing in the same directions as the unit vector $\mathbf{v}_\Gamma$. With this matching term in the functional, cf. (5), any irregularities in the given level lines will be preserved, which is then stretched smoothly across the level lines by the regularizers. Also, by varying the parameter α we can control the shape of the reconstructed surface, see the numerical section for an illustration. For our model, we therefore choose $\mathbf{v}_\Gamma$ to be the unit normal vector on the level lines.

The regularizer terms. To explain our choice of regularizer, we consider the following single regularizer model, where we consider only one regularizer term in the functional, with varying order, that is

$$\min_I \left\{ g|\nabla^k I|_{F\ (\Omega)} - \alpha\nabla I \cdot \mathbf{v}_{\Gamma\ (\Gamma)} + \theta|I - I_\Sigma|^2_{2\ (\Sigma)} \right\}, \tag{6}$$

for $k = 1, 2, \ldots$. The notations are the same as before, cf. (5). Choosing α small compared to both g and θ, the optimal solution of the problem (6) at any position away from Σ and Γ would give $\nabla^k I \sim 0$. For $k = 1$, it corresponds to $\nabla^1 I \sim 0$, in other words no change in the height value, thereby producing a staircase effect on the surface. For $k = 2$, this corresponds to the Hessian $\nabla^2 I \sim 0$, implying no change in the gradient of the height, thereby preserving the slope, consequently the shape, of

the surface. For our model, it is enough to use the two, or a combination of the two. Increasing the number k introduces oscillations to the solution, becoming unbounded near the boundary, see the numerical section for an illustration.

Our model is a combination of a first order and a second order isotropic regularizer with an anisotropic matching term, which is different from the anisotropic third order regularizer of [15], in the sense that the anisotropy in our model is separated from the regularizer, while in the latter the anisotropy is built in the regularizer itself.

The choice of parameters. The choice of parameters g, h, α and θ is crucial for our model In general the data fidelity parameter θ is set to be large enough to ensure the reconstructed surface is consistent with the given data. The choice of vector matching parameter α depends on the surface to be reconstructed, for instance, the bigger the value the steeper the structure becomes. The parameter h (for the first order total variation) is used for flat or almost flat structure. The larger its value the stronger the flattening is. The parameter g (for the second order total variation) is used for sloped structure smooothly connecting the level lines across. The larger the g the smoother the surface between the level lines becomes.

3. The numerical algorithm. In this section, we introduce our algorithm based on the augmented Lagrangian for solving the optimization problem (1), cf. [10, 24] for general literature on augmented Lagrangian method, and [32, 35] for its applications in image processing.

We denote $\hat{\alpha}$ and $\hat{\theta}$ as follows,

$$\hat{\alpha} = \begin{cases} \alpha, & \text{on } \Gamma \\ 0, & \text{in } \Omega\backslash\Gamma, \end{cases} \qquad \text{and} \quad \hat{\theta} = \begin{cases} \theta, & \text{on } \Sigma \\ 0, & \text{in } \Omega\backslash\Sigma. \end{cases}$$

The problem (1) is thus defined on the whole of domain Ω as,

$$\min_I \left\{ g|\nabla(\nabla I)|_F + h|\nabla I| - \hat{\alpha}\nabla I \cdot \mathbf{v}_\Gamma + \hat{\theta}|I - I_\Sigma|_2^2 \right\}. \tag{7}$$

Accordingly, we introduce auxiliary variables for the derivatives in L^1-terms turning the unconstrained optimization problem into a constrained optimization problem, cf. e.g., [24, 32, 35].

We set $\mathbf{Q} := \nabla\mathbf{E}$, $\mathbf{P} := \nabla I$ and $\mathbf{E} := \mathbf{P}$, where $\mathbf{E} \in \mathbb{R}^2$ and $\mathbf{P} \in \mathbb{R}^2$ are 2-dimensional vectors, and $\mathbf{Q} \in \mathbb{R}^{2\times 2}$ is a 2-by-2 matrix. Using the new variables, we get the following constrained minimization problem.

$$\min_{\mathbf{Q},\mathbf{P},I} \left\{ g|\mathbf{Q}|_F + h|\mathbf{P}| - \hat{\alpha}\mathbf{P} \cdot \mathbf{v}_\Gamma + \hat{\theta}|I - I_\Sigma|_2^2 \right\},$$

subject to

$$\mathbf{P} - \nabla I = 0, \quad \mathbf{E} - \mathbf{P} = 0, \quad \text{and} \quad \mathbf{Q} - \nabla\mathbf{E} = 0.$$

Assigning to each constraint a Lagrange multiplier and a penalty term, the Lagrangian

functional reads as follows,

$$\begin{aligned}(8)\quad \mathscr{L}(\mathbf{Q},\mathbf{P},\mathbf{E},I;\mathbf{\Lambda}_Q,\mathbf{\Lambda}_P,\mathbf{\Lambda}_E) \;=\; & g|\mathbf{Q}|_F \;+\; h|\mathbf{P}| \;-\; \hat{\alpha}\mathbf{P}\cdot\mathbf{v}_\Gamma \;+\; \hat{\theta}|I-I_\Sigma|^2 \\ & +\mathbf{\Lambda}_Q\cdot(\mathbf{Q}-\nabla\mathbf{E}) \;+\; \frac{c_Q}{2}|\mathbf{Q}-\nabla\mathbf{E}|_F^2 \\ & +\mathbf{\Lambda}_P\cdot(\mathbf{P}-\nabla I) \;+\; \frac{c_P}{2}|\mathbf{P}-\nabla I|^2 \\ & +\mathbf{\Lambda}_E\cdot(\mathbf{E}-\mathbf{P}) \;+\; \frac{c_E}{2}|\mathbf{E}-\mathbf{P}|^2,\end{aligned}$$

where $\mathbf{\Lambda}_Q$, $\mathbf{\Lambda}_P$, and $\mathbf{\Lambda}_E$ are the Lagrange multipliers, c_Q, c_P, and c_E are the positive penalty parameters. The augmented Lagrangian method is to seek the saddle point of the following problem:

$$(9)\qquad \min_{\mathbf{Q},\mathbf{P},\mathbf{E},I}\ \max_{\mathbf{\Lambda}_Q,\mathbf{\Lambda}_P,\mathbf{\Lambda}_E} \mathscr{L}(\mathbf{Q},\mathbf{P},\mathbf{E},I;\mathbf{\Lambda}_Q,\mathbf{\Lambda}_P,\mathbf{\Lambda}_E).$$

For the solution, we solve its associated system of optimality conditions with an iterative procedure, see Algorithm 3.1 and Algorithm 3.2. For the convenience, we use $\mathbf{\Lambda} := (\mathbf{\Lambda}_Q,\mathbf{\Lambda}_P,\mathbf{\Lambda}_E)$ to denote the set of Lagrange multipliers.

Algorithm 3.1: The augmented Lagrangian for the problem (9)

Input: $\mathbf{v}_\Gamma$ (The unit vector along Γ) and I_Σ (The elevation along Σ)

Output: I (The elevation on the whole domain Ω)

Set $k=0$;

Initialize each of $\mathbf{Q}^0$, $\mathbf{P}^0$, $\mathbf{E}^0$, I^0 and $\mathbf{\Lambda}^0$ to be zero;

while *not converged* **do**

  Update k = k + 1 ;

  Solve for $\mathbf{Q},\mathbf{P},\mathbf{E}$ and I:

$$(10)\qquad (\mathbf{Q}^k,\mathbf{P}^k,\mathbf{E}^k,I^k) = \arg\min_{\mathbf{Q},\mathbf{P},\mathbf{E},I} \mathscr{L}(\mathbf{Q},\mathbf{P},\mathbf{E},I;\mathbf{\Lambda}^{k-1});$$

  Update $\mathbf{\Lambda} = (\mathbf{\Lambda}_Q,\mathbf{\Lambda}_P,\mathbf{\Lambda}_E)$:

$$\mathbf{\Lambda}_Q^k = \mathbf{\Lambda}_Q^{k-1} + c_Q(\mathbf{Q}^k - \nabla\mathbf{E}^k),$$

$$\mathbf{\Lambda}_P^k = \mathbf{\Lambda}_P^{k-1} + c_P(\mathbf{P}^k - \nabla I^k),$$

$$\mathbf{\Lambda}_E^k = \mathbf{\Lambda}_E^{k-1} + c_E(\mathbf{E}^k - \mathbf{P}^k);$$

end

Return $I = I^k$;

Because the variables $\mathbf{Q}$, $\mathbf{P}$, $\mathbf{E}$ and I in $\mathscr{L}(\mathbf{Q},\mathbf{P},\mathbf{E},I;\mathbf{\Lambda}^{k-1})$ are coupled together in the minimization problem (10), it is difficult to solve them simultaneously. We split the minimization problem into four sub minimization problems, and solve them alternatingly until convergence, cf. Algorithm 3.2. Typically we need only one iteration ($L=1$).

We state an important relation, cf. Remark 1, which is used in our algorithm to give us a simpler approach to solve the first two sub-minimization problems of Algorithm 3.2.

Algorithm 3.2: Solve the minimization problem (10).

Input: $\mathbf{Q}^{k-1}$, $\mathbf{P}^{k-1}$, $\mathbf{E}^{k-1}$, I^{k-1} and $\mathbf{\Lambda}^{k-1}$

Output: $\mathbf{Q}^{k}$, $\mathbf{P}^{k}$, $\mathbf{E}^{k}$ and I^{k}

Set $l = 0$;

Initialize $\mathbf{Q}^{k,0} = \mathbf{Q}^{k-1}$, $\mathbf{P}^{k,0} = \mathbf{P}^{k-1}$, $\mathbf{E}^{k,0} = \mathbf{E}^{k-1}$ and $I^{k,0} = I^{k-1}$;

while *not converged and* $l < L$ **do**

  Solve for $\mathbf{Q}$ (the $\mathbf{Q}$-subproblem)

$$\mathbf{Q}^{k,l+1} = \arg\min_{\mathbf{Q}} \mathscr{L}(\mathbf{Q}, \mathbf{P}^{k,l}, \mathbf{E}^{k,l}, I^{k,l}; \mathbf{\Lambda}^{k-1});$$

  Solve for $\mathbf{P}$ (the $\mathbf{P}$-subproblem)

$$\mathbf{P}^{k,l+1} = \arg\min_{\mathbf{P}} \mathscr{L}(\mathbf{Q}^{k,l+1}, \mathbf{P}, \mathbf{E}^{k,l}, I^{k,l}; \mathbf{\Lambda}^{k-1});$$

  Solve for $\mathbf{E}$ (the $\mathbf{E}$-subproblem)

$$\mathbf{E}^{k,l+1} = \arg\min_{\mathbf{E}} \mathscr{L}(\mathbf{Q}^{k,l+1}, \mathbf{P}^{k,l+1}, \mathbf{E}, I^{k,l}; \mathbf{\Lambda}^{k-1});$$

  Solve for I (the I-subproblem)

$$I^{k,l+1} = \arg\min_{I} \mathscr{L}(\mathbf{Q}^{k,l+1}, \mathbf{P}^{k,l+1}, \mathbf{E}^{k,l+1}, I; \mathbf{\Lambda}^{k-1});$$

  Update $l = l + 1$;

end

Return $\mathbf{Q}^{k} = \mathbf{Q}^{k,l}$, $\mathbf{P}^{k} = \mathbf{P}^{k,l}$, $\mathbf{E}^{k} = \mathbf{E}^{k,l}$ and $I^{k} = I^{k,l}$;

REMARK 1. *If A and B are two matrices such that $A = \lambda B$ for some non-negative scalar λ, then we say that A is compatible with B. It is easy to see that $A/|A|_F = B/|B|_F$.*

The Q-subproblem: The first problem is to solve for $\mathbf{Q}$, freezing the other variables, in (10).

$$\mathbf{Q}^* = \arg\min_{\mathbf{Q}} \left\{ g|\mathbf{Q}|_F + \mathbf{\Lambda}_Q \cdot \mathbf{Q} + \frac{c_Q}{2}|\mathbf{Q} - \nabla\mathbf{E}|_F^2 \right\}. \tag{11}$$

We will find a closed form solution to this. The corresponding optimality condition, or the Euler-Lagrange equation, is as follows

$$\frac{g}{c_Q}\frac{\mathbf{Q}^*}{|\mathbf{Q}^*|_F} + \mathbf{Q}^* = \nabla\mathbf{E} - \frac{\mathbf{\Lambda}_Q}{c_Q}.$$

Since g and c_Q are both positive numbers and $|\mathbf{Q}^*|_F$ is a positive number, the matrices $\mathbf{Q}^*$ and $(\nabla\mathbf{E} - \mathbf{\Lambda}_Q/c_Q)$ become compatible in the sense of REMARK 1 because $(g/c_Q\ |\mathbf{Q}^*|_F^{-1} + 1)\mathbf{Q}^* = \nabla\mathbf{E} - \mathbf{\Lambda}_Q/c_Q$. According to which, we can replace the matrix $\mathbf{Q}^*/|\mathbf{Q}^*|_F$ with $(\nabla\mathbf{E} - \mathbf{\Lambda}_Q/c_Q)/|\nabla\mathbf{E} - \mathbf{\Lambda}_Q/c_Q|_F$ in the above equation. Moving the first term to the right hand side we get

$$\mathbf{Q}^* = \left(1 - \frac{g}{c_Q|\nabla\mathbf{E} - \frac{\mathbf{\Lambda}_Q}{c_Q}|_F}\right)\left(\nabla\mathbf{E} - \frac{\mathbf{\Lambda}_Q}{c_Q}\right).$$

Again since we have already seen that in the sense of Remark 1, $\mathbf{Q}^*$ and $(\nabla\mathbf{E} - \mathbf{\Lambda}_Q/c_Q)$ are compatible, the coefficient $\left(1 - g/c_Q\ |\nabla\mathbf{E} - \mathbf{\Lambda}_Q/c_Q|_F^{-1}\right)$ must also be non-negative. Hence

$$\mathbf{Q}^* = \max\left\{0, 1 - \frac{g}{c_Q|\nabla\mathbf{E} - \frac{\mathbf{\Lambda}_Q}{c_Q}|_F}\right\}\left(\nabla\mathbf{E} - \frac{\mathbf{\Lambda}_Q}{c_Q}\right),$$

which is the solution to the $\mathbf{Q}$-subproblem.

The P-subproblem: The second problem is to solve for $\mathbf{P}$, freezing the other variables, in (10).

(12)

$$\mathbf{P}^* = \arg\ \min_{\mathbf{P}}\left\{h|\mathbf{P}| - \hat{\alpha}\mathbf{P}\cdot\mathbf{v}_\Gamma + (\mathbf{\Lambda}_P - \mathbf{\Lambda}_E)\cdot\mathbf{P} + \frac{c_P}{2}|\mathbf{P} - \nabla I|^2 + \frac{c_E}{2}|\mathbf{E} - \mathbf{P}|^2\right\}.$$

We will find a closed form solution to this. The corresponding Euler-Lagrange equation is as follows

$$\frac{h}{c_P + c_E}\frac{\mathbf{P}^*}{|\mathbf{P}^*|} + \mathbf{P}^* = \mathbf{X},$$

where $\mathbf{X} := (c_P\nabla I + c_E\mathbf{E} - \mathbf{\Lambda}_P + \mathbf{\Lambda}_E + \hat{\alpha}\mathbf{v}_\Gamma)/(c_P + c_E)$. Since h, c_P and c_E are positive numbers and $|\mathbf{P}^*|$ is a positive number, the vectors $\mathbf{P}^*$ and $\mathbf{X}$ become compatible in the sense of Remark 1 because $(h/(c_P + c_E)\ |\mathbf{P}^*|^{-1} + 1)\mathbf{P}^* = \mathbf{X}$. According to which, we can replace the vector $\mathbf{P}^*/|\mathbf{P}^*|$ with $\mathbf{X}/|\mathbf{X}|$ in the above equation. Moving the first term to the right hand side we get

$$\mathbf{P}^* = \left(1 - \frac{h}{(c_P + c_E)|\mathbf{X}|}\right)\mathbf{X}.$$

Again since we have already seen that in the sense of Remark 1, $\mathbf{P}^*$ and $\mathbf{X}$ are compatible, the coefficient $(1 - h/(c_P + c_E)\ |\mathbf{X}|^{-1})$ must also be non-negative. Hence

$$\mathbf{P}^* = \max\left\{0, 1 - \frac{h}{(c_P + c_E)|\mathbf{X}|}\right\}\mathbf{X},$$

which is the solution to to the $\mathbf{P}$-subproblem.

The E-subproblem: The third problem is to solve for $\mathbf{E}$, freezing the other variables, in (10).

$$\mathbf{E}^* = \arg\min_{\mathbf{E}}\left\{\mathbf{\Lambda}_E\cdot\mathbf{E} + \frac{c_E}{2}|\mathbf{E} - \mathbf{P}|^2 - \mathbf{\Lambda}_Q\cdot\nabla\mathbf{E} + \frac{c_Q}{2}|\mathbf{Q} - \nabla\mathbf{E}|_F^2\right\}. \tag{13}$$

We will find a closed form solution to this. The corresponding Euler-Lagrange equation is the following.

$$\nabla\cdot(\nabla\mathbf{E}^*) - \frac{c_E}{c_Q}\mathbf{E}^* = \frac{1}{c_Q}\mathbf{\Lambda}_E - \frac{c_E}{c_Q}\mathbf{P} + \frac{1}{c_Q}\nabla\cdot\mathbf{\Lambda}_Q + \nabla\cdot\mathbf{Q},$$

which is a set of two inhomogeneous modified Helmholtz equations, one equation for each component of $\mathbf{E} = (E_1, E_2)$, with the following Neumann boundary conditions,

$$\nabla E_1 \cdot \nu = (\mathbf{Q}_1 + \frac{1}{c_Q}\mathbf{\Lambda}_{Q1}) \cdot \nu, \quad \text{and} \quad \nabla E_2 \cdot \nu = (\mathbf{Q}_2 + \frac{1}{c_Q}\mathbf{\Lambda}_{Q2}) \cdot \nu,$$

where $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are the row vectors of the matrix $\mathbf{Q}$, and $\mathbf{\Lambda}_{Q1}$ and $\mathbf{\Lambda}_{Q2}$ are the corresponding Lagrange multipliers, respectively. ν is the outward unit normal vector on the boundary of the domain. Each equation is solved in the same way as follows.

Solving the inhomogeneous modified Helmholtz equation:

$$\triangle u(x, y) - \lambda u(x, y) = F(x, y), \tag{14}$$

with a Neumann boundary condition and λ a positive scalar, also known as the inhomogeneous modified Helmholtz equation. A fast solver based on discrete cosine transform similar for the Poisson equation, cf. [33, 6] and also cf. [5] for details, is given below.

The discrete version of Laplace operator in (14) is the matrix

$$\begin{bmatrix} -1 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & 1 & -1 \end{bmatrix},$$

with the help of discrete cosine transformation matrix $C \in \mathbb{R}^{N\times N}$ and singular value decomposition, we get the following decomposition,

$$-C^\top \begin{bmatrix} 0 & \\ & \Sigma^2 \end{bmatrix} C,$$

where $\Sigma = \text{diag}(\sigma_1, \cdots, \sigma_{N1})$ is the diagonal matrix with its entries representing the singular values $\sigma_k = 2\sin(\pi k/2/N)$ for $k = 1, 2, \cdots, N-1$. Substituting the decomposition back into (14), we get

$$-uC^\top \begin{bmatrix} 0 & \\ & \Sigma_x^2 \end{bmatrix} C - C^\top \begin{bmatrix} 0 & \\ & \Sigma_y^2 \end{bmatrix} Cu - \lambda u = F.$$

A further transformation by using $\tilde{u} := CuC^\top$ and $\tilde{F} := CFC^\top$ results with

$$-\tilde{u} \begin{bmatrix} 0 & \\ & \Sigma_x^2 \end{bmatrix} - \begin{bmatrix} 0 & \\ & \Sigma_y^2 \end{bmatrix} \tilde{u} - \lambda\tilde{u} = \tilde{F}.$$

The solution to the above equation can be obtained by a direct entrywise division due to linearity of the equation and the non-singularity of the coefficient matrix (non zero λ), giving

$$\tilde{u} = \tilde{F}./M,$$

where M is the $N \times N$ coefficient matrix defined as

$$M = -\begin{bmatrix} 0 & \sigma_{1,x}^2 & \cdots & \sigma_{N-2,x}^2 & \sigma_{N-1,x}^2 \\ \sigma_{1,y}^2 & \sigma_{1,x}^2+\sigma_{1,y}^2 & \cdots & \sigma_{N-2,x}^2+\sigma_{1,y}^2 & \sigma_{N-1,x}^2+\sigma_{1,y}^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sigma_{N-2,y}^2 & \sigma_{1,x}^2+\sigma_{N-2,y}^2 & \cdots & \sigma_{N-2,x}^2+\sigma_{N-2,y}^2 & \sigma_{N-1,x}^2+\sigma_{N-2,y}^2 \\ \sigma_{N-1,y}^2 & \sigma_{1,x}^2+\sigma_{N-1,y}^2 & \cdots & \sigma_{N-2,x}^2+\sigma_{N-1,y}^2 & \sigma_{N-1,x}^2+\sigma_{N-1,y}^2 \end{bmatrix} - \lambda.$$

The solution to (14) is thus

$$u = C^\top((CFC^\top)./M)C.$$

Note that the discrete cosine transformations are easy to implement in MATLAB using commands `dct2` for the forward transform ($C[\cdot]C^\top$) and `idct2` for the inverse transform ($C^\top[\cdot]C$).

The I-subproblem: The fourth problem is to solve for I, freezing the other variables, in (10).

$$I^* = \arg\min_I \left\{ \hat{\theta}|I - I_\Sigma|^2 - \mathbf{\Lambda}_P \cdot \nabla I + \frac{c_P}{2}|\mathbf{P} - \nabla I|^2 \right\}. \tag{15}$$

We will find a fast solver to this. The corresponding Euler-Lagrange equation is as the following.

$$\nabla \cdot (\nabla I^*) - \frac{2\hat{\theta}}{c_P} I^* = \nabla \cdot \mathbf{P} + \frac{1}{c_P}\nabla \cdot \mathbf{\Lambda}_P - \frac{2\hat{\theta}}{c_P} I_\Sigma, \tag{16}$$

with the Neumann boundary condition

$$\nabla I^* \cdot \nu = (\mathbf{P} + \frac{1}{c_P}\mathbf{\Lambda}_P) \cdot \nu, \tag{17}$$

where the ν is the outward unit normal vector on the boundary of the domain. The above equation is not a standard modified Helmholtz equation because the coefficient $2\hat{\theta}/c_P$ is a scalar function. We solve this equation with the conjugate gradient method, cf. e.g. [24], in MATLAB using the function `pcg`. Using the diagonal preconditioner is the simplest, yet effective in this case, cf. FIG. 2, a typical convergence result is shown.

4. Numerical experiments. In this section we present our experiments, on the choice of our vector matching term and the regularizers in the proposed model, on the effectiveness of the model on simple geometries, on its being able to accurately represent structures like edges and geometries, and finally on its effectiveness on real 3D maps with very few level lines comparing it with the state of the art model.

We use the augmented Lagrangian method of section 3 for the solution. The parameters of the augmented Lagrangian method are set as $c_Q = 1$, $c_P = 1$ and $c_E = 1$ throughout our experiment, unless otherwise stated. The information needed are level curves or level lines, often provided as point clouds. In practice, these points are normally oriented, and is therefore easy to construct level lines from them. In case they are not oriented, we need to determine their orientations. To do that we

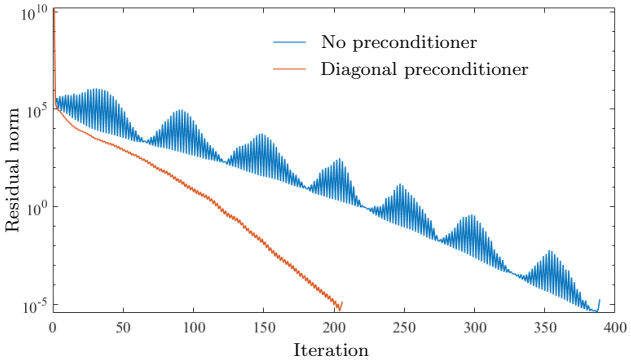

FIG. 2. *Convergence of the preconditioned conjugate gradient method solving (16)-(17). Residual norm against iteration. The diagonal preconditioner is very effective in this case.*

simply connect the nearest given points with same level value to form the the level lines. In some cases the nearest points cannot be identified, for which, we first run the isotropic model ($\alpha = 0$ in (5)), corresponding to the model without vector matching. Once the isotropic surface is obtained, the MATLAB function called `contour` is used to get the level lines which are then used as guidelines to find the order of the given points. A threshold is used to determine the connectivity between the points, once exceeded the level lines are then considered disconnected.

On the vector matching term. With this experiment we justify our choice of vector matching term $V(\nabla I, \mathbf{v}_\Gamma)$ in the minimization (1), from the two options (2) and (4). To see the difference, we choose the following simple example: level lines (with height value) parallel to the y-axis, cf. FIG. 3 and FIG. 4 (blue lines on the floor or red lines on the surface). Vector $\mathbf{v}_\Gamma$ is defined only on the lowest and highest level lines. The regularization and data fidelity parameters are kept the same as $g = 1$, $h = 0$ and $\theta = 10^5$ throughout this experiment.

We consider the tangent vector matching first, model (3), this is illustrated in FIG. 3. As mentioned earlier, there are two solutions, either $\nabla I^* = \mathbf{0}$, which is reflected in the last two sub-figures, or $\nabla I^* \perp \mathbf{v}_\Gamma$ with $\nabla I^* \neq \mathbf{0}$, which is reflected in the first two sub-figures. There are basically two shapes that can be obtained through the model (3), cf. FIG. 3, flat structures being created between the level lines where vectors are given, and parabolic structures otherwise. $\mathbf{v}_\Gamma$ is chosen so that it is not parallel to tangent $\boldsymbol{\tau}_\Gamma$ (to the level lines) in the last two sub-figures.

We now consider the normal vector matching term, giving as the model (5). The results are shown in FIG. 4. As shown in the figure, varying the parameter α, we get different shapes, from parabolic to flattened. Note that, based on the normal vector matching, we are able to generate flat structures even when the first order regularizer has been inactive (corresponding to $h = 0$), see the last sub-figure. The value of the parameter α for the vector matching term is set equal to 1.5, 2, -0.1 and -0.2 respectively in FIG. 4. A second example showing similar results is presented in FIG. 5, where cones of different shapes, from concave to convex, have been obtained by varying the parameter α ($\alpha = -1, 0, 1$, and 2).

On the regularizer terms. The second experiment is to study the effect of different orders of the regularizer. To do this we use only one regularizer term in our

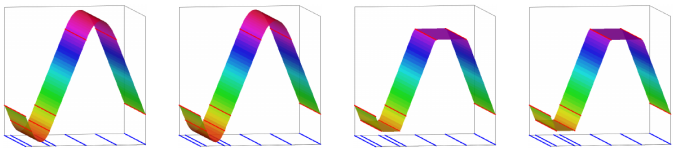

FIG. 3. *Illustrating the use of tangent vector matching term (2) in the model (1) to get (3). Surfaces are reconstructed from the level lines (the blue lines on the floor, or the red lines on the surface) with the vector $\mathbf{v}_\Gamma$ prescribed only on the level lines with the largest and the smallest height values. Here, only the second order regularization is used. $\boldsymbol{\tau}_\Gamma$ is the tangent vector of the given level lines. The first two figures correspond to the case $\nabla I^* \perp \mathbf{v}_\Gamma$ (no matter how big the value of α is, the surface is the same isotropic surface) and the last two figures correspond to the case $\nabla I^* = 0$ (in which case the surface is a flattened structure).*

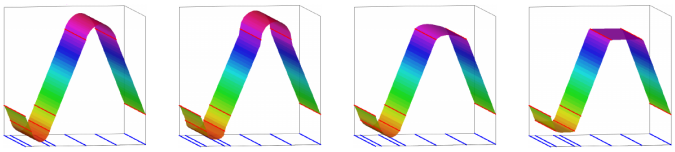

FIG. 4. *Illustrating the use of normal vector matching (4) in the model (1), giving (5). Surfaces are reconstructed from given level lines (the blue lines on the floor, or the red lines on the surface), with unit normal vectors prescribed only on the level lines with the largest and the smallest height values. Here, only the second order regularizer is used. The different shapes are reconstructed using different values of α.*

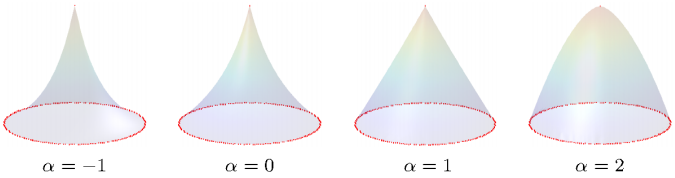

FIG. 5. *Illustrating the use of normal vector matching (4) in our model (1), giving (5). Surfaces are reconstructed from one level line (the red line on the floor) and one height in the center (the red point on the top), plus unit normal vectors prescribed only on the level line with the smallest height value. Here, only the second order regularizer is used. Cones with different shapes, from concave to convex, are reconstructed using different values of α.*

model and fix the vector matching term. The following minimization, same as (6), is considered

$$\min_I g|\nabla^k I|_{F\ (\Omega)} - \alpha \nabla I \cdot \mathbf{v}_{\Gamma\ (\Gamma)} + \theta |I - I_\Sigma|^2_{2\ (\Sigma)},$$

where the first term is the k^{th} order total variation (TV), the second term is the normal vector matching term, and the third term is the data fidelity term. g, α and θ are the scalar parameters. Γ and Σ are the set points where the unit normal vector and elevation/height value are given, respectively. For clarity, we apply the model to

one space dimension (1D). The results are shown in FIG. 6, each column represents a fixed order from $k = 1$ to 5, and each row corresponds to a particular test. The height values are given at red points and the vectors are marked with red arrows. Note that, in 1D, the unit vector is the value one with a sign for the direction, i.e. ± 1. In FIG. 6, the positive sign corresponds to the arrow pointing to the right and vice versa. The parameters are kept the same throughout this experiment as $g = 1$ for the regularizer, $\alpha = 1$ for the vector matching term, and $\theta = 10$ for the data fidelity.

The experiment shows that orders higher than two, when combined with the normal vector matching, may induce oscillation in the shape of the reconstructed surface, as well as unbounded surface at the boundary. A combination of the first order and the second order regularizer is enough to reconstruct most shapes. This will see in the following experiments.

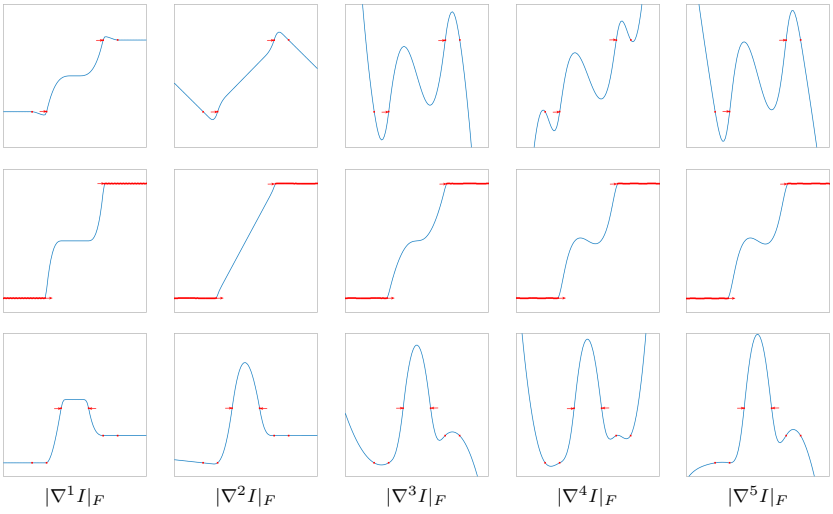

FIG. 6. *Showing The effect of using different order of regularizer in the reconstruction in 1D. The profiles are reconstructed using (6); sparse data given are the height values (in red), and gradient direction using arrows (in blue). Each row corresponds to a specific test, and the columns corresponds to the orders from 1 to 5 of the regularizer. The first order results in flattened structure while the second order favours slopes in the structure. The higher order regularizer introduces oscillation in the solution, and makes it unbounded at the boundary.*

The proposed model on regular structures. The third experiment is to validate the proposed model (5) where both regularizers and the normal vector matching are used. We study its effectiveness in accurately representing simple geometries like edges and corners.

We first test the proposed model on a 1D signal with the mixed shape of sine and rectangular waves, cf. FIG. 7. Only a few isolated points (red points) with height values and unit normal vectors (± 1) are given. The normal vector matching is only used in the last sub-figure, with the parameter $\alpha = 1.5$. Throughout this test, the parameter $h = 1.5$ for the first order total variation (TV), i.e., $|\nabla^1 I|_F$, and the parameter $g = 4$ for the second order TV, i.e., $|\nabla^2 I|_F$.

The result as shown in FIG. 7, demonstrates that the proposed model (5) is able

to accurately recover the shape with corners. While the first order TV favors the staircase structures and the second TV preserves sloped structures, the combination of the first order and the second order TVs without vector matching reconstructs both flattened and sloped structures quite well. However it still fails to be accurate on the top and in valleys.

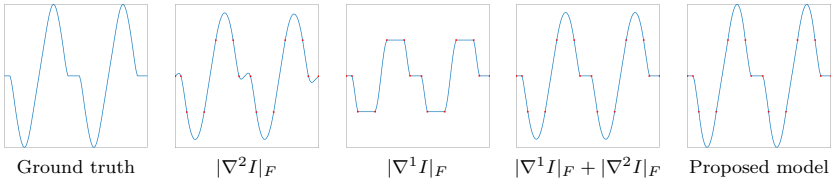

FIG. 7. *Showing the effect of using vector matching in combination with both first order and second order total variation regularizer in the proposed model (5), as compared to the second, third and fourth figures from left where the vector matching has not been used.*

Next, we apply our model in 2D, on a geometry that looks like the Maya pyramid, cf. FIG. 8. There are three level lines available (red lines on the surface) with both height values and unit normal vectors, either given or extracted from the geometry of the underlying level lines. Note that, to accurately reconstruct the structures, we need to adjust the parameters of each term in our model (5). The data fidelity parameter θ, in the experiment shown in FIG. 8, is chosen to $\theta = 10^5$. The regularizer parameter $g = 10$ in the second sub-figure, while $h = 1$ in the third sub-figure. For sloped structures such as the walls of the base and the walls of the top, we simply use only the second order regularizer for the last two sub-figures, where $h = 0$. In these two sub-figures, $g = 10$ and $g = 1$ for the walls of the base and the walls of the top, respectively. For flattened structures like the horizontal area of the pyramid, we use only the first order regularizer in the last two sub-figures, where $g = 0$ and $h = 10$. The normal vector matching is only used in the last sub-figure, where $\alpha = 1$ along the level lines except the ones are intersection between the base and the top of the pyramid, where $\alpha = 11$. A higher value of the parameter α setting on this intersection is to get sharper edges for the top.

Similar to the 1D test showing in FIG. 7, the first order TV favors the flat structures and the second TV preserves sloped structures. The combination of the first order and the second order TVs without vector matching reconstructs both flat and sloped structures but fails to accurately represent edges and corners. The proposed model (5) reconstructs the Maya pyramid more accurately representing the edges and corners.

The third test is to justify the capability of the proposed model (5) in controlling of the reconstructed shape by adjusting the vector matching parameter α. The results are shown in FIG. 9. Different to the previous test, there is no information given now at the center of the 2D domain, cf.FIG. 9 and FIG. 8. The parameters are kept the same as those in the last sub-figure of FIG. 8 except α on the innermost given level line (in red), cf. FIG. 8 where $\alpha = 2.5$, $\alpha = 2$, $\alpha = 0$ and $\alpha = -2$ for each test respectively. The test shows that varying parameter α changes the shape, from a convex shape to a concave structure.

The fourth test is on a semi-sphere. For each test we have different number of contours (the red lines on the surface), cf. FIG. 10. Both the height value and the unit normal vectors are given on these contours. In this experiment, the regularizer

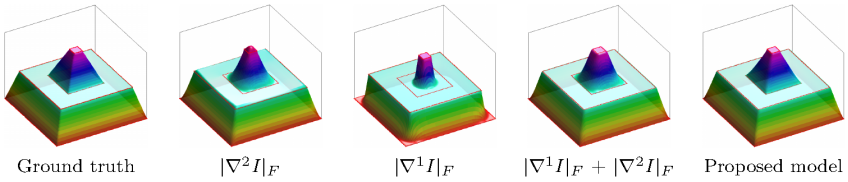

FIG. 8. *2D illustration showing the effect of using vector matching in combination with both first order and second order total variation regularizer in the proposed model (5), in the last figure, as compared to the first, second and third figures from left where the vector matching has not been used.*

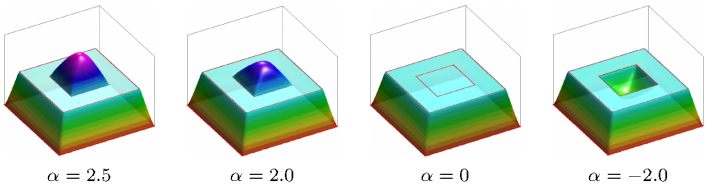

FIG. 9. *2D illustration showing how easily we can control the reconstructed surface through adjusting the vector matching parameter α in the proposed model (5).*

parameters have been $g = 1$ and $h = 0$, and the data fidelity parameter has been $\theta = 10^5$. The vector matching parameter has been $\alpha = 0.5$ along the contours, except the one with the largest radius, where $\alpha = 2.85$.

As shown here, even with only one single contour, as in the first sub-figure of FIG. 10, the proposed model (5) is still able to reconstruct the surface reasonably well. As we increase the number of contours, the reconstructed surface becomes more and more close to the perfect semi-sphere.

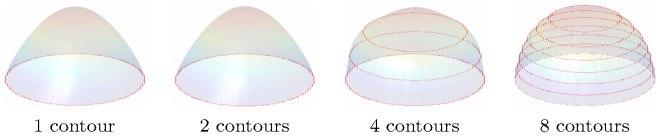

FIG. 10. *Showing the effect of level lines in the proposed model (5), increasing the number of contours from left to right. With only one level line (the leftmost figure), the parameter α alone is enough to control the shape of the surface. The surface improves as the number of level lines increases.*

Real map I. In this experiment we apply our model to real maps with clear crease (kinks) and clear valleys, where we only have a few level lines to reconstruct the map from, cf. FIG. 12. Later we show how the reconstruction improves as the level lines increase, cf. FIG. 13.

To see the effectiveness of our model we first test it without the vector matching ($\alpha = 0$), cf. FIG. 11. In this test there are only three level lines available (blue lines on the floor or red lines on the surface), where the height values are given. The first two extensions are using the first order total variation $|\nabla^1 I|_F$ and the second order

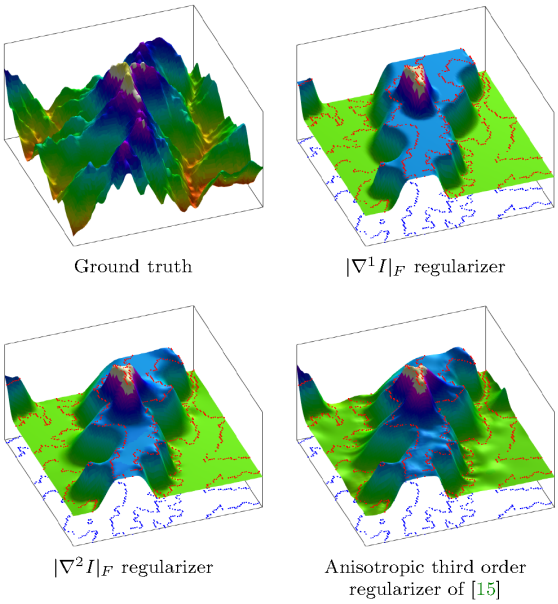

FIG. 11. *Illustrating the challenges in reconstruction a real 3D map (mountain and valley) with only few level lines. The reconstructions here are from only three level lines (in blue on the plane, in red on the surface). The first two reconstructions, the upper-right and the lower-left sub-figure, are based on the first order and the second order total variation regularizer, respectively, but without any vector matching. The last reconstruction (the lower-right sub-figure) is based on the anisotropic third order regularizer proposed in [15, $R_1^{(3)}(u)$ on page 5].*

total variation $|\nabla^2 I|_F$, respectively. The reconstructed surfaces in both runs show difficult to reconstruct the valley in the ground truth. $\theta = 10^5$, $g = 0$ and $h = 10$ for the top right sub-figure, $g = 10$ and $h = 0$ for the bottom left sub-figure. The first one shows typical staircase effect while the second one favors slope. However in the valley lacking of information the only possible solution is the flattened structure. This is what we have already seen in previous experiments.

In the last sub-figure of FIG. 11, we present the reconstruction using the anisotropic third order regularizer model of [15, $R_1^{(3)}(u)$ in model (4) on page 5]. The reconstructed surface improves a bit area still fails in reconstructing the valley, we believe it is because the level lines are too few to have enough similarity between the level lines which the algorithm requires, making it difficult to reconstruct the valley. For the experiment, we have solved the problem in [15] using our algorithm based on the augmented Lagrangian.

We now apply our model (5) involving the normal vector matching term, on the three level lines. The results are shown in FIG. 12. The unit normal vector $\mathbf{v}_\Gamma$ needed in the model are extracted from the level line C_Γ. Two possible choices $\hat{\mathbf{v}}_\Gamma = \mathbf{v}_\Gamma \cdot |\hat{\mathbf{v}}_\Gamma| = \pm\partial_t^\perp C_\Gamma$ where $\partial_t C_\Gamma$, with C_Γ being the parametrization of the level line, denotes the tangent vector along C_Γ obtained via finite differences between discrete points along the level line.

In the reconstruction, the correct sign of the unit normal vector $\mathbf{v}_\Gamma$ is not known

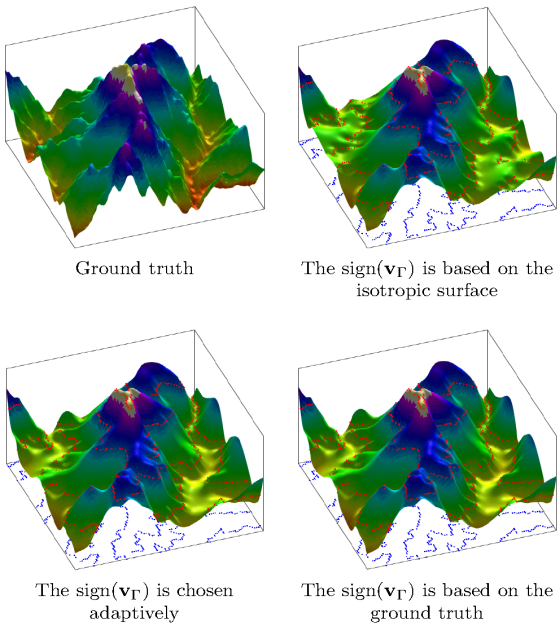

FIG. 12. *Reconstructions of the same real 3D map (mountain and valley) as in* FIG. 11, *from its three level lines (blue lines on the floor or red lines on the surface), using our proposed model (5). The different sub-figures here correspond to the different choices of sign for the unit normal vector* $\mathbf{v}_\Gamma$ *in the model.*

in advance. To determine the sign we first solve the isotropic model, i.e. the model (5)) with $\alpha = 0$, in order to get the ∇I. We then look at the value of $\varrho := \mathbf{v}_\Gamma \cdot \nabla I$. If $\varrho < 0$, we set $\mathbf{v}_\Gamma = -\mathbf{v}_\Gamma$. The surface presented in the upper-right corner in FIG. 12 has been reconstructed in this way, showing already a significant improvement over the reconstructions presented in FIG. 11, where we see a more accurate representation of the mountain, as well as the two rivers in the image.

This is further improved, as we update the sign($\mathbf{v}_\Gamma$) gradually instead of everywhere at once, cf. the figure in the lower-left corner of FIG. 12. We let Γ_ε be the subset of Γ where we update the sign($\mathbf{v}_\Gamma$). In each iteration, we look at the absolute value of the directional derivative, i.e. $\varepsilon(x) = |\mathbf{v}_\Gamma(x) \cdot \nabla I(x)|$, and compare it with a threshold $\varepsilon_{threshold}$. Starting with the empty set $\Gamma_\varepsilon = \varnothing$, we gradually include parts of Γ whenever $\varepsilon(x)$ is found to be larger than the threshold in those parts. In other words, if $\varepsilon(x) > \varepsilon_{threshold}$ on Γ then update $\Gamma_\varepsilon \leftarrow \Gamma_\varepsilon \cup \{x \in \Gamma : \varepsilon(x) > \varepsilon_{threshold}\}$. At start we have the isotropic case, i.e. $\alpha = 0$. $\alpha \neq 0$ on Γ_ε. If $\varrho(x) < 0$ on $x \in \Gamma_\varepsilon$ then we set $\mathbf{v}_\Gamma = -\mathbf{v}_\Gamma$ at that point. As the iteration continues, we expect $\Gamma_\varepsilon = \Gamma$. Gradually updating the level lines has given even better reconstruction which is very close to what we could achieve if sign($\mathbf{v}_\Gamma$) is extracted from the ground truth, cf. the figure in the lower-right corner of FIG. 12).

As we can see, even with a few number of level lines, we have an almost perfect reconstruction. However, it becomes better as we increase the number of level lines, cf. FIG. 13.

In figures 12–13, the parameters for the regularizers has been $g = 1$ and $h = 0$, the data fidelity parameter has been $\theta = 10^5$, and the matching term parameter has been $\alpha = 30$. In addition, $h = 0.1$ has been on the boundary. $\varepsilon_{threshold} = 0.2$ has been used to determine the sign($\mathbf{v}_\Gamma$).

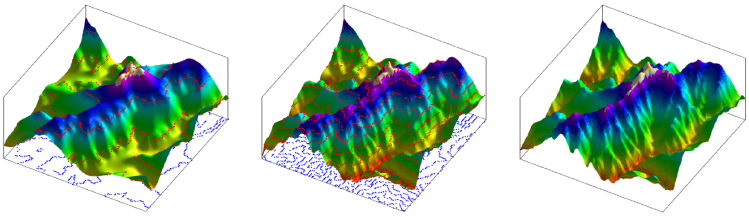

FIG. 13. *The real 3D map seen from the side, with increasing number of level lines from left to right. The surface reconstruction improves as the number of level lines increases.*

Real map II. In the experiment we compare our model with the third order anisotropic model of [15]. For the experiment we use two of the examples from [15], reproducing their results using our algorithm based on the augmented Lagrangian. The results are presented in FIG. 14, one example in each row. For our model, $\theta = 10^5$ (data fidelity), $\alpha = 10$ (vector matching) and $g = 0.1$ (the second order regularizer). For both, $h = 0$ inside and $h = 2$ on the boundary. $\varepsilon_{threshold} = 0.1$ has been used to determine sign($\mathbf{v}_\Gamma$). The parameters for augmented Lagrangian are set as $c_Q = 20$, $c_P = 1$ and $c_E = 50$ in this experiment. As seen from the figure, with enough level lines, both models perform well in capturing the anisotropy, however, our model managed to capture even the small variations, details like the small hill top etc.. This is because the model uses the geometry of the level lines.

Real map III. In this experiment we apply our model to reconstruct mixed structures with regular and irregular geometries. The top left sub-figure of FIG. 15 is the ground truth, which we reconstruct from its three level lines. Just as before, we apply our model with or without the vector matching and we observe the clear importance of having the vector matching. For the experiment, $\theta = 10^5$ (data fidelity) on Σ and $h = 0.1$ at boundary. For the top right sub-figure $g = 10^{-2}$ and $h = 0$ while $g = 0$ and $h = 50$ for the bottom left sub-figure. In the last sub-figure, $g = 10^{-2}$ for the area with mountain, $h = 50$ and $h = 5$ for the flat base and tip of the building, respectively. $\alpha = 10$ and $\varepsilon_{threshold} = 0.2$ has been used to determine sign($\mathbf{v}_\Gamma$).

5. Conclusions. We have proposed a model consisting of a vector matching term (normal vector matching) to account for the anisotropy, together with a first-order and a second-order total variation regularizer (isotropic) under a fidelity constraint on the height. The model is able to effectively capture the irregularity along the given level lines, e.g., kinks or creases along level lines, and is able to recover surfaces from only a small number of level lines.

We have proposed an effective way, based on the augmented Lagrangian method, to solve the model. In our algorithm, each sub-problem has either a closed form solution or a fast solver. We have derived closed form solution for the minimization problems containing L^1 term, provided a simple approach to derive the solution. For the inhomogenous modified Helmholtz equation (IMHE) with constant coefficient, we

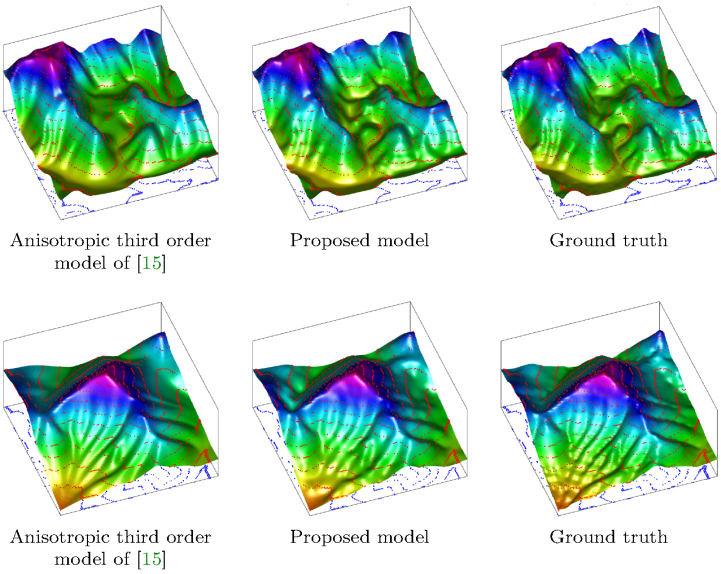

FIG. 14. *Comparing the proposed model with the anisotropic third order model proposed in [15, $R_1^{(3)}(u)$ on page 5] on two different real 3D maps from the same paper, one map in each row, with eleven level lines.*

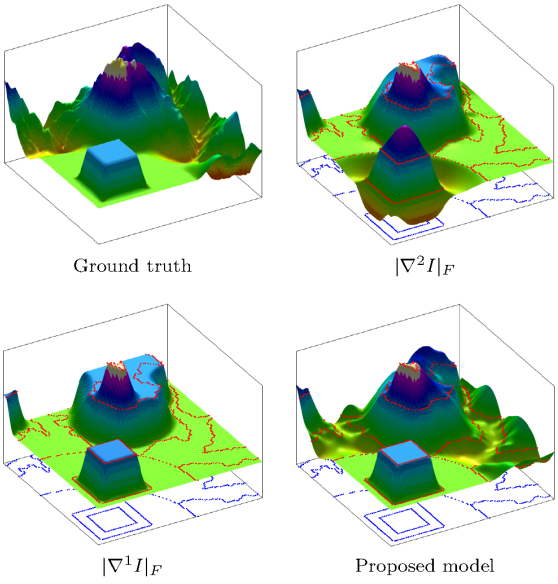

FIG. 15. *Illustrating the effectiveness of the proposed model (5) in reconstructing real 3D map with mixed structures, e.g. mountains, valleys, and buildings, from few level lines. The reconstructions here correspond to three level lines.*

have given a closed form solution based on the discrete cosine transform. For the IMHE with variable coefficient (scalar function), we use the preconditioned conjugate gradient method, using the simplest yet effective preconditioner, the diagonal preconditioner.

The choice of parameters in our model plays a crucial role in the reconstruction. This choice may be turned automatically through some machine learning algorithm. This is a topic of future work.

Acknowledgments. We thank Carola-Bibiane Schönlieb and Jan Lellmann for the fruitful discussions in the beginning of this project, and particularly to Jan Lellmann for verifying our reconstruction in FIG. 11 (the lower-right sub-figure) using their model.

REFERENCES

[1] A. AGRAWAL, R. RASKAR, AND R. CHELLAPPA, *What is the range of surface reconstructions from a gradient field?*, in European conference on computer vision, Springer, 2006, pp. 578–591.

[2] L. ALVAREZ, F. GUICHARD, P.-L. LIONS, AND J.-M. MOREL, *Axioms and fundamental equations of image processing*, Archive for rational mechanics and analysis, 123 (1993), pp. 199–257.

[3] J. C. CARR, W. R. FRIGHT, AND R. K. BEATSON, *Surface interpolation with radial basis functions for medical imaging*, IEEE transactions on medical imaging, 16 (1997), pp. 96–107.

[4] V. CASELLES, J.-M. MOREL, AND C. SBERT, *An axiomatic approach to image interpolation*, IEEE Transactions on image processing, 7 (1998), pp. 376–386.

[5] C. A. ELO, *Image denoising algorithms based on the dual formulation of total variation*, master's thesis, The University of Bergen, 2009.

[6] C. A. ELO, A. MALYSHEV, AND T. RAHMAN, *A dual formulation of the tv-stokes algorithm for image denoising*, in International Conference on Scale Space and Variational Methods in Computer Vision, Springer, 2009, pp. 307–318.

[7] R. FRANKE, *Scattered data interpolation: tests of some methods*, Mathematics of computation, 38 (1982), pp. 181–200.

[8] R. T. FRANKOT AND R. CHELLAPPA, *A method for enforcing integrability in shape from shading algorithms*, IEEE Transactions on pattern analysis and machine intelligence, 10 (1988), pp. 439–451.

[9] Y. GINGOLD, T. IGARASHI, AND D. ZORIN, *Structured annotations for 2d-to-3d modeling*, in ACM SIGGRAPH Asia 2009 papers, ACM Press, 2009, pp. 1–9.

[10] R. GLOWINSKI AND P. LE TALLEC, *Augmented Lagrangian and operator-splitting methods in nonlinear mechanics*, SIAM, 1989.

[11] J. HAHN, J. QIU, E. SUGISAKI, L. JIA, X.-C. TAI, AND H. S. SEAH, *Stroke-based surface reconstruction*, Numerical Mathematics: Theory, Methods and Applications, 6 (2013), pp. 297–324.

[12] J. HAHN, X.-C. TAI, S. BOROK, AND A. M. BRUCKSTEIN, *Orientation-matching minimization for image denoising and inpainting*, International journal of computer vision, 92 (2011), pp. 308–324.

[13] T. IGARASHI, S. MATSUOKA, AND H. TANAKA, *Teddy: a sketching interface for 3d freeform design*, in ACM SIGGRAPH 2006 Courses, ACM Press, 2006, pp. 11–es.

[14] O. A. KARPENKO AND J. F. HUGHES, *Smoothsketch: 3d free-form shapes from complex sketches*, in ACM SIGGRAPH 2006 Papers, ACM Press, 2006, pp. 589–598.

[15] J. LELLMANN, J.-M. MOREL, AND C.-B. SCHÖNLIEB, *Anisotropic third-order regularization for sparse digital elevation models*, in International Conference on Scale Space and Variational Methods in Computer Vision, Springer, 2013, pp. 161–173.

[16] W. G. LITVINOV, T. RAHMAN, AND X.-C. TAI, *A modified tv-stokes model for image processing*, SIAM Journal on Scientific Computing, 33 (2011), pp. 1574–1597.

[17] S. MASNOU AND J.-M. MOREL, *Level lines based disocclusion*, in Proceedings 1998 International Conference on Image Processing. ICIP98 (Cat. No. 98CB36269), IEEE, 1998, pp. 259–263.

[18] J. MEINGUET, *Surface spline interpolation: basic theory and computational aspects*, in Approximation Theory and Spline Functions, Springer, 1984, pp. 127–142.

[19] T. MEYER, *Coastal elevation from sparse level curves*, Summer project under the guidance of

T. Wittman, A. Bertozzi, and A. Chen, UCLA, (2011).

[20] D. Meyers, S. Skinner, and K. Sloan, *Surfaces from contours*, ACM Transactions On Graphics (TOG), 11 (1992), pp. 228–258.

[21] L. Mitas and H. Mitasova, *Spatial interpolation*, Geographical information systems: principles, techniques, management and applications, 1 (1999).

[22] A. Nealen, T. Igarashi, O. Sorkine, and M. Alexa, *Fibermesh: designing freeform surfaces with 3d curves*, in ACM SIGGRAPH 2007 papers, ACM Press, 2007, pp. 41–es.

[23] H.-S. Ng, T.-P. Wu, and C.-K. Tang, *Surface-from-gradients without discrete integrability enforcement: A gaussian kernel approach*, IEEE Transactions on pattern analysis and machine intelligence, 32 (2009), pp. 2085–2099.

[24] J. Nocedal and S. Wright, *Numerical optimization*, Springer Science & Business Media, 2006.

[25] L. Olsen, F. F. Samavati, M. C. Sousa, and J. A. Jorge, *Sketch-based mesh augmentation.*, in SBM, 2005, pp. 43–52.

[26] L. Olsen, F. F. Samavati, M. C. Sousa, and J. A. Jorge, *Sketch-based modeling: A survey*, Computers & Graphics, 33 (2009), pp. 85–103.

[27] N. Petrovic, I. Cohen, B. J. Frey, R. Koetter, and T. S. Huang, *Enforcing integrability for surface reconstruction algorithms using belief propagation in graphical models*, in Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, vol. 1, IEEE, 2001, pp. I–I.

[28] M. Prasad and A. Fitzgibbon, *Single view reconstruction of curved surfaces*, in 2006 IEEE computer society conference on computer vision and pattern recognition (CVPR'06), vol. 2, IEEE, 2006, pp. 1345–1354.

[29] T. Rahman, X.-C. Tai, and S. Osher, *A TV-Stokes denoising algorithm*, in International Conference on Scale Space and Variational Methods in Computer Vision, Springer, 2007, pp. 473–483.

[30] T. Simchony, R. Chellappa, and M. Shao, *Direct analytical methods for solving poisson equations in computer vision problems*, IEEE transactions on pattern analysis and machine intelligence, 12 (1990), pp. 435–446.

[31] X.-C. Tai, S. Borok, and J. Hahn, *Image denoising using tv-stokes equation with an orientation-matching minimization*, in International Conference on Scale Space and Variational Methods in Computer Vision, Springer, 2009, pp. 490–501.

[32] X.-C. Tai and C. Wu, *Augmented lagrangian method, dual methods and split bregman iteration for rof model*, in International conference on scale space and variational methods in computer vision, Springer, 2009, pp. 502–513.

[33] C. Van Loan, *Computational frameworks for the fast Fourier transform*, SIAM, 1992.

[34] B. Wu, T. Rahman, and X.-C. Tai, *Sparse-data based 3d surface reconstruction for cartoon and map*, in International Conference on Imaging, Vision and Learning based on Optimization and PDEs, Springer, 2016, pp. 47–64.

[35] C. Wu, J. Zhang, and X.-C. Tai, *Augmented lagrangian method for total variation restoration with non-quadratic fidelity*, Inverse Problems & Imaging, 5 (2011), p. 237.

[36] T.-P. Wu, C.-K. Tang, M. S. Brown, and H.-Y. Shum, *Shapepalettes: interactive normal transfer via sketching*, in ACM SIGGRAPH 2007 papers, ACM Press, 2007, pp. 44–es.

[37] R. C. Zeleznik, K. P. Herndon, and J. F. Hughes, *Sketch: An interface for sketching 3d scenes*, in ACM SIGGRAPH 2006 Courses, ACM Press, 2006, pp. 9–es.

[38] L. Zhang, G. Dugas-Phocion, J.-S. Samson, and S. M. Seitz, *Single-view modelling of free-form scenes*, The Journal of Visualization and Computer Animation, 13 (2002), pp. 225–235.

Paper C

Multidimensional TV-Stokes for Image Processing

Bin Wu, Xue-Cheng Tai, and Talal Rahman *Submitted. arXiv:2009.11971*

MULTIDIMENSIONAL TV-STOKES FOR IMAGE PROCESSING

BIN WU*, XUE-CHENG TAI†, AND TALAL RAHMAN*

Abstract. A complete multidimential TV-Stokes model is proposed based on smoothing a gradient field in the first step and reconstruction of the multidimensional image from the gradient field. It is the correct extension of the original two dimensional TV-Stokes to multidimensions. Numerical algorithm using the Chambolle's semi-implicit dual formula is proposed. Numerical results applied to denoising 3D images and movies are presented. They show excellent performance in avoiding the staircase effect, and preserving fine structures.

Key words. Total variation, Denoising, Restoration

1. Introduction. As one of the major techniques in computer vision, image processing is widely used in the community. One of the important tasks of image processing is image denoising which is used for removing noises from corrupted images. It plays a crucial and distinguished role, e.g. being used as a pre-step to any processing or analysis of the images. The noise in a corrupted image may depend on the type of instrument used, the procedure applied, and the environment in which the image is taken. Noises are characterized as Gaussian or salt-pepper, additive or multiplicative [3]. We consider only the Gaussian noise in this paper, which is additive.

Searching for a more general and robust algorithm to handle the problem of image denoising has been one of the major interests in the image processing community. With the pioneering work of Rudin, Osher and Fatemi (cf. [18]) based on the total variation minimization for image denoising, published in 1992, their model has been widely used as the basis for most denoising algorithms because of its elegant mathematics, simplicity of implementation, and effectiveness. Allthough the model is able to preserve edges in the image, which is important for an image, it has the limitation that it exhibits staircase effect or produces patchy images, cf. [16]. A number of algorithms now exist that are based on the total variation minimization, which try to overcome this limitation, including high order algorithms, cf. [13, 21], the training based algorithms, cf. [5], iterative regularization algorithms, cf. [15, 4, 14], and algorithms based on the TV-Stokes, cf. [20, 16, 12, 9, 8, 7]. There are also excellent algorithms that are not based on the variational approach, e.g., the non-local algorithms, cf. [1], the block matching algorithm, cf. [6], and the low rank algorithm, cf. [11].

The purpose of this paper is to propose a model which can be used for the data of any dimensions, e.g. 3D medical data, 2D video, and dynamic volume data. The model is expected to be simple but effective with only a small number of parameters. The TV-Stokes model is the right candidate for this. We extend the TV-Stokes model from 2D to an arbitrary dimensions, and we call it the multi-dimensional TV-Stokes. Even though the paper addresses the task of denoising, The model can be extended for use in other data restoration tasks, like image inpainting, image compression, etc. .

The structure of the paper is organized as follows. In Section 2, we present the first step of the multi-dimensional TV-Stokes and its dual formulation. is addressed

*Department of Computer Science, Electrical Engineering and Mathematical Sciences, Western Norway University of Applied Sciences, Inndalsveien 28, 5063 Bergen, Norway (bin.wu@hvl.no, talal.rahman@hvl.no).

†Department of Mathematics, University of Bergen, Allégaten 41, 5007 Bergen, Norway (xuecheng.tai@uib.no).

accompanying with related analysis. In the Section 3, Chambolle's semi-implicit iteration is presented for the minimization. The key part of the multi-dimensional TV-Stokes is the orthogonal projection onto the constrained spaces, which is presented in Section 4, based on the singular value decomposition of the finite difference matrix D. Section 5 gives the second step of the proposed model, height map reconstruction using a surface fitting technique. Numerical results are presented in Section 6, comparing the proposed model with the Rudin-Osher-Fatemi (ROF) model.

2. The vector field smoothing step of the TV-Stokes. The first step of the TV-Stokes is the vector field smoothing step. Its required that the desired smoothed vector field is a gradient field, that is $\boldsymbol{g} = \nabla u$ where $\boldsymbol{g} \in \mathbb{R}^{N_1 \times \ldots N_d \times d}$ and $u \in \mathbb{R}^{N_1 \times \ldots N_d}$ is a d-dimensional grayscale image. In the case $d = 2$, it is the same as requiring $\nabla \times \boldsymbol{g} = 0$ for the gradient field, which is equivalent to $\nabla \cdot \boldsymbol{g}^{\perp} = 0$, which is exactly the constraint used in the TV-Stokes model for 2D.

Let $u \in \mathbb{R}^{N_1 \times \cdots \times N_d}$ denote the components of a d-dimensional grayscale image with components $u_{i_1 \ldots i_d}, i_j = 1, \ldots, N_j$ for $j = 1, \ldots, d$. When a linear transformation defined by a matrix $T \in \mathbb{R}^{N_i \times N_i}$ is applied to the i-th dimension of u, its result will be denoted by $T_{\times i} d$. For instance,

$$(T_{\times 2} u)_{j i_1 i_3} = \sum_{i_2} T_{j i_2} u_{i_1 i_2 i_3}.$$

The gradient field of u is $\boldsymbol{g} = (g_1, g_2, \ldots, g_d) = \nabla u \in \mathbb{R}^{N_1 \times \cdots \times N_d \times d}$, where ∇ signifies a discrete gradient operator with homogeneous Neumann boundary conditions. Namely, the gradient operator $\nabla u = (D_{\times 1} u, D_{\times 2} u, \ldots, D_{\times d} u)$ is determined by means of the following $N_i \times N_i$ differentiation matrix

$$D = \begin{bmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \\ & & & 0 & 0 \end{bmatrix}. \tag{1}$$

Given a d-dimensional image $u^o \in \mathbb{R}^{N_1 \times \cdots \times N_d}$ corrupted by a Gaussian noise, a constrained variant of the Rudin-Osher-Fatemi model [18] consists in finding a gradient field $\boldsymbol{g} = (g_1, \ldots, g_d) \in \mathbb{R}^{N_1 \times \cdots \times N_d \times d}$ satisfying the minimization problem

$$\min_{\substack{\boldsymbol{g} = \nabla u \\ u \in \mathbb{R}^{N_1 \times \cdots \times N_d}}} \left[\|\nabla \boldsymbol{g}\|_1 + \frac{1}{2\lambda} \|\boldsymbol{g} - \boldsymbol{g}^o\|_2^2 \right], \tag{2}$$

where $\lambda > 0$ is a suitable scalar parameter, $\boldsymbol{g}^o = \nabla u^o$ and

$$h = \nabla \boldsymbol{g} = \begin{pmatrix} D_{\times 1} g_1 & D_{\times 2} g_1 & \cdots & D_{\times d} g_1 \\ D_{\times 1} g_2 & D_{\times 2} g_2 & \cdots & D_{\times d} g_2 \\ \vdots & \vdots & \cdots & \vdots \\ D_{\times 1} g_d & D_{\times 2} g_d & \cdots & D_{\times d} g_d \end{pmatrix} \in \mathbb{R}^{N_1 \times \cdots \times N_d \times d \times d}.$$

The (isotropic) norms in (2) are as follows:

$$\|\boldsymbol{g}\|_2 = \sqrt{\textstyle\sum_{i_1 \ldots i_d} g_{1, i_1 \ldots i_d}^2 + g_{2, i_1 \ldots i_d}^2 + \cdots + g_{d, i_1 \ldots i_d}^2},$$
$$\|h\|_1 = \textstyle\sum_{i_1 \ldots i_d} \sqrt{\sum_{l,m=1}^{d} h_{lm, i_1 \ldots i_d}^2}.$$

The linear subspace $V = \{\boldsymbol{g} = \nabla u \colon u \in \mathbb{R}^{N_1 \times \cdots \times N_d}\} \subset \mathbb{R}^{N_1 \times \cdots \times N_d \times d}$ is the range of the orthogonal projector

$$\Pi = \nabla(\nabla^*\nabla)^\dagger \nabla^*, \tag{3}$$

where ∇^* is the adjoint to the ∇ operator, and $\dagger$ denotes the Moore-Penrose pseudoinverse. By the way, $\boldsymbol{g}^o = \Pi \boldsymbol{g}^o \in V$.

The continuous functional

$$\|\nabla \boldsymbol{g}\|_1 + \frac{1}{2\lambda}\|\boldsymbol{g} - \boldsymbol{g}^o\|_2^2, \quad \boldsymbol{g} \in V,$$

is strictly convex, its minimum is unique and attained in the closed ball $\{\boldsymbol{g} \colon \|\boldsymbol{g} - \boldsymbol{g}^o\|_2 \le \|\boldsymbol{g}^o\|_2\}$.

By the aid of a dual variable

$$p = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1d} \\ p_{21} & p_{22} & \cdots & p_{2d} \\ \vdots & \vdots & \cdots & \vdots \\ p_{d1} & p_{d2} & \cdots & p_{dd} \end{pmatrix} \in \mathbb{R}^{N_1 \times \cdots \times N_d \times d \times d}$$

such that $\|\nabla \boldsymbol{g}\|_1 = \max\limits_{\|p\|_\infty \le 1} \langle \nabla \boldsymbol{g}, p\rangle = \max\limits_{\|p\|_\infty \le 1} \langle \boldsymbol{g}, \nabla^* p\rangle$ the model (2) reduces to the equivalent min-max formulation

$$\min_{\boldsymbol{g} \in V} \max_{\|p\|_\infty \le 1} F(\boldsymbol{g}, p), \quad \text{where} \quad F(\boldsymbol{g}, p) = \langle \boldsymbol{g}, \nabla^* p\rangle + \frac{1}{2\lambda}\langle \boldsymbol{g} - \boldsymbol{g}^o, \boldsymbol{g} - \boldsymbol{g}^o\rangle. \tag{4}$$

Here the max norm is $\|p\|_\infty = \max_{i_1 \dots i_d} \sqrt{\sum_{l,m=1}^d |p_{lm,i_1 \dots i_d}|^2}$.

The generalized minimax theorem, cf. [19] , justifies the equality

$$\min_{\boldsymbol{g} \in V} \max_{|p| \le 1} F(\boldsymbol{g}, p) = \max_{|p| \le 1} \min_{\boldsymbol{g} \in V} F(\boldsymbol{g}, p),$$

so that the model (2) reduces to the max-min formulation

$$\max_{\|p\|_\infty \le 1} \min_{\boldsymbol{g} \in V} \langle \boldsymbol{g}, \nabla^* p\rangle + \frac{1}{2\lambda}\langle \boldsymbol{g} - \boldsymbol{g}^o, \boldsymbol{g} - \boldsymbol{g}^o\rangle. \tag{5}$$

In spite of a possible nonuniqueness of p, the solution $\boldsymbol{g}$ obtained by means of (5) is unique. Indeed, if $(\boldsymbol{g}_*, p_*)$ and $(\boldsymbol{g}, p)$ are solutions of (4) and (5) respectively, then $F(\boldsymbol{g}_*, p) = F(\boldsymbol{g}, p)$ by Lemma 36.2 in [17]. The strict convexity of the mapping $\boldsymbol{g} \mapsto F(\boldsymbol{g}, p)$ implies $\boldsymbol{g} = \boldsymbol{g}_*$.

Since the constraint $\boldsymbol{g} \in V$ is equivalent to the condition $\Pi \boldsymbol{g} = \boldsymbol{g}$, the term $\langle \boldsymbol{g}, \nabla^* p\rangle$ in (5) can be replaced by $\langle \boldsymbol{g}, \nabla^* p\rangle = \langle \boldsymbol{g}, \Pi\nabla^* p\rangle$. Now we observe that the unique solution

$$\boldsymbol{g} = \boldsymbol{g}^o - \lambda \Pi \nabla^* p \tag{6}$$

of the unconstrained problem $\min_{\boldsymbol{g} \in \mathbb{R}^{N_1 \times \cdots \times N_d \times d}} \langle \boldsymbol{g}, \Pi\nabla^* p\rangle + \frac{1}{2\lambda}\langle \boldsymbol{g} - \boldsymbol{g}^o, \boldsymbol{g} - \boldsymbol{g}^o\rangle$ also solves the constrained problem $\min_{\boldsymbol{g} \in V} \langle \boldsymbol{g}, \nabla^* p\rangle + \frac{1}{2\lambda}\langle \boldsymbol{g} - \boldsymbol{g}^o, \boldsymbol{g} - \boldsymbol{g}^o\rangle$.

Under the condition (6)

$$F(\boldsymbol{g}, p) = \frac{1}{2\lambda}\|\boldsymbol{g}^o\|_2^2 - \frac{1}{2\lambda}\|\boldsymbol{g}^o - \lambda\Pi\nabla^* p\|_2^2,$$

and we arrive at the minimum distance problem

$$\min_{\|p\|_\infty \le 1} \|\boldsymbol{g}^o - \lambda\Pi\nabla^* p\|_2. \tag{7}$$

2.1. Chambolle's iteration. We solve the minimization problem (7) using Chambolle's iteration, cf. [2, 8]. Accordingly, the Karush-Kuhn-Tucker conditions for (7) are given by the equation

$$\nabla\left(\Pi\nabla^* p - \boldsymbol{g}^o/\lambda\right) + \left|\nabla\left(\Pi\nabla^* p - \boldsymbol{g}^o/\lambda\right)\right| \cdot p = 0, \tag{8}$$

where the dot signifies the entrywise product. Solution of (7) is approximated by the semi-implicit iteration

$$\begin{aligned} p^0 &= 0, \\ p^{k+1} &= \frac{p^k - \tau\nabla\left(\Pi\nabla^* p^k - \dfrac{\boldsymbol{g}^o}{\lambda}\right)}{\max\left\{1, \left|p^k - \tau\nabla\left(\Pi\nabla^* p^k - \dfrac{\boldsymbol{g}^o}{\lambda}\right)\right|\right\}}, \end{aligned} \tag{9}$$

where the maximum and division are both entrywise operations.

LEMMA 1. *The semi-implicit iteration (9) is 1-Lipschitz [10] for*

$$\tau \le \frac{2}{\|\nabla\Pi\nabla^*\|_2}. \tag{10}$$

Proof. Each step of (9) uses two mappings: $p \mapsto p - \tau\nabla\left(\Pi\nabla^* p - \boldsymbol{g}^o/\lambda\right)$ and $q \mapsto q/\max(1, |q|)$. The first mapping is linear and 1-Lipschitz if and only if $\|I - \tau\nabla\Pi\nabla^*\|_2 \le 1$, where I is the identity transformation. The second mapping is always 1-Lipschitz.

The SVD of the matrix D allows us to show that 1-Lipschitz holds when

$$\tau \le \frac{1}{2d}. \tag{11}$$

2.2. Implementation of $\Pi = \nabla(\nabla^*\nabla)^\dagger\nabla^*$. The intricate part in our algorithm is the implementation of the Π operation. We describe here how this can be done efficiently. The singular value decomposition of the differentiation matrix D is

$$D = \begin{bmatrix} 0 & S \\ 1 & 0 \end{bmatrix} \Sigma C,$$

where the diagonal Σ has the entries $\Sigma_{ii} = 2\sin\dfrac{\pi(i-1)}{2N}$, $i = 1, \ldots, N$. The orthogonal $N \times N$ matrix C of the discrete cosine transform has the entries

$$\begin{aligned} C_{1j} &= \sqrt{\frac{1}{N}}, \\ C_{ij} &= \sqrt{\frac{2}{N}}\cos\frac{\pi(i-1)(2j-1)}{2N}, \quad i = 2, \ldots, N, \quad j = 1, \ldots, N. \end{aligned}$$

The orthogonal $(N-1)\times(N-1)$ matrix S of the discrete sine transform has the entries

$$S_{ij}=\sqrt{\frac{2}{N}}\sin\frac{\pi ij}{N},\quad i,j=1,\dots,N-1.$$

We can use the singular value decomposition of D to compute $\|\nabla\|_2$. Indeed, $\|\nabla\|_2=\|\nabla\|_2=2\|\left[\sin\frac{\pi(N_1-1)}{2N_1},\dots,\sin\frac{\pi(N_d-1)}{2N_d}\right]\|_2\approx 2\sqrt{d}$.

We implement $(\nabla^*\nabla)^\dagger$ by the fast Fourier transform. Since $\nabla^*\nabla u=(D^TD)_{\times 1}u+(D^TD)_{\times 2}u+\ldots+(D^TD)_{\times d}u$, the equation $\nabla^*\nabla u=f$ is equivalent to

$$\Sigma^2_{\times 1}\widehat{u}+\Sigma^2_{\times 2}\widehat{u}+\ldots+\Sigma^2_{\times d}\widehat{u}=\widehat{f},$$

where $\widehat{u}=C_{\times 1}C_{\times 2}\dots C_{\times d}u$ and $\widehat{f}=C_{\times 1}C_{\times 2}\dots C_{\times d}f$. The components of $\widehat{u}$ and $\widehat{f}$ are related by the equalities

$$\widehat{u}_{i_1\cdots i_d}(\Sigma_{i_1i_1}+\Sigma_{i_2i_2}+\ldots+\Sigma_{i_di_d})=\widehat{f}_{i_1\cdots i_d},\quad i_k=1,\dots,N_k.$$

Note that $\Sigma_{i_1i_1}+\Sigma_{i_2i_2}+\ldots+\Sigma_{i_di_d}$ equals zero if $i_1=\ldots=i_d=1$ and is positive otherwise. Hence

$$\widehat{u}_{1\cdots 1}=0,\quad \widehat{u}_{i_1\cdots i_d}=\widehat{f}_{i_1\cdots i_d}/(\Sigma_{i_1i_1}+\Sigma_{i_2i_2}+\ldots+\Sigma_{i_di_d})\quad \text{if } i_1+\ldots+i_d>d.$$

Recall that $u=C^T_{\times 1}C^T_{\times 2}\cdots C^T_{\times d}\widehat{u}$. Moreover, multiplication with the matrices C and C^T should be carried out by the fast Fourier transform FFT.

3. The scalar field reconstruction step of the TV-Stokes. We extend the second step or the image reconstruction step of the original TV-Stokes [16] for 2D to multidimensions. Accordingly, for the given d-dimensional image $u^o\in\mathbb{R}^{N_1\times\cdots\times N_d}$, which is corrupted by a Gaussian noise, a vector matching model [16] is applied which satisfies the minimization problem,

$$\min_{u\in\mathbb{R}^{N_1\times\cdots\times N_d}}\left[\|\nabla u\|_1+\frac{1}{2\lambda}\|u-u^o\|_2^2-\nabla u\cdot\frac{\boldsymbol{g}}{|\boldsymbol{g}|}\right],\tag{12}$$

where $\lambda>0$ is a suitable scalar parameter, $\nabla u=(D_{\times 1}u,D_{\times 2}u,\dots,D_{\times d}u)$ and

$$\|\nabla u\|_1=:\|s\|_1=\sum_{i_1\dots i_d}\sqrt{\sum_{l=1}^{d}s^2_{l,i_1\dots i_d}}.$$

Note that the $\boldsymbol{g}$ is obtained from the first step and thus, in the reconstruction, it is a known term. By completing the quadratic term, the above problem (12) can be reformed as

$$\min_{u\in\mathbb{R}^{N_1\times\cdots\times N_d}}\left[\|\nabla u\|_1+\frac{1}{2\lambda}\|u+\lambda\nabla\cdot\frac{\boldsymbol{g}}{|\boldsymbol{g}|}-u^o\|_2^2\right],\tag{13}$$

which is strictly convex, its minimum is unique and attained in the closed ball $\left\{u\colon\left\|u-(u^o-\lambda\nabla\cdot\frac{\boldsymbol{g}}{|\boldsymbol{g}|})\right\|_2\le\left\|u^o-\lambda\nabla\cdot\frac{\boldsymbol{g}}{|\boldsymbol{g}|}\right\|_2\right\}$.

With the aid of a dual variable $p = (p_1, p_2, \cdots, p_d)$ such that $\|\nabla u\|_1 = \max_{\|p\|_\infty \le 1} \langle \nabla u, p\rangle = \max_{\|p\|_\infty \le 1} \langle u, \nabla^* p\rangle$ the model (12) reduces to the equivalent min-max formulation

$$\begin{aligned} &\min_{u\in\mathbb{R}^{N_1\times\cdots\times N_d}} \max_{\|p\|_\infty\le 1} F(u,p), \\ &\text{where} \quad F(u,p) = \langle u, \nabla^* p\rangle + \frac{1}{2\lambda}\|u-u^o\|_2^2 - \nabla u\cdot\frac{\boldsymbol{g}}{|\boldsymbol{g}|}. \end{aligned} \tag{14}$$

Here the max norm is $\|p\|_\infty = \max_{i_1\ldots i_d} \sqrt{\sum_{l=1}^d |p_{l,i_1\ldots i_d}|^2}$.

The generalized minimax theorem, cf. [19], justifies the equality

$$\min_{u\in\mathbb{R}^{N_1\times\cdots\times N_d}} \max_{|p|_\infty\le 1} F(u,p) = \max_{|p|_\infty\le 1} \min_{u\in\mathbb{R}^{N_1\times\cdots\times N_d}} F(u,p),$$

so that the model (12) reduces to the max-min formulation

$$\max_{\|p\|_\infty\le 1} \min_{u\in\mathbb{R}^{N_1\times\cdots\times N_d}} \langle u, \nabla^* p\rangle + \frac{1}{2\lambda}\|u-u^o\|_2^2 - \nabla u\cdot\frac{\boldsymbol{g}}{|\boldsymbol{g}|}. \tag{15}$$

In spite of a possible nonuniqueness of p, the solution u obtained by means of (15) is unique. Indeed, if (u_*, p_*) and (u, p) are solutions of (14) and (15) respectively, then $F(u_*, p) = F(u, p)$ by (Lemma 36.2 of [17]). The strict convexity of the mapping $u \mapsto F(u, p)$ implies $u = u_*$. The unique solution is obtained as follows

$$u = u^o - \lambda(\nabla^* p + \nabla\cdot\frac{\boldsymbol{g}}{|\boldsymbol{g}|}) \tag{16}$$

Under the condition (16)

$$F(u,p) = \frac{1}{2\lambda}\Big\|u^o - \lambda\nabla\cdot\frac{\boldsymbol{g}}{|\boldsymbol{g}|}\Big\|_2^2 - \frac{1}{2\lambda}\Big\|u^o - \lambda\nabla\cdot\frac{\boldsymbol{g}}{|\boldsymbol{g}|} - \lambda\nabla^* p\Big\|_2^2,$$

and we arrive at the minimum distance problem

$$\min_{\|p\|_\infty\le 1} \Big\|u^o - \lambda\nabla\cdot\frac{\boldsymbol{g}}{|\boldsymbol{g}|} - \lambda\nabla^* p\Big\|_2. \tag{17}$$

As in [2, 8] the Karush-Kuhn-Tucker conditions for (17) are given by the equation

$$\nabla\left(-\frac{u^o}{\lambda} + \nabla\cdot\frac{\boldsymbol{g}}{|\boldsymbol{g}|} + \nabla^* p\right) + \left|\nabla\left(-\frac{u^o}{\lambda} + \nabla\cdot\frac{\boldsymbol{g}}{|\boldsymbol{g}|} + \nabla^* p\right)\right|\cdot p = 0, \tag{18}$$

where the dot signifies the entrywise product. Solution of (17) is approximated by the semi-implicit iteration

$$\begin{aligned} p^0 &= 0, \\ p^{k+1} &= \frac{p^k - \tau\nabla\left(\nabla\cdot\frac{\boldsymbol{g}}{|\boldsymbol{g}|} + \nabla^* p - \frac{u^o}{\lambda}\right)}{\max\left\{1, \left|p^k - \tau\nabla\left(\nabla\cdot\frac{\boldsymbol{g}}{|\boldsymbol{g}|} + \nabla^* p - \frac{u^o}{\lambda}\right)\right|\right\}}, \end{aligned} \tag{19}$$

where the maximum and division are entrywise operations.

LEMMA 2. *The semi-implicit iteration (19) is 1-Lipschitz for*

$$\tau \leq \frac{2}{\|\nabla\nabla^*\|_2}. \tag{20}$$

Proof. Each step of (19) uses two mappings:

$$p \mapsto p - \tau\nabla\left(\nabla \cdot \frac{\boldsymbol{g}}{|\boldsymbol{g}|} + \nabla^* p - \frac{u^o}{\lambda}\right)$$

and $q \mapsto q/\max(1, |q|)$. The first mapping is linear and 1-Lipschitz if and only if $\|I - \tau\nabla\nabla^*\|_2 \leq 1$, where I is the identity transformation. The second mapping is always 1-Lipschitz.

The SVD of the matrix D allows us to show that 1-Lipschitz holds when

$$\tau \leq \frac{1}{2d}. \tag{21}$$

4. Numerical experiments. We first test the proposed model on some 3d static data, which are computer tomography data from two lung tissues of human. In FIG. 1, FIG. 3 and FIG. 4, 3 groups of orthogonal slices are presented, with respect to xz-plane, xy-plane, and yz-plane. In these figures, the last rows show relative positions of those slices in their embedded volume data. In the first column of each figures, the 3-slice view is taken on the raw data, which contains some unknown noise. The noise type is supposed to be Gaussian. In the middle column and the last column of each figures, comparison is presented between the denoised data of using ROF model and the proposed multidimensional TV-Stokes model. The one (last column) using the proposed multidimensional TV-Stokes model results higher fidelity to the raw data than the one (middle column) using ROF model and preserves more fine structures as well as better continuity of tube-like structures. FIG. 2 and FIG. 5 show the isosurfaces of the volume data. In these figures, the results from proposed model (last column) well trade-off the smoothness and details preservation. It observes that the fine tube-like structures disappeared while the ROF model smooths the noisy data.

Another experiment is performed on 2d dynamic data, where new 3d data are obtained by stacking all 2d data along temporal dimension. FIG. 6 and FIG. 7 show the results of using proposed multidimensional TV-Stokes on movie / monitoring video denoising. In FIG. 6, smooth cheek and sharp edges are well balanced, showing in row two and row four. In escalator case, cf. FIG. 7, shape edges of structures and digitals are preserved, showing in row two and row four.

Acknowledgments. We thank Dr. Andreas Langer for interesting discussion and for providing us with the 3d data.

REFERENCES

[1] A. BUADES, B. COLL, AND J. MOREL, *A Non-Local Algorithm for Image Denoising*, in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005, pp. 60–65.

[2] A. CHAMBOLLE, *An Algorithm for Total Variation Minimization and Applications*, Journal of Mathematical Imaging and Vision, 20 (2004), pp. 89–97.

[3] T. F. CHAN AND J. SHEN, *Image processing and analysis: variational, PDE, wavelet, and stochastic methods*, SIAM, 2005.

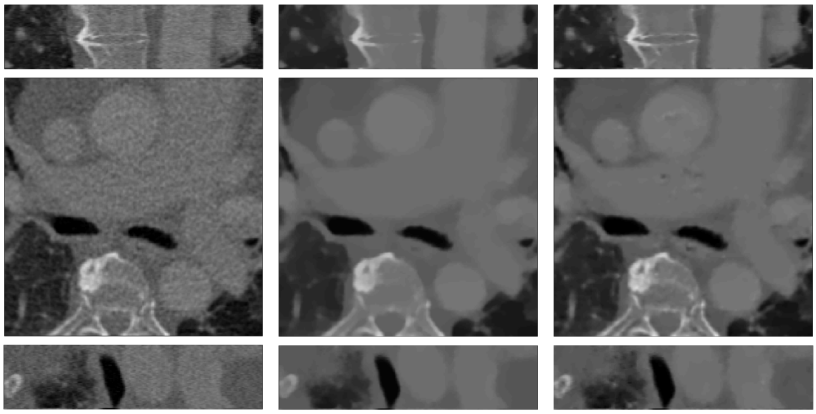

FIG. 1. *Visualization of computer tomography data of human lung tissue A via 3 groups of orthogonal slices. Row one, row two and row three present respectively the xz-slices, the xy-slices and the yz-slices. Column (a) visualizes the raw data with some noise (Gaussian noise). Column (b) visualizes the data denoised with ROF model. Column (c) visualizes the data denoised with proposed multidimensional TV-Stokes model.*

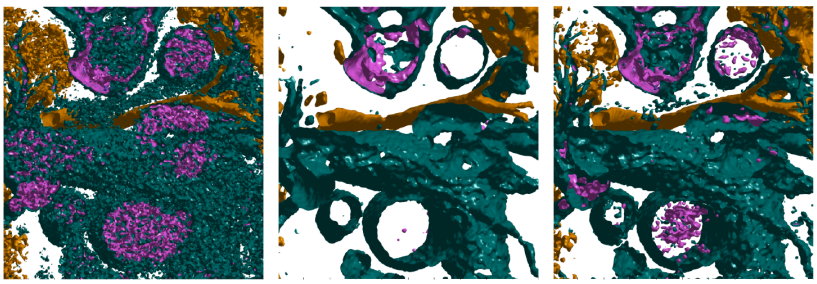

FIG. 2. *Isosurface show of computer tomography data of human lung tissue A. (a) raw data with some noise (Gaussian noise). (b) data denoised with ROF model. (c) data denoised with proposed multidimensional TV-Stokes model.*

[4] M. R. CHAREST, M. ELAD, AND P. MILANFAR, *A general iterative regularization framework for image denoising*, 2006 IEEE Conference on Information Sciences and Systems, CISS 2006 - Proceedings, 1 (2006), pp. 452–457.

[5] Y. CHEN AND T. POCK, *Trainable Nonlinear Reaction Diffusion: A Flexible Framework for Fast and Effective Image Restoration*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 39 (2017), pp. 1256–1272, https://arxiv.org/abs/1508.02848.

[6] K. DABOV, A. FOI, V. KATKOVNIK, AND K. EGIAZARIAN, *Image Denoising by Sparse 3D Transformation-Domain Collaborative Filtering*, IEEE Transactions on Image Processing, 16 (2007), pp. 1–16, http://ieeexplore.ieee.org/stamp/stamp.jsp?tp={&}arnumber=4271520, https://arxiv.org/abs/arXiv:1011.1669v3.

[7] C. A. ELO, *Image Denoising Algorithms Based on the Dual Formulation of Total Variation*, master thesis, University of Bergen, 2009, http://hdl.handle.net/1956/3367.

[8] C. A. ELO, A. MALYSHEV, AND T. RAHMAN, *A Dual Formulation of the TV-Stokes Algorithm for Image Denoising*, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 5567 LNCS (2009), pp. 307–318.

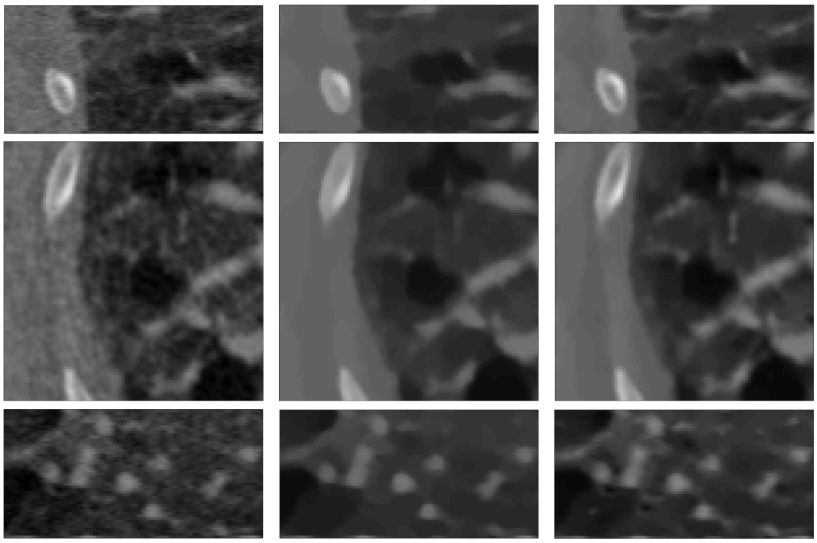

FIG. 3. *Visualization of computer tomography data of human lung tissue B via 3 groups of orthogonal slices. Row one, row two and row three present respectively the xz-slices, the xy-slices and the y-slices. Column (a) visualizes the raw data with some noise (Gaussian noise). Column (b) visualizes the data denoised with ROF model. Column (c) visualizes the data denoised with proposed multidimensional TV-Stokes model.*

[9] J. HAHN, C. WU, AND X.-C. TAI, *Augmented Lagrangian Method for Generalized TV-Stokes Model*, Journal of Scientific Computing, 50 (2012), pp. 235–264, http://link.springer.com/10.1007/s10915-011-9482-6.

[10] J. HEINONEN, *Lectures on analysis on metric spaces*, Springer Science & Business Media, 2012.

[11] H. JI, C. LIU, Z. SHEN, AND Y. XU, *Robust Video Denoising Using Low Rank Matrix Completion*, in 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE, 2010, pp. 1791–1798, http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5539849.

[12] W. G. LITVINOV, T. RAHMAN, AND X.-C. TAI, *A Modified TV-Stokes Model for Image Processing*, SIAM Journal on Scientific Computing, 33 (2011), pp. 1574–1597.

[13] M. LYSAKER AND X. C. TAI, *Iterative Image Restoration Combining Total Variation Minimization and a Second-Order Functional*, International Journal of Computer Vision, 66 (2006), pp. 5–18.

[14] L. MARCINKOWSKI AND T. RAHMAN, *An Iterative Regularization Algorithm for the TV-Stokes in Image Processing*, in Parallel Processing and Applied Mathematics, Lecture Notes in Computer Science, vol. 9574, Springer, Cham, 2016, pp. 381–390.

[15] S. OSHER, M. BURGER, D. GOLDFARB, J. XU, AND W. YIN, *An Iterative Regularization Method for Total Variation-Based Image Restoration*, Multiscale Modeling & Simulation, 4 (2005), pp. 460–489.

[16] T. RAHMAN, X. TAI, AND S. OSHER, *A TV-Stokes Denoising Algorithm*, in Scale Space and Variational Methods in Computer Vision, Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2007, pp. 473–483.

[17] R. ROCKAFELLAR, *Convex analysis*, Princeton University Press, Princeton, NJ, 1997.

[18] L. I. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear Total Variation Noise Removal Algorithm*, Physica D: Nonlinear Phenomena, 60 (1992), pp. 259–268, http://www.csee.wvu.edu/~xinl/courses/ee565/total_variation.pdf, https://arxiv.org/abs/arXiv:1011.1669v3.

[19] M. SION, *On General Minimax Theorems*, Pacific Journal of Mathematics, 18 (1957), pp. 171–176.

[20] X.-C. TAI, S. OSHER, AND R. HOLM, *Image Inpainting Using a TV-Stokes Equation*, CAM

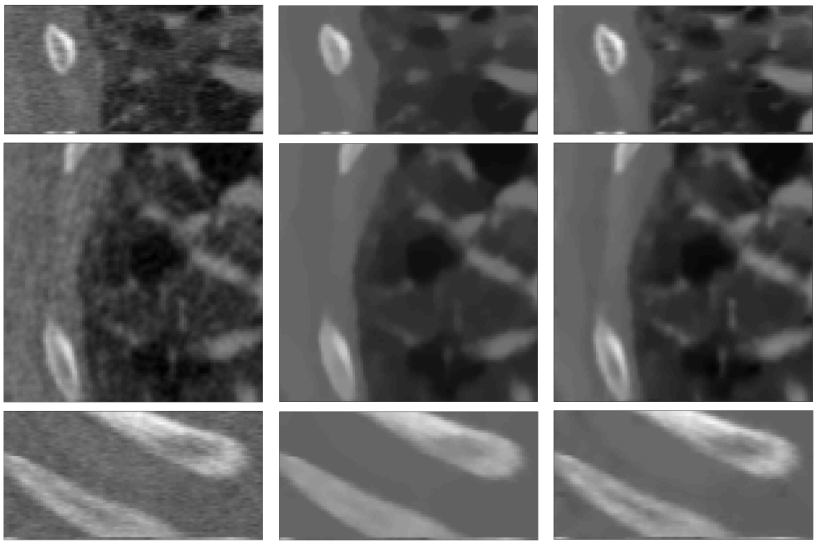

FIG. 4. *Visualization of computer tomography data of human lung tissue B via 3 groups of orthogonal slices. In contrast to* FIG. 3*, different observation slices are presented. Row one, row two and row three present respectively the xz-slices, the xy-slices and the yz-slices. Column (a) visualizes the raw data with some noise (Gaussian noise). Column (b) visualizes the data denoised with ROF model. Column (c) visualizes the data denoised with proposed multidimensional TV-Stokes model.*

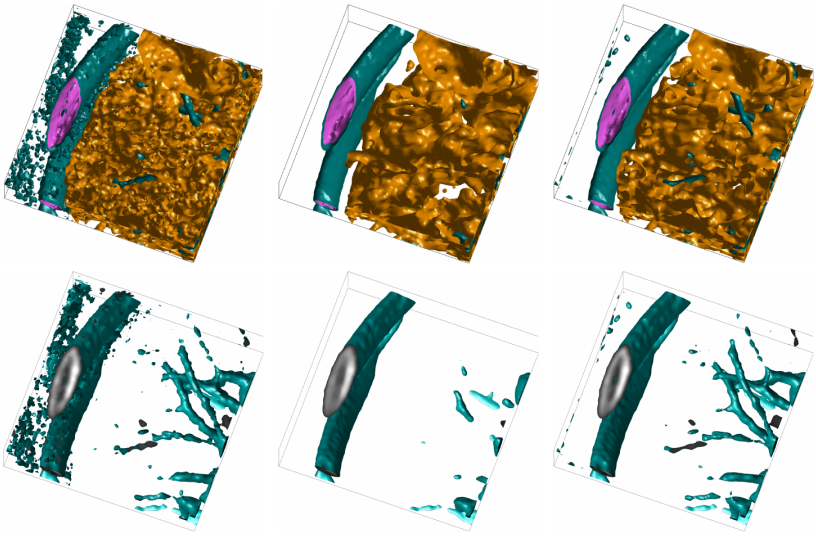

FIG. 5. *Isosurface show of computer tomography data of human lung tissue B. (a) raw data with some noise (Gaussian noise). (b) data denoised with ROF model. (c) data denoised with proposed multidimensional TV-Stokes model. In row two, only tube-like structures are presented.*

FIG. 6. *Contrast of initial noisy movie and its denoised copy, by using proposed multidimensional TV-Stokes model. Row one and row three are sequences from initial movie. Row two and row four are sequences from denoised results.*

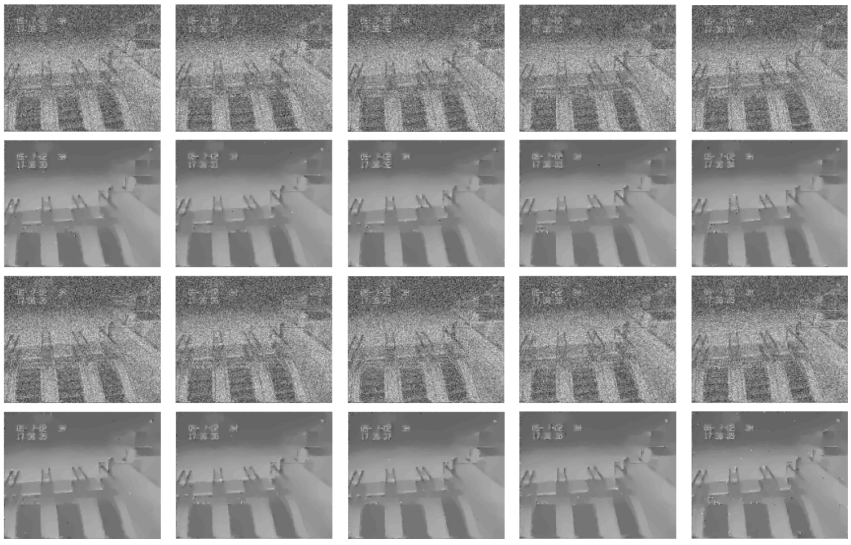

FIG. 7. *Contrast of initial noisy monitoring video and its denoised copy, by using proposed multidimensional TV-Stokes model. Row one and row three are sequences from initial video. Row two and row four are sequences from denoised results.*

report, 06 (2006), pp. 1–20, ftp://ftp.math.ucla.edu/pub/camreport/cam06-01.pdf.

[21] B. WU, T. RAHMAN, AND X. C. TAI, *Sparse-Data Based 3D Surface Reconstruction for Cartoon and Map*, Mathematics and Visualization, 0 (2018), pp. 47–64.

Paper D

Iterative Regularization Algorithms for Image Denoising with the TV-Stokes Model

Bin Wu, Leszek Marcinkowski, Xue-Cheng Tai, and Talal Rahman *Submitted. arXiv:2009.11976*

ITERATIVE REGULARIZATION ALGORITHMS FOR IMAGE DENOISING WITH THE TV-STOKES MODEL

BIN WU*, LESZEK MARCINKOWSKI†, XUE-CHENG TAI‡, AND TALAL RAHMAN*

Abstract. We propose a set of iterative regularization algorithms for the TV-Stokes model to restore images from noisy images with Gaussian noise. These are some extensions of the iterative regularization algorithm proposed for the classical Rudin-Osher-Fatemi (ROF) model for image reconstruction, a single step model involving a scalar field smoothing, to the TV-Stokes model for image reconstruction, a two steps model involving a vector field smoothing in the first and a scalar field smoothing in the second. The iterative regularization algorithms proposed here are Richardson's iteration like. We have experimental results that show improvement over the original method in the quality of the restored image. Convergence analysis and numerical experiments are presented.

Key words. Image processing, Total variation minimization

AMS subject classifications. 68Q25, 68R10, 68U05

1. Introduction. Recovering an image from a noisy and blurry image is an inverse problem which is possible to be solved via variational methods, using total variation regularization, e.g., cf. [21, 7, 8, 1, 20, 30, 24, 22, 11, 12, 29, 26, 28, 14, 3, 15, 6, 9, 27, 2, 13, 16]. In this paper, we only focus on the denoising problem in image processing. Considering a noisy image $f : \Omega \mapsto \mathbb{R}$, where Ω is a bounded open subset of $\mathbb{R}^2$, the problem is to find a decomposition such that $f = u + v$, where u is the signal, and v is the noise. Let us consider this problem as an optimization problem. The simplest model can be the least square fitting, in other words, to find the minimizer in the squared L^2 space:

$$u = \arg\min_u \|u - f\|^2_{(\Omega)}.$$

The notation $\|\cdot\|$, in this paper, refers to a L^2 norm if there is no specific subscript. This model, however, only works when we know the structure of u otherwise there is only a trivial solution $u = f$. It is obvious that, without sufficient priori, to find a decomposition is an ill-posed inverse problem, cf. e.g., [8, 1]. A regularizer is thus necessary. The Tikhonov regularizer is the first one used in this problem in history. In general, we define regularizer as

$$J_p(u) := \int_\Omega |\nabla u|^p.$$

The Tikhonov regularizer is the case where $p = 2$. For the models with this regularizer, it is difficult to preserve edges while smoothing noise. Rudin, Osher, and Fatemi proposed a model (ROF) with a regularizer where $p = 1$. To be simplified in presentation, in this paper, we denote $J(\cdot)$ equipped with $p = 1$ as default. The regularizer is thus the BV seminorm where $BV(\Omega)$ means the space of functions with the bounded variation on Ω. The ROF model successfully enhances the capability in

*Department of Computer Science, Electrical Engineering and Mathematical Sciences, Western Norway University of Applied Sciences, Inndalsveien 28, 5063 Bergen, Norway (bin.wu@hvl.no, talal.rahman@hvl.no).

†Faculty of Mathematics, University of Warsaw, Banacha 2, 02-097 Warszawa, Poland (leszek.marcinkowski@mimuw.edu.pl).

‡Department of Mathematics, University of Bergen, Allégaten 41, 5007 Bergen, Norway (xuecheng.tai@uib.no).

edge preserving. However, it suffers a staircase effect which makes the restored image patternized. There are many models to overcome this problem, for instance, the high order regularization, LOT model, and TV-Stokes. The TV-Stokes model is defined as follows:

$$\boldsymbol{\tau} = \arg \min_{\substack{\boldsymbol{\tau} \in BV(\Omega) \\ \nabla \cdot \boldsymbol{\tau} = 0}} \Big\{ J(\boldsymbol{\tau}) + H(\boldsymbol{\tau}, \boldsymbol{\tau}^0) \Big\}, \tag{1}$$

and

$$u = \arg \min_{u \in BV(\Omega)} \Big\{ J(u) - \langle \nabla u, \frac{\boldsymbol{\tau}^\perp}{|\boldsymbol{\tau}^\perp|} \rangle + H(u, f) \Big\}, \tag{2}$$

where $\boldsymbol{\tau}, \boldsymbol{\tau}^0 \in \mathbb{R}^2$ are vectors, $\nabla \boldsymbol{\tau}$ inside $J(\boldsymbol{\tau})$ is a 2×2 matrix, i.e., the gradient of the vector $\boldsymbol{\tau}$; $H(u, f) := \frac{\eta}{2} \|u - f\|^2$ stands the quadratic fidelity with a scale parameter η; $\langle \cdot, \cdot \rangle$ denotes the inner product. In this model, a smoothed tangent field $\boldsymbol{\tau}$ is firstly obtained by solving the minimization problem (1) under a divergence-free constraint, and subsequently the restored u is obtained by a kind of vector matching under a limited deviation which is formulated as (2).

The ROF model is also considered defective in some cases for signal and noise decomposition, cf. e.g., [18, 17]. There are many ways to handle this problem, for instance, Meyer's model, cf. [17], the Vese and Osher's approximated Meyer's model, cf. [25], the Osher, Solé, and Vese's model, cf. [19]. There is another stream of methods which handle the problem through the iterative way, e.g., [18] applies an iterative algorithm on the ROF model. To heuristically introduce this algorithm, we start with a typical ROF model as follows,

$$u^1 = \arg \min_{u \in BV(\Omega)} \Big\{ J(u) + H(u, f) \Big\}. \tag{3}$$

We then calculate unit normal vector $\frac{\mathbf{n}^1}{|\mathbf{n}^1|} = \frac{\nabla u^1}{|\nabla u^1|}$ and perform a vector matching step

$$u^2 = \arg \min_{u \in BV(\Omega)} \Big\{ J(u) - \langle \nabla u, \frac{\mathbf{n}^1}{|\mathbf{n}^1|} \rangle + H(u, f) \Big\}, \tag{4}$$

The optimal condition for (3), namely the Euler-Lagrange equation, is

$$-\nabla \cdot \frac{\nabla u^1}{|\nabla u^1|} + \eta(u^1 - f) = 0. \tag{5}$$

Slightly reforming (4) with adjoint, cf. [18], and substituting the relation (5) into the reformed (4), we get

$$u^2 = \arg \min_{u \in BV(\Omega)} \Big\{ J(u) - \langle u, \eta(f - u^1) \rangle + H(u, f) \Big\}.$$

By completing the square with some added constants, the above minimization problem is equivalent to the following

$$u^2 = \arg \min_{u \in BV(\Omega)} \Big\{ J(u) + H(u, f + f - u^1) \Big\}. \tag{6}$$

Algorithm 1 The iterative regularization proposed by Osher et al.

Initialize $k = 0$, $v^0 = 0$;
repeat
  $u = \arg\min_{u \in BV(\Omega)} \{J(u) + H(u, f + v^k)\}$;
  Update noise: $v^{k+1} = f + v^k - u$;
  $k = k + 1$;
until satisfied;
return u.

It implies that the matching step exactly equivalents to a ROF model. The initial image f is accordingly replaced by f added with a 'noise' $(f - u^1)$ obtained from the previous step. By an induction, Algorithm 1 has been proposed by Osher and his coworkers, cf. [18] for more details,.

Instead of considering the given image with accumulative noise, we consider a direct process on noise part in this paper. The idea can source back to the 'twicing' method proposed by Tukey, which corrects the approximate solution obtained from the first step by repeating the same processing on its residual. We noted that this idea is the modified Richardson iteration which has been generalized to image restoration problems by Michael Charest Jr. and his coworkers based on scalar valued functional. By defining the operation of finding the solution of a minimization problem as $T(\cdot)$, and starting with initial iterate u^0, the analog of the modified Richardson iteration reads as follows,

$$u^{k+1} = u^k + T(f - u^k), \tag{7}$$

which is equivalent to the following in terms of residuals

$$r_{ex}^{k+1} = r_{ex}^k - T(r_{ex}^k), \tag{8}$$

where the exact residual is defined as $r_{ex}^k = f - u^k$. It also can be derived that $u^k = u^0 + \sum_{i=0}^{k-1} T(r_{ex}^i)$ and $r_{ex}^k = r_{ex}^0 - \sum_0^{k-1} T(r_{ex}^i)$. Let us call the above the Richardson-like iteration. In this paper, we present several Richardson-like iterative algorithms based on the TV-Stokes model.

The paper is organized as follows. Our main results including the proposed algorithms are in section 2, experimental results are in section 3, and the conclusions follow in section 4.

2. Proposed algorithms and their convergence analysis. In this section, we present several variants of the iterative regularization algorithm for the TV-Stokes model. Those algorithms are quite simple and of the form of Richardson iteration. We prove their convergence based on the Bregman distance.

2.1. Prelimits. Before we start to present the algorithm, let first consider two equivalent minimization problems.

LEMMA 1. *$\forall \boldsymbol{\tau} \in \mathbb{R}^2$, define operator Π, such that $\Pi(\boldsymbol{\tau}) = (I - \nabla\triangle^\dagger\nabla\cdot)\boldsymbol{\tau}$. The constrained problem (1) is equivalent to the following unconstrained problem*

$$\boldsymbol{\tau} = arg \min_{\boldsymbol{\tau} \in BV(\Omega)} \Big\{ J(\Pi\boldsymbol{\tau}) + H(\boldsymbol{\tau}, \boldsymbol{\tau}^0) \Big\}. \tag{9}$$

Proof. Let $\mathbf{p}$ be the dual variable such that $\mathbf{p} \in C_c^1(\Omega, \mathbb{R}^4)$ and $|\mathbf{p}| \leq 1$, the minimization problem in (1) is thus

$$\min_{\boldsymbol{\tau}} \max_{\substack{\lambda, \mathbf{p} \\ |p| \leq 1}} \left\{ \int_\Omega \langle \boldsymbol{\tau}, \nabla \cdot \mathbf{p} \rangle + \frac{\eta}{2} (\boldsymbol{\tau} - \boldsymbol{\tau}^0)^2 + \langle \lambda, \nabla \cdot \boldsymbol{\tau} \rangle d\mathbf{x} \right\}, \tag{10}$$

where $\lambda \in \mathbb{R}$ is the Lagrange multiplier. By the Minimax theorem, cf. [23], we can firstly consider the minimization problem with respect to $\boldsymbol{\tau}$ as well as the the maximization with respect to λ freezing $\mathbf{p}$. The corresponding Euler-Lagrange equations are

$$\nabla \cdot \mathbf{p} + \eta(\boldsymbol{\tau} - \boldsymbol{\tau}^0) - \nabla \lambda = 0, \tag{11}$$

and

$$\nabla \cdot \boldsymbol{\tau} = 0.$$

Taking the divergence for the both sides of (11), we obtain the following relation with the help of Moore-Penrose pseudoinverse

$$\lambda = \Delta^\dagger \nabla \cdot \nabla \cdot \mathbf{p}.$$

The considered problem (10) is accordingly

$$\min_{\boldsymbol{\tau}} \max_{\substack{\mathbf{p} \\ |p| \leq 1}} \left\{ \int_\Omega \langle \boldsymbol{\tau}, \nabla \cdot \mathbf{p} \rangle + \frac{\eta}{2} (\boldsymbol{\tau} - \boldsymbol{\tau}^0)^2 + \langle \Delta^\dagger \nabla \cdot \nabla \cdot \mathbf{p}, \nabla \cdot \boldsymbol{\tau} \rangle d\mathbf{x} \right\},$$

which is exactly same as follows with adjoint

$$\min_{\boldsymbol{\tau}} \max_{\substack{\mathbf{p} \\ |p| \leq 1}} \left\{ \int_\Omega \langle \boldsymbol{\tau}, (I - \nabla \Delta^\dagger \nabla \cdot) \nabla \cdot \mathbf{p} \rangle + \frac{\eta}{2} (\boldsymbol{\tau} - \boldsymbol{\tau}^0)^2 d\mathbf{x} \right\}.$$

Rewriting with Π operator as we defined, it is

$$\min_{\boldsymbol{\tau}} \max_{\substack{\mathbf{p} \\ |p| \leq 1}} \left\{ \int_\Omega \langle \boldsymbol{\tau}, \Pi(\nabla \cdot \mathbf{p}) \rangle + \frac{\eta}{2} (\boldsymbol{\tau} - \boldsymbol{\tau}^0)^2 d\mathbf{x} \right\},$$

which is equivalent to the following primal problem

$$\min_{\boldsymbol{\tau}} \left\{ \int_\Omega |\nabla \Pi(\boldsymbol{\tau})| + \frac{\eta}{2} (\boldsymbol{\tau} - \boldsymbol{\tau}^0)^2 d\mathbf{x} \right\}.$$

□

It is worth to mention that $\Pi(\cdot)$ is exactly the orthogonal projection to the divergence free subspace, cf. [11]. The consequent property is that, for all $\boldsymbol{\tau} \in Y$ such that $Y = \{\mathbf{m} : \nabla \cdot \mathbf{m} = 0\}$, we have $\Pi(\boldsymbol{\tau}) = \boldsymbol{\tau}$.

Let us recall the second step of the TV-Stokes model, cf. (2), as follows.

$$u = \arg \min_{u \in BV(\Omega)} \left\{ J(u) - \alpha \langle \nabla u, \frac{\boldsymbol{\tau}^\perp}{|\boldsymbol{\tau}^\perp|} \rangle + H(u, f) \right\}, \tag{12}$$

where α is the parameter for orientation matching term $-\langle \nabla u, \frac{\boldsymbol{\tau}^\perp}{|\boldsymbol{\tau}^\perp|} \rangle$. When $\alpha = 1$, the above minimization problem degenerates to (2). By completing the square, we

can reform the above problem as follows

$$u = \arg\min_{u\in BV(\Omega)} \left\{ J(u) + H(u, f - \frac{\alpha}{\eta}\nabla\cdot\frac{\boldsymbol{\tau}^{\perp}}{|\boldsymbol{\tau}^{\perp}|}) \right\}. \tag{13}$$

Observing (13), we can find out that there is an optimal decomposition $f = u + \frac{\alpha}{\eta}\nabla\cdot\frac{\boldsymbol{\tau}^{\perp}}{|\boldsymbol{\tau}^{\perp}|}$ corresponding to the fidelity parameter $\eta/2$. According to Meyer's theory, cf. [17], $\frac{\alpha}{\eta}\nabla\cdot\frac{\boldsymbol{\tau}^{\perp}}{|\boldsymbol{\tau}^{\perp}|}$ can be read as the high frequency part, which represents the fine structures and the noise part of the corrupted image f. Since it is the indistinguishable part from noise for ROF model, we can roughly say it is also of Gaussian distribution with mean 0 if the noise is white Gaussian, that is $\frac{\alpha}{\eta}\nabla\cdot\frac{\boldsymbol{\tau}^{\perp}}{|\boldsymbol{\tau}^{\perp}|} \sim N(0,\sigma^2)$ for some unknown variance σ^2. Another thing we can find out from (13) is that the model results smooth u with the variance $\sigma^2 = 1/\eta$, cf. e.g., [4, 9], considering Gaussian noise only. Up to here, we can find that the TV-Stokes model tends to find an optimal image u which is close to "clean" image $f - \frac{\alpha}{\eta}\nabla\cdot\frac{\boldsymbol{\tau}^{\perp}}{|\boldsymbol{\tau}^{\perp}|}$, where the noisy part $\frac{\alpha}{\eta}\nabla\cdot\frac{\boldsymbol{\tau}^{\perp}}{|\boldsymbol{\tau}^{\perp}|}$ has an amplitude equal to its variance level while $\alpha = 1$, that is $\sigma^2 = 1/\eta$.

In our proposed algorithms, we only consider Richardson-like iterations, applied to the residual. For Gaussian noise, the residual r satisfies $r \sim N(0,\sigma^2)$ for some unknown variance σ^2. It is natural to assume that a fixed percentage of the residual is the uncertain part, cf. $\alpha \in [0,1]$ in (13). When $\alpha = 0$, the model reduces to the ROF model.

The following lemmas are necessary for the convergence analysis.

LEMMA 2. *Given $a, b, c \in L^2(\Omega;\mathbb{R})$, where $b \sim N(0,\sigma^2)$, $a \sim N(b,\sigma_1^2)$ and $c \sim N(b,\sigma_2^2)$, such that $\sigma_1^2 \le \sigma_2^2$, then $\|a\| \le \|c\|$.*

Proof. Since $a \sim N(b,\sigma_1^2)$, $a - b$ is a Gaussian distribution such that $a - b \sim N(0,\sigma_1^2)$. Since b and $a - b$ are two independent Gaussian distributions, their sum $a = b + (a - b)$ is another Gaussian distribution such that $a \sim N(0,\sigma^2 + \sigma_1^2)$.

Similarly, the sum $c = b + (c - b)$ is also a Gaussian distribution such that $c \sim N(0,\sigma^2 + \sigma_2^2)$.

And now, since $\sigma_1^2 \le \sigma_2^2$, it follows that $\|a\| \le \|c\|$. ∎

LEMMA 3. *Suppose $u \in BV(\Omega;\mathbb{R})$. Consider two minimization problems same as the second step of TV-Stokes, i.e., $\min_u\{J(u) - \alpha\langle\nabla u, \mathbf{v}/|\mathbf{v}|\rangle + \eta(u-f)^2/2\}$, corresponding to two different fidelity parameters, η_1 and η_2, such that $\eta_1 \le \eta_2$. If $u_1 = \arg\min_u\{J(u) - \alpha\langle\nabla u, \mathbf{v}/|\mathbf{v}|\rangle + \eta_1(u-f)^2/2\}$ and $u_2 = \arg\min_u\{J(u) - \alpha\langle\nabla u, \mathbf{v}/|\mathbf{v}|\rangle + \eta_2(u-f)^2/2\}$, then $\|u_1 - f\|^2 \ge \|u_2 - f\|^2$.*

Proof. Rewrite the minimization problem for η_2 as follows.

$$\min_u \left\{ J(u) - \alpha\langle\nabla u, \frac{\mathbf{v}}{|\mathbf{v}|}\rangle + \frac{\eta_1}{2}(u-f)^2 + \frac{\eta_2-\eta_1}{2}(u-f)^2 \right\}. \tag{14}$$

Since u_2 is the minimizer of (14), the functional has following inequality

$$\begin{aligned} &\int_\Omega J(u_2) - \alpha\langle\nabla u_2, \frac{\mathbf{v}}{|\mathbf{v}|}\rangle + \frac{\eta_1}{2}(u_2-f)^2 + \frac{\eta_2-\eta_1}{2}(u_2-f)^2 \\ &\le \int_\Omega J(u_1) - \alpha\langle\nabla u_1, \frac{\mathbf{v}}{|\mathbf{v}|}\rangle + \frac{\eta_1}{2}(u_1-f)^2 + \frac{\eta_2-\eta_1}{2}(u_1-f)^2 \\ &\le \int_\Omega J(u_2) - \alpha\langle\nabla u_2, \frac{\mathbf{v}}{|\mathbf{v}|}\rangle + \frac{\eta_1}{2}(u_2-f)^2 + \frac{\eta_2-\eta_1}{2}(u_1-f)^2. \end{aligned} \tag{15}$$

Note that the first two terms in the functional are actually same as the minimization problem for η_1. u_1 therefore minimizes the energy composed by this two terms.

Comparing the first line and the last line of (15), we obtain $\|u_1 - f\|^2 \geq \|u_2 - f\|^2$ by using relation $\eta_1 \leq \eta_2$. ∎

LEMMA 4. *Consider a minimization problems with ROF model, that is* $r^* = \arg\min_r\{J(r) + \eta(r - r^0)^2/2\}$. *Define* $\eta = \beta/\gamma$ *and* $\beta \in (1, +\infty)$. *There exists a constant* $\gamma > 0$ *for any* $r^0 \neq 0$ *such that* $r^* \neq 0$.

We will use the consequences from Meyer's theory, cf. [17], to prove this lemma.

Proof. According to Meyer's theory, [17, Lemma 4 and Theorem 3 on p. 32], if $\|r^0\|_* > 1/\eta$, the ROF model generates a non-trivial decomposition, $r^0 = r^* + v$, that is $r^* \neq 0$ for any $r^0 \neq 0$.

According to [17, Lemma 3 on p. 31], if $r^0 \in L^2(\mathbb{R}^2)$, then $|\int \tilde{r}(x)r^0(x)dx| \leq \|\tilde{r}\|_{BV}\|r^0\|_*$. By simply replacing $\tilde{r}(x)$ with $r^0(x)$, we obtain $\|r^0\|^2/\|r^0\|_{BV} \leq \|r^0\|_*$. While $r^0 \neq 0$ and not constant over the entire domain Ω, which is nature by considered problem, $\|r^0\|^2/\|r^0\|_{BV} > 0$. Define $\gamma := \|r^0\|^2/\|r^0\|_{BV}$. Since $\beta \in (1, +\infty)$ and $\eta = \beta/\gamma$, we have $\gamma > 1/\eta$. To sum up, the inequality, $\|r^0\|_* > 1/\eta$, holds and, as the consequence from Meyer's theory, $r^* \neq 0$ while $r^0 \neq 0$. ∎

REMARK 1. *For a discrete system, equipped with the finite center difference scheme for example, the divergence* $\nabla \cdot \mathbf{g}$ *can be expressed as follows.*

$$r^0 = \nabla \cdot \mathbf{g} = \partial_1 g_1 + \partial_2 g_2 = \frac{g_1^+ - g_1^-}{2h} + \frac{g_2^+ - g_2^-}{2h},$$

where $^+$ *and* $^-$ *denote the forward and backward positions, respectively.* h *is the uniform discretized unit. We thus obtain the following inequality from the above definition.*

$$\begin{aligned}
|r^0|^2 &= \left(\frac{g_1^+ - g_1^-}{2h} + \frac{g_2^+ - g_2^-}{2h}\right)^2 \\
&\leq 2\left(\frac{g_1^+ + g_2^+}{2h}\right)^2 + 2\left(\frac{g_1^- + g_2^-}{2h}\right)^2 \\
&\leq \frac{1}{h^2}((g_1^+)^2 + (g_2^+)^2 + (g_1^-)^2 + (g_2^-)^2) \\
&\leq \frac{4}{h^2}(\overline{g_1^2} + \overline{g_2^2}) \\
&\sim \frac{4}{h^2}(g_1^2 + g_2^2),
\end{aligned}$$

where $\overline{(\cdot)}$ *denotes the average value over the finite volume. Consider the* L^∞ *norm, we obtain*

$$\|r^0\|_\infty \leq \frac{2}{h}\|\mathbf{g}\|_\infty.$$

Since $\|r^0\|_*$ *is the infimum of* $\|\mathbf{g}\|_\infty$, *we find* $\|r^0\|_\infty h/2 \leq \|r^0\|_*$ *corresponding to* $\gamma = \|r^0\|_\infty h/2$. *In practice,* $h = 1$.

2.2. Iterative regularization for the first step of TV-Stokes. In this subsection, we only consider the Richardson-like iteration on the first step of TV-Stokes model. Following from the unconstrained problem (9), with the help of Π operator, our proposed Richardson-like algorithm is as follows,

Algorithm 2 Iterative regularization applied to the 1^{st} step of TV-Stokes

1: Initialize $k = 0$, $\boldsymbol{\tau} = 0$, $\mathbf{r}_{ex}^0 = \nabla^\perp f$;
2: **repeat**
3: $\quad k = k + 1$;
4: $\quad \mathbf{r}^k = \arg\min_{\mathbf{r}\in BV(\Omega)}\{J(\Pi\mathbf{r}) + H(\mathbf{r}, \mathbf{r}_{ex}^{k-1})\}$;
5: $\quad \mathbf{r}_{ex}^k = \mathbf{r}_{ex}^{k-1} - \mathbf{r}^k$;
6: $\quad \boldsymbol{\tau} = \boldsymbol{\tau} + \mathbf{r}^k$;
7: **until** satisfied;
8: $u = \arg\min_{u\in BV(\Omega)}\{J(u) - \langle\nabla u, \frac{\boldsymbol{\tau}^\perp}{|\boldsymbol{\tau}^\perp|}\rangle + H(u, f)\}$;
9: **return** u.

Let us define a convex functional $Q^{\mathbf{s}^{k-1}}(\mathbf{r})^k$ for each iteration in Algorithm 2 with

$$Q^{\mathbf{s}^{k-1}}(\mathbf{r})^k = H(\mathbf{r}, 0) + J(\Pi\mathbf{r}) - J(\Pi\mathbf{r}^{k-1}) - \langle\mathbf{s}^{k-1}, \mathbf{r} - \mathbf{r}^{k-1}\rangle, \tag{16}$$

where $\mathbf{r}^{k-1}$ denotes the minimizer for $Q^{\mathbf{s}^{k-2}}(\mathbf{r})^{k-1}$, and $\mathbf{r}_{ex}^k := \mathbf{r}_{ex}^{k-1} - \mathbf{r}^k$ is the exact residual, giving $\mathbf{r}_{ex}^0 = \nabla^\perp f$. By defining $\mathbf{s}^{k-1} := \eta\mathbf{r}_{ex}^{k-1}$, considering the problem $\mathbf{r}^k = \arg\min_{\mathbf{r}\in BV(\Omega)} Q^{\mathbf{s}^{k-1}}(\mathbf{r})^k$, we have

$$\begin{aligned}
\mathbf{r}^k &= \arg\min_{\mathbf{r}\in BV(\Omega)} Q^{\mathbf{s}^{k-1}}(\mathbf{r})^k \\
&= \arg\min_{\mathbf{r}\in BV(\Omega)}\{H(\mathbf{r}, 0) + J(\Pi\mathbf{r}) - J(\Pi\mathbf{r}^{k-1}) - \langle\eta\mathbf{r}_{ex}^{k-1}, \mathbf{r} - \mathbf{r}^{k-1}\rangle\} \\
&= \arg\min_{\mathbf{r}\in BV(\Omega)}\{H(\mathbf{r}, 0) + J(\Pi\mathbf{r}) - \langle\eta\mathbf{r}_{ex}^{k-1}, \mathbf{r}\rangle\} \\
&= \arg\min_{\mathbf{r}\in BV(\Omega)}\{H(\mathbf{r}, \mathbf{r}_{ex}^{k-1}) + J(\Pi\mathbf{r})\}, \qquad (17)
\end{aligned}$$

which implies the considered problem $\mathbf{r}^k = \arg\min_{\mathbf{r}\in BV(\Omega)} Q^{\mathbf{s}^{k-1}}(\mathbf{r})^k$ is equivalent to the problem listed on the line 4 in Algorithm 2 since the terms in $Q^{\mathbf{s}^{k-1}}(\mathbf{r})^k$ without $\mathbf{r}$ are constants for the k^{th} iteration. The Algorithm 2 is therefore can be reformed as Algorithm 3

2.2.1. Well-definedness of iterates. Let us start with a simple case without iteration, specifically for a fixed k. The considered minimization problem is (18). For such a given problem, we can find the solution exists and is unique.

LEMMA 5. *Let* $\boldsymbol{\mathcal{R}} = \{\mathbf{r}|\mathbf{r} \in BV(\Omega;\mathbb{R}^2), \Pi\mathbf{r} = \mathbf{r}\}$, $F(\mathbf{r}) = J(\Pi\mathbf{r}) + H(\mathbf{r}, \mathbf{z})$ *and* $\Pi\mathbf{z} = \mathbf{z}$. *Consider the problem to find* $\mathbf{r}^*$ *such that*

$$\mathbf{r}^* \in \boldsymbol{\mathcal{R}}, \quad F(\mathbf{r}^*) = \inf_{\mathbf{r}\in\boldsymbol{\mathcal{R}}} F(\mathbf{r}). \tag{18}$$

The solution for this problem exists and is unique.

Proof. Let

$$m := \inf_{\mathbf{r}\in\boldsymbol{\mathcal{R}}} F(\mathbf{r}), \tag{19}$$

Algorithm 3 Bregmanized version of the iterative regularization Algorithm 2

Initialize $k = 0$, $\boldsymbol{\tau} = 0$, $\mathbf{r}_{ex}^0 = \nabla^\perp f$;
repeat
  $k = k + 1$;
  $\mathbf{r}^k = \arg\min_{\mathbf{r}\in BV(\Omega)} Q^{\mathbf{s}^{k-1}}(\mathbf{r})^k$;
  $\mathbf{r}_{ex}^k = \mathbf{r}_{ex}^{k-1} - \mathbf{r}^k$;
  $\mathbf{s}^k = \eta \mathbf{r}_{ex}^k$;
  $\boldsymbol{\tau} = \boldsymbol{\tau} + \mathbf{r}^k$;
until satisfied;
$u = \arg\min_{u\in BV(\Omega)}\{J(u) - \langle \nabla u, \frac{\boldsymbol{\tau}^\perp}{|\boldsymbol{\tau}^\perp|}\rangle + H(u, f)\}$;
return u.

and $\{\mathbf{r}^j\}$ is a minimizing sequence such that

$$\mathbf{r}^j \in \boldsymbol{\mathcal{R}}, \quad \lim_{j\to+\infty} F(\mathbf{r}^j) \to m. \tag{20}$$

Define an equivalent BV-norm [10] as

$$\|\mathbf{r}\|_{BV(\Omega;\mathbb{R}^2)} = \int_\Omega |\nabla \mathbf{r}| d\mathbf{x} + \|\mathbf{r}\|. \tag{21}$$

Followed from the above definition, we have that the sequence $\{\mathbf{r}^j\}$ is bounded in $BV(\Omega;\mathbb{R}^2)$, and consequently there exists a convergent subsequence $\{\mathbf{r}^k\}$ such that

$$\mathbf{r}^k \underset{\boldsymbol{\mathcal{R}}}{\rightharpoonup} \hat{\mathbf{r}} \tag{22}$$

By the lower semicontinuity of F, we have

$$\underline{\lim}_{\mathbf{r}^k \rightharpoonup \hat{\mathbf{r}}} F(\mathbf{r}^k) \geq F(\hat{\mathbf{r}}) \tag{23}$$

Due to (20), we obtain

$$m \geq F(\hat{\mathbf{r}}). \tag{24}$$

But owing to (19), $m \leq F(\hat{\mathbf{r}})$. Consequently $\hat{\mathbf{r}}$ is indeed a minimizer. Furthermore, because F is strictly convex, the solution is unique. ∎

Proposition 6. *Setting* $\mathbf{r}_{ex}^0 = \nabla^\perp f$, $\mathbf{s}^0 := \eta \mathbf{r}_{ex}^0$, *and* $\mathbf{q}^k = \partial H(\mathbf{r}^k, 0) = \eta \mathbf{r}^k$, *for each* $k \in \mathbb{N}$, *there is an unique minimizer* $\mathbf{r}^k$ *of* $Q^{\mathbf{s}^{k-1}}(\mathbf{r})^k$, *and a subgradient* $\mathbf{s}^k \in \partial J(\Pi \mathbf{r}^k)$ *such that*

$$\mathbf{s}^k + \mathbf{q}^k = \mathbf{s}^{k-1} \tag{25}$$

Proof. The well-definedness for each iteration follows directly from Lemma 5. The relation between $\mathbf{s}$ and $\mathbf{q}$ is proved by induction. For k = 1, we have

$$\arg\min_{\mathbf{r}\in BV(\Omega)} Q^{\mathbf{s}^0}(\mathbf{r})^1 = \arg\min_{\mathbf{r}\in BV(\Omega)} \{H(\mathbf{r}, \mathbf{r}_{ex}^0) + J(\Pi \mathbf{r})\}.$$

The relation $\mathbf{s}^1 + \mathbf{q}^1 = \mathbf{s}^0$ holds by defining $\mathbf{s}^1 = \eta \mathbf{r}_{ex}^1$ which exactly can be deduced by the relations of $\mathbf{q}^1 = \eta \mathbf{r}^1$, $\mathbf{s}^0 = \eta \mathbf{r}_{ex}^0$ and $\mathbf{r}_{ex}^1 = \mathbf{r}_{ex}^0 - \mathbf{r}^1$. Taking the observation of optimal condition for the case $k = 1$

$$\partial J(\Pi \mathbf{r}^1) + \eta \mathbf{r}^1 - \mathbf{s}^0 \ni 0,$$

which is the same as

$$\partial J(\Pi \mathbf{r}^1) \ni \eta \mathbf{r}_{ex}^1,$$

we have $\mathbf{s}^1 \in \partial J(\Pi \mathbf{r}^1)$, cf. [18] where ∂J is the subgradient of J in Euclidean space $\mathbb{R}^2$. Assuming that $\mathbf{s}^{k-1} = \eta \mathbf{r}_{ex}^{k-1} \in \partial J(\Pi \mathbf{r}^{k-1})$ holds, the k^{th} case is

$$\begin{aligned}
&\arg \min_{\mathbf{r} \in BV(\Omega)} Q^{\mathbf{s}^{k-1}}(\mathbf{r})^k \\
&\quad = \quad \arg \min_{\mathbf{r} \in BV(\Omega)} \{\frac{\eta}{2}\mathbf{r}^2 + J(\Pi \mathbf{r}) - J(\Pi \mathbf{r}^{k-1}) - \langle \eta \mathbf{r}_{ex}^{k-1}, \mathbf{r} - \mathbf{r}^{k-1} \rangle\} \\
&\quad = \quad \arg \min_{\mathbf{r} \in BV(\Omega)} \{\frac{\eta}{2}(\mathbf{r} - \mathbf{r}_{ex}^{k-1})^2 + J(\Pi \mathbf{r})\}.
\end{aligned}$$

The optimal condition is accordingly

$$\eta \mathbf{r}^k - \mathbf{s}^{k-1} + \partial J(\Pi \mathbf{r}^k) \ni 0.$$

Since $\mathbf{s}^{k-1} = \eta \mathbf{r}_{ex}^{k-1}$ and $\mathbf{q}^k = \partial H(\mathbf{r}^k, 0) = \eta \mathbf{r}^k$, it is easy to find that $\mathbf{s}^k = \eta(\mathbf{r}_{ex}^{k-1} - \mathbf{r}^k) = \eta \mathbf{r}_{ex}^k \in \partial J(\Pi \mathbf{r}^k)$, and thus we obtain (25). ∎

2.2.2. Convergence analysis. We define the generalized Bregman distance associated with $J(\Pi(\cdot))$ as follows

$$D^{\mathbf{s}}(\mathbf{w}, \mathbf{m}) := J(\Pi \mathbf{w}) - J(\Pi \mathbf{m}) - \langle \mathbf{s}, \mathbf{w} - \mathbf{m} \rangle,$$

where $\mathbf{s}$ is the subgradient for $J(\Pi \mathbf{m})$.

PROPOSITION 7. *The sequence $H(\mathbf{r}^k, 0)$ is monotonically nonincreasing, and*

$$H(\mathbf{r}^k, 0) \le H(\mathbf{r}^k, 0) + D^{\mathbf{s}^{k-1}}(\mathbf{r}^k, \mathbf{r}^{k-1}) \le H(\mathbf{r}^{k-1}, 0), \tag{26a}$$

$$D^{\mathbf{s}^k}(\mathbf{r}, \mathbf{r}^k) + D^{\mathbf{s}^{k-1}}(\mathbf{r}^k, \mathbf{r}^{k-1}) + H(\mathbf{r}^k, 0) \le H(\mathbf{r}, 0) + D^{\mathbf{s}^{k-1}}(\mathbf{r}, \mathbf{r}^{k-1}), \tag{26b}$$

subject to $k \in \mathbb{N} \setminus \{1\}$.

Proof. Since $D^{\mathbf{s}^{k-1}}(\mathbf{r}^k, \mathbf{r}^{k-1})$ is nonnegative, it is easy to find

$$H(\mathbf{r}^k, 0) \le H(\mathbf{r}^k, 0) + D^{\mathbf{s}^{k-1}}(\mathbf{r}^k, \mathbf{r}^{k-1}) = Q^{\mathbf{s}^{k-1}}(\mathbf{r}^k)^k.$$

Because $\mathbf{r}^k$ is the minimizer of $Q^{\mathbf{s}^{k-1}}(\mathbf{r})^k$, we have

$$Q^{\mathbf{s}^{k-1}}(\mathbf{r}^k)^k \le Q^{\mathbf{s}^{k-1}}(\mathbf{r}^{k-1})^k = H(\mathbf{r}^{k-1}, 0),$$

which implies (26a).

$$\begin{aligned}
&D^{\mathbf{s}^k}(\mathbf{r}, \mathbf{r}^k) - D^{\mathbf{s}^{k-1}}(\mathbf{r}, \mathbf{r}^{k-1}) + D^{\mathbf{s}^{k-1}}(\mathbf{r}^k, \mathbf{r}^{k-1}) \\
=&J(\Pi \mathbf{r}) - J(\Pi \mathbf{r}^k) - \langle \mathbf{s}^k, \mathbf{r} - \mathbf{r}^k \rangle \\
&-J(\Pi \mathbf{r}) + J(\Pi \mathbf{r}^{k-1}) + \langle \mathbf{s}^{k-1}, \mathbf{r} - \mathbf{r}^{k-1} \rangle \\
&+J(\Pi \mathbf{r}^k) - J(\Pi \mathbf{r}^{k-1}) - \langle \mathbf{s}^{k-1}, \mathbf{r}^k - \mathbf{r}^{k-1} \rangle \\
=&\langle \mathbf{s}^{k-1} - \mathbf{s}^k, \mathbf{r} - \mathbf{r}^k \rangle \\
=&\langle \mathbf{q}^k, \mathbf{r} - \mathbf{r}^k \rangle.
\end{aligned}$$

The relation $\mathbf{s}^{k-1} - \mathbf{s}^k = \mathbf{q}^k$ has been used here according to Proposition 25. The $\mathbf{q}^k$ is the subgradient of $H(\mathbf{r}^k, 0)$. By the definition of subgradient, we have

$$D^{\mathbf{s}^k}(\mathbf{r}, \mathbf{r}^k) - D^{\mathbf{s}^{k-1}}(\mathbf{r}, \mathbf{r}^{k-1}) + D^{\mathbf{s}^{k-1}}(\mathbf{r}^k, \mathbf{r}^{k-1}) = \langle \mathbf{q}^k, \mathbf{r} - \mathbf{r}^k \rangle \leq H(\mathbf{r}, 0) - H(\mathbf{r}^k, 0),$$

and thus we obtain (26b). ∎

There is a direct result from this relation (26b). If there exists a minimizer $\mathbf{r}$ of $H(\cdot, 0)$, by using (26b), we have

$$\begin{aligned} D^{\mathbf{s}^k}(\mathbf{r}, \mathbf{r}^k) &\leq D^{\mathbf{s}^k}(\mathbf{r}, \mathbf{r}^k) + D^{\mathbf{s}^{k-1}}(\mathbf{r}^k, \mathbf{r}^{k-1}) \\ &\leq D^{\mathbf{s}^k}(\mathbf{r}, \mathbf{r}^k) + D^{\mathbf{s}^{k-1}}(\mathbf{r}^k, \mathbf{r}^{k-1}) + H(\mathbf{r}^k, 0) - H(\mathbf{r}, 0) \\ &\leq D^{\mathbf{s}^{k-1}}(\mathbf{r}, \mathbf{r}^{k-1}). \end{aligned} \tag{27}$$

It implies that, for each iteration, the Bregman distance to optimal $\mathbf{r}$ is getting shorter.

THEOREM 8. *If $\mathbf{r} \in BV(\Omega; \mathbb{R}^2)$ is the minimizer of $H(\cdot, 0)$ subject to $k \in \mathbb{N} \setminus \{1\}$, then $\mathbf{r}^k$ converges and*

$$H(\mathbf{r}^k, 0) \leq \frac{J(\Pi \mathbf{r}) - J(\Pi \mathbf{r}^1) - \langle \mathbf{s}^1, \mathbf{r} - \mathbf{r}^1 \rangle}{k-1}, \tag{28}$$

moreover,

$$\boldsymbol{\tau}^k = \sum_{i=1}^{k} \mathbf{r}^i,$$

converges to $\boldsymbol{\tau}^0$.

Proof. Taking the sum of (26b), we obtain

$$D^{\mathbf{s}^k}(\mathbf{r}, \mathbf{r}^k) + \sum_{i=2}^{k} \left[D^{\mathbf{s}^{i-1}}(\mathbf{r}^i, \mathbf{r}^{i-1}) + H(\mathbf{r}^i, 0) - H(\mathbf{r}, 0) \right] \leq D^{\mathbf{s}^1}(\mathbf{r}, \mathbf{r}^1). \tag{29}$$

Since $H(\mathbf{r}^k, 0)$ is monotonically nonincreasing,

$$(k-1) \left[D^{\mathbf{s}^{k-1}}(\mathbf{r}^k, \mathbf{r}^{k-1}) + H(\mathbf{r}^k, 0) - H(\mathbf{r}, 0) \right] \leq J(\Pi \mathbf{r}) - J(\Pi \mathbf{r}^1) - \langle \mathbf{s}^1, \mathbf{r} - \mathbf{r}^1 \rangle.$$

Because $\mathbf{r}$ is the minimizer of $H(\cdot, 0)$ and $D^{\mathbf{s}^{k-1}}(\mathbf{r}^k, \mathbf{r}^{k-1})$ is nonnegative, we obtain (28). It implies that, when $k \to \infty$, $\mathbf{r}^k$ converges to 0 with rate

$$\|\mathbf{r}^k\| \leq \sqrt{\frac{J(\Pi \mathbf{r}) - J(\Pi \mathbf{r}^1) - \langle \mathbf{s}^1, \mathbf{r} - \mathbf{r}^1 \rangle}{k-1}} = \mathcal{O}((k-1)^{-1/2}).$$

From the definition,

$$\mathbf{r}^k = \arg\min_{\mathbf{r}} \left\{ \int_\Omega |\nabla \Pi(\mathbf{r})| + \frac{\eta}{2} (\mathbf{r} - (\nabla^\perp f - \sum_{i=1}^{k-1} \mathbf{r}^i))^2 d\mathbf{x} \right\},$$

we obtain the optimal condition

$$\partial J(\Pi(\mathbf{r}^k)) + \eta(\mathbf{r}^k - (\nabla^\perp f - \sum_{i=1}^{k-1} \mathbf{r}^i)) \ni 0.$$

Since $\mathbf{r}^k$ converges to 0 while $k \to \infty$ and, and therefor $\partial J(\Pi(\mathbf{r}^\infty)) = 0$, we obtain $\sum_{i=1}^{\infty} \mathbf{r}^i = \nabla^\perp f$. ∎

2.3. Iterative regularization for the second step of TV-Stokes. In this subsection, we consider the Richardson iteration on the second step of TV-Stokes model. Similar to Algorithm 2, the proposed algorithm is accordingly the following as Algorithm 4.

Algorithm 4 Iterative regularization applied to the 2^{nd} step of TV-Stokes

$\boldsymbol{\tau} = \arg\min_{\boldsymbol{\tau} \in BV(\Omega;\mathbb{R}^2)} \{J(\Pi\boldsymbol{\tau}) + H(\boldsymbol{\tau}, \nabla^\perp f)\}$;
Initialize $k = 0$, $u = 0$, $r_{ex}^0 = f$;
repeat
  $k = k + 1$;
  $r^k = \arg\min_{r \in BV(\Omega;\mathbb{R})} \{J(r) + H^k(r, r_{ex}^{k-1} - \frac{\alpha}{\eta^k} \nabla \cdot \frac{\boldsymbol{\tau}^\perp}{|\boldsymbol{\tau}^\perp|})\}$;
  $r_{ex}^k = r_{ex}^{k-1} - r^k$;
  $u = u + r^k$;
until satisfied;
return u.

2.3.1. Well-definedness of iterates. Let us start with a simple case without iteration, specifically for a fixed k. The considered minimization problem is shown as (31). For a such given problem, we can find the solution exists and is unique.

LEMMA 9. *Let* $\mathcal{R} = \{r | r \in BV(\Omega;\mathbb{R})\}$, $F(r) = J(r) + H(r, z)$. *Consider the problem to find* r^* *such that*

$$r^* \in \mathcal{R}, \quad F(r^*) = \inf_{r \in \mathcal{R}} F(r). \tag{31}$$

The solution for this problem exists and is unique.

Proof. Let

$$\hat{m} := \inf_{r \in \mathcal{R}} F(r), \tag{32}$$

and $\{r^j\}$ is a minimizing sequence such that

$$r^j \in \mathcal{R}, \quad \lim_{j \to +\infty} F(r^j) \to \hat{m}. \tag{33}$$

Define an equivalent BV-norm as

$$\|r\|_{BV(\Omega;\mathbb{R})} = \int_\Omega |\nabla r| d\mathbf{x} + \|r\|. \tag{34}$$

Followed from the above definition, we have that the sequence $\{r^j\}$ is bounded in $BV(\Omega;\mathbb{R})$, and consequently there exists a convergent sub-sequence $\{r^k\}$ such that

$$r^k \underset{\mathcal{R}}{\rightharpoonup} \hat{r} \tag{35}$$

By the lower semicontinuity of F, we have

$$\lim_{r^k \to \hat{r}} F(r^k) \geq F(\hat{r}) \tag{36}$$

Due to relation (33), we obtain

$$\hat{m} \geq F(\hat{r}). \tag{37}$$

But owing to (32), $\hat{m} \leq F(\hat{r})$. Consequently $\hat{r}$ is indeed a minimizer. Furthermore, because F is strictly convex, the solution is unique. ∎

The well-definedness for each iteration follows directly from the above Lemma 9. In the iterations, we choose the fidelity parameter η^k for each iteration as shown in Lemma 4 such that $\eta^k = \max(\beta/\gamma, \eta^{k-1})$, where $\beta \in (1, +\infty)$.

Consider the following minimizations for iterations k and $k+1$ ($k \in \mathbb{N}$)

$$r^k = \arg\min_{r \in BV(\Omega;\mathbb{R})} \left\{ J(r) + H^k(r, r_{ex}^{k-1} - \frac{\alpha}{\eta^k} \nabla \cdot \frac{\boldsymbol{\tau}^\perp}{|\boldsymbol{\tau}^\perp|}) \right\}; \tag{38}$$

$$\tilde{r}^{k+1} = \arg\min_{r \in BV(\Omega;\mathbb{R})} \left\{ J(r) + H^k(r, r_{ex}^{k} - \frac{\alpha}{\eta^k} \nabla \cdot \frac{\boldsymbol{\tau}^\perp}{|\boldsymbol{\tau}^\perp|}) \right\}; \tag{39}$$

$$r^{k+1} = \arg\min_{r \in BV(\Omega;\mathbb{R})} \left\{ J(r) + H^{k+1}(r, r_{ex}^{k} - \frac{\alpha}{\eta^{k+1}} \nabla \cdot \frac{\boldsymbol{\tau}^\perp}{|\boldsymbol{\tau}^\perp|}) \right\}. \tag{40}$$

The Euler-Lagrangian equation for k iteration (38) is

$$\partial J(r^k) + \eta^k (r^k - r_{ex}^{k-1} + \frac{\alpha}{\eta^k} \nabla \cdot \frac{\boldsymbol{\tau}^\perp}{|\boldsymbol{\tau}^\perp|}) \ni 0.$$

The subgradient of J thus can be determined as $s^k := \partial J(r^k) = \eta^k (r_{ex}^k - \frac{\alpha}{\eta^k} \nabla \cdot \frac{\boldsymbol{\tau}^\perp}{|\boldsymbol{\tau}^\perp|})$ for $k \in \mathbb{N}$. When $k = 1$, we set $r_{ex}^0 := f$.

Definition 10. *Define* $\tilde{Q}^{\tilde{s}^k}(r)^{k+1} := H^k(r,0) + J(r) - J(r^k) - \langle s^k, r - r^k \rangle$. *Setting* $\tilde{q}^{k+1} := \partial H^k(\tilde{r}^{k+1}, 0) = \eta^k \tilde{r}^{k+1}$, *for each* $k \in \mathbb{N}$, *there is an unique minimizer* $\tilde{r}^{k+1}$ *of* $\tilde{Q}^{\tilde{s}^k}(r)^{k+1}$, *and a subgradient* $\tilde{s}^{k+1} \in \partial J(\tilde{r}^{k+1})$ *such that*

$$\tilde{s}^{k+1} + \tilde{q}^{k+1} = s^k \tag{41}$$

The relation of $\tilde{s}$, s and $\tilde{q}$ is easy to be obtained.

Lemma 11. *For a given* $k \in \mathbb{N} \setminus \{1\}$, *assume* $r_{ex}^k, r^{k+1} \in L^2(\Omega;\mathbb{R})$ *such that* $r^{k+1} \sim N(0, (\sigma^{k+1})^2)$, $r_{ex}^k \sim N(r^{k+1}, (\sigma_{ex}^k)^2)$, *then* $\|r_{ex}^{k+1}\| \leq \|r_{ex}^k\|$.

Proof. Since $r_{ex}^k \sim N(r^{k+1}, (\sigma_{ex}^k)^2)$, $r_{ex}^{k+1} = r_{ex}^k - r^{k+1}$ is a Gaussian distribution such that $r_{ex}^{k+1} \sim N(0, (\sigma_{ex}^k)^2)$. r^{k+1} as given is also a Gaussian distribution such that $r^{k+1} \sim N(0, (\sigma^{k+1})^2)$. Since r^{k+1} and $r_{ex}^k - r^{k+1}$ are two independent Gaussian distributions, the sum $r_{ex}^k = r^{k+1} + (r_{ex}^k - r^{k+1})$ is another Gaussian distribution such that $r_{ex}^k \sim N(0, (\sigma_{ex}^k)^2 + (\sigma^{k+1})^2)$.

We obtain $\|r_{ex}^{k+1}\| \leq \|r_{ex}^k\|$ since $(\sigma_{ex}^k)^2 \leq (\sigma_{ex}^k)^2 + (\sigma^{k+1})^2$. ∎

2.3.2. Convergence analysis. We define the generalized Bregman distance associated with $J(\cdot)$ as follows

$$D^s(w, m) := J(w) - J(m) - \langle s, w - m \rangle.$$

Proposition 12.

$$H^k(\tilde{r}^{k+1}, 0) \leq H^k(\tilde{r}^{k+1}, 0) + D^{s^k}(\tilde{r}^{k+1}, r^k) \leq H^k(r^k, 0), \tag{42}$$

$$\begin{aligned} &D^{\tilde{s}^{k+1}}(r, \tilde{r}^{k+1}) + D^{\tilde{s}^k}(\tilde{r}^{k+1}, \tilde{r}^k) + H^k(\tilde{r}^{k+1}, 0) \\ \leq &\frac{\eta^k - \eta^{k-1}}{\eta^k} H^k(r_{ex}^k, 0) + D^{\tilde{s}^k}(r, \tilde{r}^k) + (\eta^k - \eta^{k-1}) \langle \tilde{r}^{k+1} - r_{ex}^k, r \rangle, \end{aligned} \tag{43}$$

subject to $J(r) < \infty$ *and* $k \in \mathbb{N}$.

Proof. Since $D^{s^k}(\tilde{r}^{k+1}, r^k)$ is non-negative, it is easy to find

$$H^k(\tilde{r}^{k+1}, 0) \leq H^k(\tilde{r}^{k+1}, 0) + D^{s^k}(\tilde{r}^{k+1}, r^k) = \tilde{Q}^{s^k}(\tilde{r}^{k+1})^{k+1}.$$

Because $\tilde{r}^{k+1}$ is the minimizer of $\tilde{Q}^{s^k}(r)^{k+1}$, we have

$$\tilde{Q}^{s^k}(\tilde{r}^{k+1})^{k+1} \leq \tilde{Q}^{s^k}(r^k)^{k+1} = H^k(r^k, 0),$$

which implies (42).

$$\begin{aligned}
&D^{\tilde{s}^{k+1}}(r, \tilde{r}^{k+1}) - D^{\tilde{s}^k}(r, \tilde{r}^k) + D^{\tilde{s}^k}(\tilde{r}^{k+1}, \tilde{r}^k) \\
&= J(r) - J(\tilde{r}^{k+1}) - \langle \tilde{s}^{k+1}, r - \tilde{r}^{k+1} \rangle \\
&\quad - J(r) + J(\tilde{r}^k) + \langle \tilde{s}^k, r - \tilde{r}^k \rangle \\
&\quad + J(\tilde{r}^{k+1}) - J(\tilde{r}^k) - \langle \tilde{s}^k, \tilde{r}^{k+1} - \tilde{r}^k \rangle \\
&= \langle \tilde{s}^k - \tilde{s}^{k+1}, r - \tilde{r}^{k+1} \rangle \\
&= \langle \tilde{s}^k - s^k + \tilde{q}^{k+1}, r - \tilde{r}^{k+1} \rangle \\
&= \langle \tilde{s}^k - s^k, r - \tilde{r}^{k+1} \rangle + \langle \tilde{q}^{k+1}, r - \tilde{r}^{k+1} \rangle.
\end{aligned}$$

Substituting $\tilde{s}^k = \eta^{k-1}(r_{ex}^k - \frac{\alpha}{\eta^{k-1}} \nabla \cdot \frac{\boldsymbol{\tau}^\perp}{|\boldsymbol{\tau}^\perp|})$ and $s^k = \eta^k(r_{ex}^k - \frac{\alpha}{\eta^k} \nabla \cdot \frac{\boldsymbol{\tau}^\perp}{|\boldsymbol{\tau}^\perp|})$ into the above transformation, we obtain

$$\begin{aligned}
&D^{\tilde{s}^{k+1}}(r, \tilde{r}^{k+1}) - D^{\tilde{s}^k}(r, \tilde{r}^k) + D^{\tilde{s}^k}(\tilde{r}^{k+1}, \tilde{r}^k) \\
&= (\eta^k - \eta^{k-1}) \langle r_{ex}^k, \tilde{r}^{k+1} - r \rangle + \langle \tilde{q}^{k+1}, r - \tilde{r}^{k+1} \rangle \\
&= \frac{\eta^k - \eta^{k-1}}{\eta^k} \langle r_{ex}^k - \tilde{r}^{k+1}, \tilde{q}^{k+1} \rangle + \frac{\eta^{k-1}}{\eta^k} \langle \tilde{q}^{k+1}, r - \tilde{r}^{k+1} \rangle + (\eta^k - \eta^{k-1}) \langle \tilde{r}^{k+1} - r_{ex}^k, r \rangle.
\end{aligned}$$

The $\tilde{q}^{k+1}$ is the subgradient of $H^k(\tilde{r}^{k+1}, 0)$. By the definition of subgradient, we have

$$\begin{aligned}
&D^{\tilde{s}^{k+1}}(r, \tilde{r}^{k+1}) - D^{\tilde{s}^k}(r, \tilde{r}^k) + D^{\tilde{s}^k}(\tilde{r}^{k+1}, \tilde{r}^k) - (\eta^k - \eta^{k-1}) \langle \tilde{r}^{k+1} - r_{ex}^k, r \rangle \\
\leq &\frac{\eta^k - \eta^{k-1}}{\eta^k} (H^k(r_{ex}^k, 0) - H^k(\tilde{r}^{k+1}, 0)) + \frac{\eta^{k-1}}{\eta^k} (H^k(r, 0) - H^k(\tilde{r}^{k+1}, 0)) \\
= &\frac{\eta^k - \eta^{k-1}}{\eta^k} H^k(r_{ex}^k, 0) - H^k(\tilde{r}^{k+1}, 0).
\end{aligned}$$

∎

The proposition (42) implies $\|\tilde{r}^{k+1}\| \leq \|r^k\|$ corresponding to (38) and (39). By Lemma 3, we can obtain $\|r^{k+1} - r_{ex}^k\| \leq \|\tilde{r}^{k+1} - r_{ex}^k\|$. With a Gaussian assumption, cf. Lemma 2, we have $\|r^{k+1}\| \leq \|\tilde{r}^{k+1}\|$. The iteration series $\|r^i\|$, $i \in \mathbb{N}$, is therefore non-increase, i.e., $\|r^{i+1}\| \leq \|r^i\|$.

If there exists a minimizer r of $H(\cdot, 0)$ with $J(r) < \infty$, by using (43), we have

$$\begin{aligned}
D^{\tilde{s}^{k+1}}(r, \tilde{r}^{k+1}) &\leq D^{\tilde{s}^{k+1}}(r, \tilde{r}^{k+1}) + D^{\tilde{s}^k}(\tilde{r}^{k+1}, \tilde{r}^k) \\
&\leq D^{\tilde{s}^{k+1}}(r, \tilde{r}^{k+1}) + D^{\tilde{s}^k}(\tilde{r}^{k+1}, \tilde{r}^k) \\
&\quad + H^k(\tilde{r}^{k+1}, 0) - (\eta^k - \eta^{k-1}) \langle \tilde{r}^{k+1} - r_{ex}^k, r \rangle \\
&\leq D^{\tilde{s}^k}(r, \tilde{r}^k) + \frac{\eta^k - \eta^{k-1}}{\eta^k} H^k(r_{ex}^k, 0).
\end{aligned} \tag{44}$$

THEOREM 13. *If $r \in BV(\Omega)$ is the minimizer of $H(\cdot, 0)$ subject to $k \in \mathbb{N}$, then r^k converges and*

$$\|r^k\|^2 \leq \frac{2D^{\tilde{s}^1}(r, \tilde{r}^1) + 2\beta\|r_{ex}^0\|_\infty}{k\eta^1}, \tag{45}$$

moreover,

$$u^k = \sum_{i=1}^{k} r^i,$$

converges to f.

Proof. Taking the sum of (43) as follows

$$\sum_{i=1}^{k}\left[D^{\tilde{s}^{i+1}}(r, \tilde{r}^{i+1}) - D^{\tilde{s}^i}(r, \tilde{r}^i) + D^{\tilde{s}^i}(\tilde{r}^{i+1}, \tilde{r}^i) + H^i(\tilde{r}^{i+1}, 0) - \frac{\eta^i - \eta^{i-1}}{\eta^i}H^i(r_{ex}^i, 0)\right],$$

we obtain

$$D^{\tilde{s}^{k+1}}(r, \tilde{r}^{k+1}) + \sum_{i=1}^{k}\left[D^{\tilde{s}^i}(\tilde{r}^{i+1}, \tilde{r}^i) + H^i(\tilde{r}^{i+1}, 0) - \frac{\eta^i - \eta^{i-1}}{\eta^i}H^i(r_{ex}^i, 0)\right] \leq D^{\tilde{s}^1}(r, \tilde{r}^1). \tag{46}$$

Due to the non-negativity of Bregmann distance, the above inequality can be rewritten as follows

$$\sum_{i=1}^{k}\left[H^i(\tilde{r}^{i+1}, 0) - \frac{\eta^i - \eta^{i-1}}{\eta^i}H^i(r_{ex}^i, 0)\right] \leq D^{\tilde{s}^1}(r, \tilde{r}^1). \tag{47}$$

Substituting $\eta^i = \frac{\beta}{\gamma^i} \sim \frac{2\beta}{\|r_{ex}^i\|_\infty}$ into the above inequality, we obtain

$$\begin{aligned}
&\sum_{i=1}^{k}\left[H^i(\tilde{r}^{i+1}, 0) - \frac{\eta^i - \eta^{i-1}}{2}\|r_{ex}^i\|^2\right] \leq D^{\tilde{s}^1}(r, \tilde{r}^1),\\
\Rightarrow &\sum_{i=1}^{k}\left[H^i(\tilde{r}^{i+1}, 0) - \frac{\beta(\|r_{ex}^{i-1}\|_\infty - \|r_{ex}^i\|_\infty)}{\|r_{ex}^i\|_\infty\|r_{ex}^{i-1}\|_\infty}\|r_{ex}^i\|^2\right] \leq D^{\tilde{s}^1}(r, \tilde{r}^1),\\
\Rightarrow &\sum_{i=1}^{k}\left[\frac{\eta^i}{2}\|\tilde{r}^{i+1}\|^2 - \beta(\|r_{ex}^{i-1}\|_\infty - \|r_{ex}^i\|_\infty)\right] \leq D^{\tilde{s}^1}(r, \tilde{r}^1).
\end{aligned}$$

Since $\|\tilde{r}^k\|$ is monotonically nonincreasing, it results

$$k\eta^1\|\tilde{r}^k\|^2 \leq 2D^{\tilde{s}^1}(r, \tilde{r}^1) + 2\beta(\|r_{ex}^0\|_\infty - \|r_{ex}^k\|_\infty) \leq 2D^{\tilde{s}^1}(r, \tilde{r}^1) + 2\beta\|r_{ex}^0\|_\infty.$$

We obtain (45). It implies that, when $k \to \infty$, r^k converges to 0 with rate

$$\|r^k\| \leq \|\tilde{r}^k\| \leq \sqrt{\frac{2D^{\tilde{s}^1}(r, \tilde{r}^1) + 2\beta\|r_{ex}^0\|_\infty}{k\eta^1}} = \mathcal{O}(k^{-1/2}).$$

∎

When r^k converges to 0, it is easy to find u^k converges to f by making a contradiction against Lemma (4).

REMARK 2. *In the proof of* THEOREM 13*, defining* $\eta^i \sim \frac{2\beta}{\|r^i_{ex}\|_\infty}$*, only the case* $\eta^{i-1} <= \eta^i$ *is considered, otherwise, as our setting* $\eta^i = \max(\beta/\gamma, \eta^{i-1})$*,* $\eta^{i-1} = \eta^i$ *so that the convergence follows directly by the fact that the term* $\frac{\eta^i - \eta^{i-1}}{\eta^i} H^i(r^i_{ex}, 0)$ *vanishes in (47).*

2.4. Iterative regularization applied separately to each of TV-Stokes steps. The Richardson-like iterations are applied separately on both two steps of TV-Stokes model is listed below, cf. Algorithm 5. The properties of well-definedness and convergence naturally follow from the separated cases addressed in previous subsections.

Algorithm 5 Separated iterative regularization for TV-Stokes

Initialize $k = 0$, $\mathbf{r}^0_{ex} = \nabla^\perp f$, $\boldsymbol{\tau} = 0$;
repeat
  $k = k + 1$;
  $\mathbf{r}^k = \arg\min_{\mathbf{r}\in BV(\Omega)}\{J(\Pi\mathbf{r}) + H(\mathbf{r}, \mathbf{r}^{k-1}_{ex})\}$;
  $\mathbf{r}^k_{ex} = \mathbf{r}^{k-1}_{ex} - \mathbf{r}^k$;
  $\boldsymbol{\tau} = \boldsymbol{\tau} + \mathbf{r}^k$;
until satisfied
Initialize $k = 0$, $r^0_{ex} := f$, $u = 0$;
repeat
  $k = k + 1$;
  $r^k = \arg\min_{r\in BV(\Omega)}\{J(r) - \alpha\nabla \cdot \frac{\boldsymbol{\tau}^\perp}{|\boldsymbol{\tau}^\perp|} + H^k(r, r^{k-1}_{ex})\}$;
  $r^k_{ex} = r^{k-1}_{ex} - r^k$;
  $u = u + r^k$;
until satisfied
return u.

3. Numerical experiments. In this section, we present our experiments on the effectiveness of proposed algorithms on smooth structures, e.g., Lena's face, on their capability in preserving both sharp edges and smooth patterns, e.g., fingerprint with clean surrounding, and finally on structures mixed with pinstripes and smooth surfaces, cf. e.g., figures of Barbara.

In our experiments, we employ the dual-formula-based method to solve the TV-Stokes model, cf. [12], where we keep the step sizes for the line search the same as 1/4 throughout the experiments. The noise resource considered in this paper is of Gaussian.

We start the experiment with Algorithm 1, applying Osher's iterative regularization algorithm on a Lena portrait, cf. FIG. 1. The associated ROF model is solved via the Chambolle dual formula, cf. [5] for the details. The initial noise level is 7.97 while the corresponding Peak Signal-Noise Ratio (PSNR) is 30.79. In this paper, the PSNR number is calculated via the MATLAB function `psnr`. The curve $\|u - g\|$ shows an optimal solution at the iteration 33 where the restored image u is most close to the clean image g in L^2, the curve $\|u - f\|$ shows the resulted image u through the iteration is converging to the initial image f. The restored image via this experiment is at noise level 5.46 and with PSNR 34.08. The result suffers the effect of stair-case inherited from ROF model.

For verifying the effectiveness of our proposed algorithms, we apply Algorithm 2,

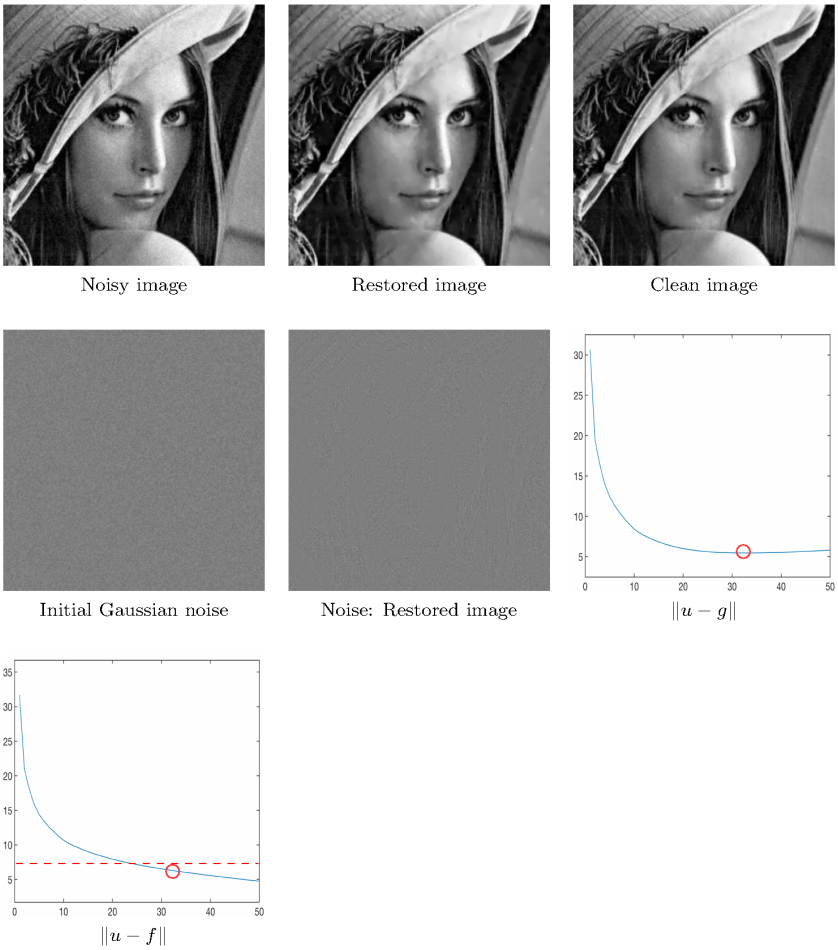

FIG. 1. *Showing the effect of applying the Osher-like iterative regularization on the ROF model, cf. Algorithm 1. The initial image is of a Gaussian noise at noise level* 7.97 *and with* $PSNR =$ 30.79. *The final restored image is at noise level* 5.46 *and with* $PSNR = 34.08$.

Algorithm 4, and Algorithm 5, respectively, on the same image, the Lena portrait, with the same noise.

We first apply Algorithm 2 on the Lena portrait, that is using the Richardson-like iterative regularization on the first step of the TV-Stokes model, cf. FIG. 2. The parameter β is set to be 6.5 and 3 for the first step and the second step, respectively. The parameter α is given to be 0.9. We observe that the optimal solution for the first step achieves at the iteration 13 in our experiment. The final restored image is at noise level 5.10 and with PSNR 34.66. The result shows an improvement of the smoothness of Lena's face.

The next experiment is applying the Algorithm 4 on the same Lena portrait,

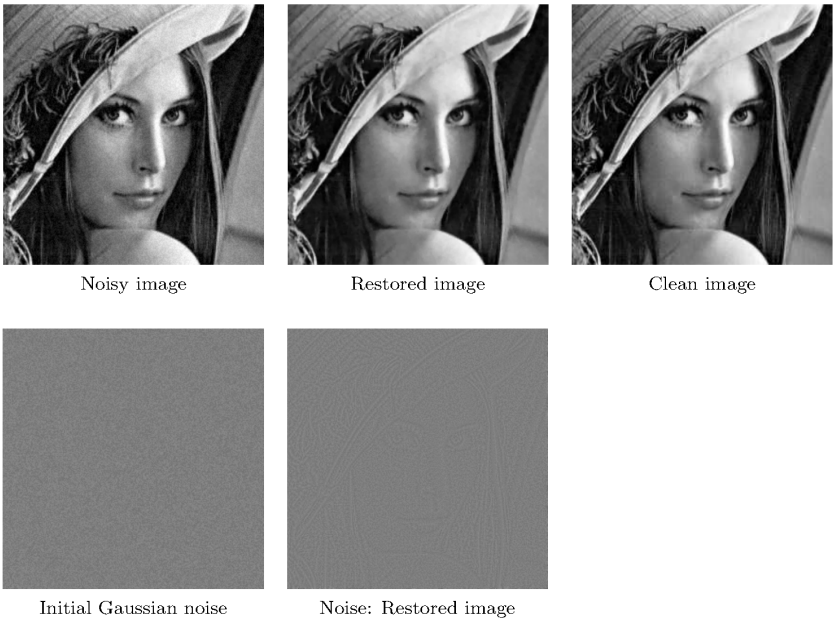

FIG. 2. *Showing the effect of applying the iterative regularization on the first step of TV-Stokes model, cf. Algorithm 2. The initial image is of a Gaussian noise at noise level* 7.97 *and with* $PSNR = 30.79$. *The final restored image is of noise level* 5.10 *and* $PSNR = 34.66$.

that is using the Richardson-like iterative regularization on the second step of the TV-Stokes model, cf. FIG. 3. The parameter β is set to be 8.0 and 2.5 for the first step and the second step, respectively. The parameter α is given to be 0.9. We observe that the optimal solution for the second step achieves at the iteration 12 in our experiment. The curve $\|u - g\|$ shows an optimal solution at the iteration 12 where the restored image u is most close to the clean image g in L^2, the curve $\|u - f\|$ shows the resulted image u through the iteration is converging to the initial image f. The final restored image is at noise level 4.90 and with PSNR 35.01. The result shows a visible improvement both in the smoothness of Lena's face and in the preserving of details of the hat, comparing to FIG. 1 via Osher-like iterative regularization. The textures and fine structures are observed to be added back to the restored image accumulatively through the iterations.

The last experiment on the same noisy Lena portrait is applying the Algorithm 5, that is using the Richardson-like iterative regularization on both the first step and the second step of the TV-Stokes model, cf. FIG. 4. The parameter β is set to be 6.5 and 2.5 for the first step and the second step, respectively. The parameter α is given to be 0.9. We observe that the optimal solution for the first step achieves at the iteration 13 and for the second step achieves at the iteration 3 in our experiment. The curve $\|u - g\|$ shows an optimal solution at the iteration 3 where the restored image u is most close to the clean image g in L^2, the curve $\|u - f\|$ shows the resulted image u through the iteration is converging to the initial image f. The final restored image is at noise level 4.94 and with PSNR 34.95. The result also shows a visible

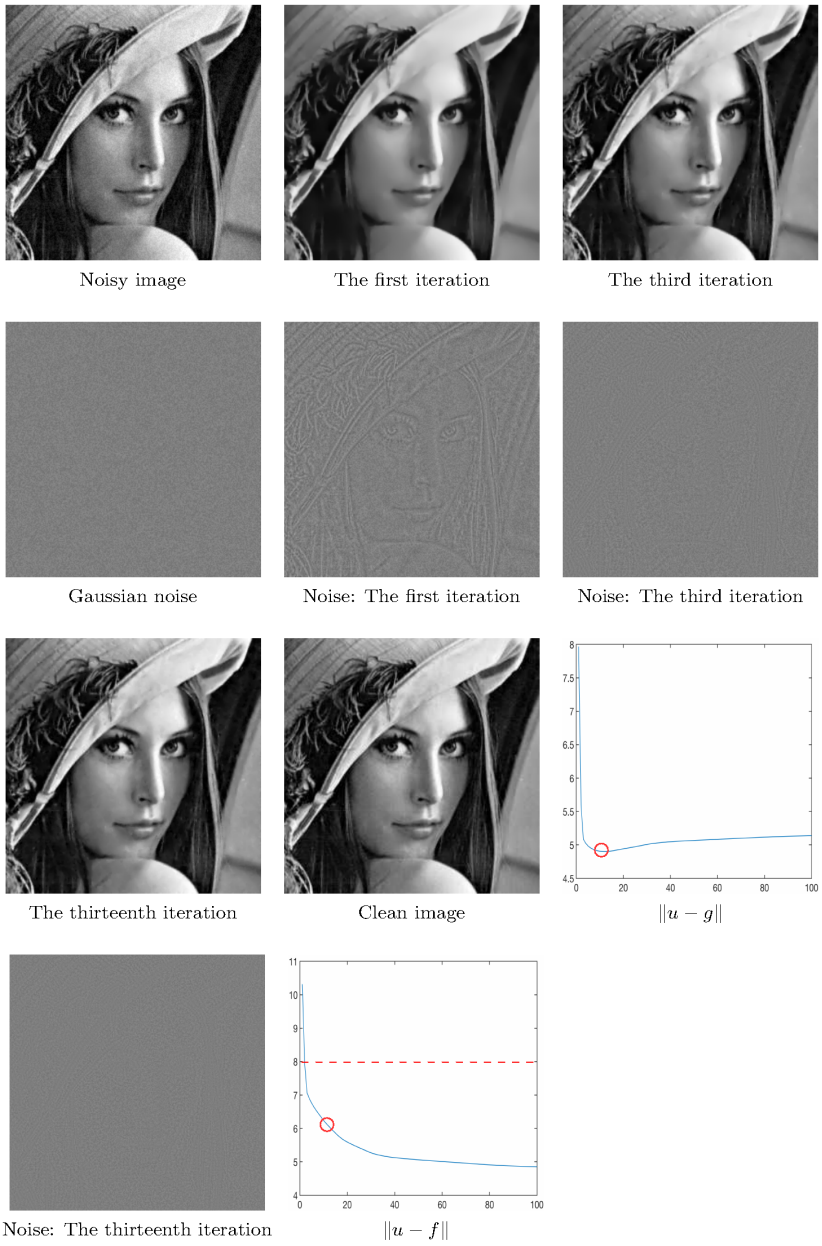

FIG. 3. *Showing the effect of applying the iterative regularization on the second step of TV-Stokes model, cf. Algorithm 4. The initial image is of a Gaussian noise at noise level* 7.97 *and with* $PSNR = 30.79$. *The final restored image is of noise level* 4.90 *and* $PSNR = 35.01$.

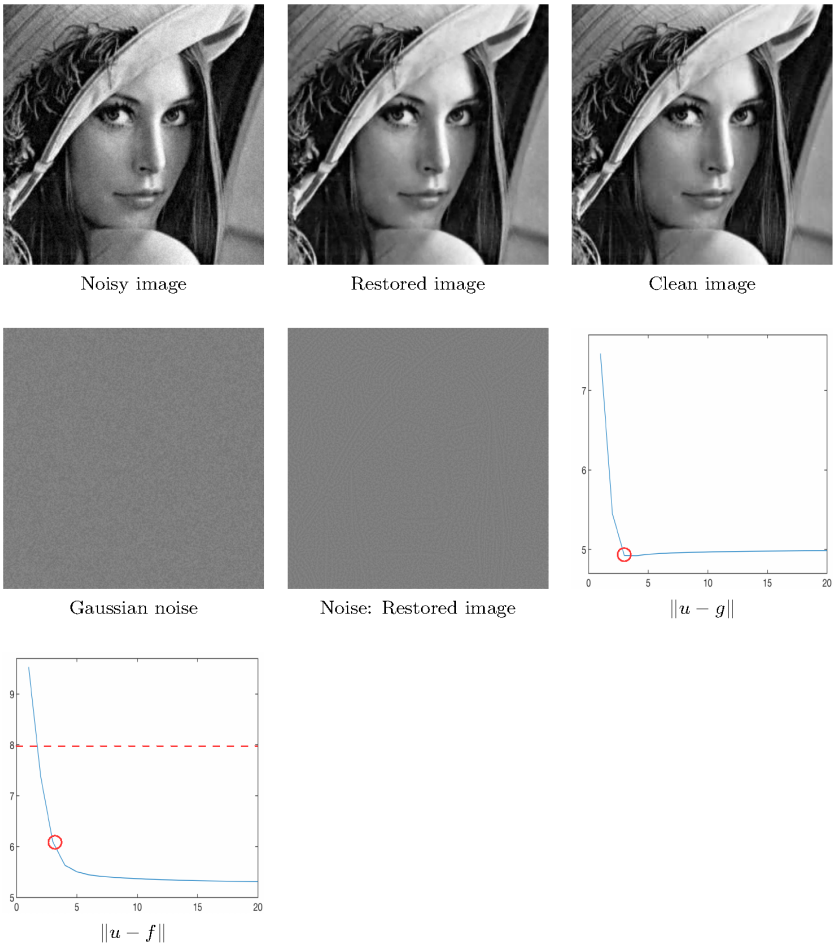

FIG. 4. *Showing the effect of applying the iterative regularization onto both the first step and the second step of TV-Stokes model, cf. Algorithm 5. The initial image is of a Gaussian noise at noise level* 7.97 *and with* $PSNR = 30.79$. *The final restored image is of noise level* 4.94 *and* $PSNR = 34.95$.

improvement both in the smoothness of Lena's face and in the preserving of details of the hat, comparing to FIG. 1 via Osher-like iterative regularization.

For convenience in comparing the results from the different algorithms, we collect all the restored images and the clean image together, cf. FIG. 5. All the proposed algorithms have a visible improvement in handling the staircase effect compared to Algorithm 1. Among them, the Algorithm 4 results of the best restoration concerning both the PSNR value and visual pleasure.

The other group of experiments is on the same Lena portrait but with heavier noise, cf. FIG. 6 and FIG. 7, where the noise level is at 15.71 and PSNR is 24.89.

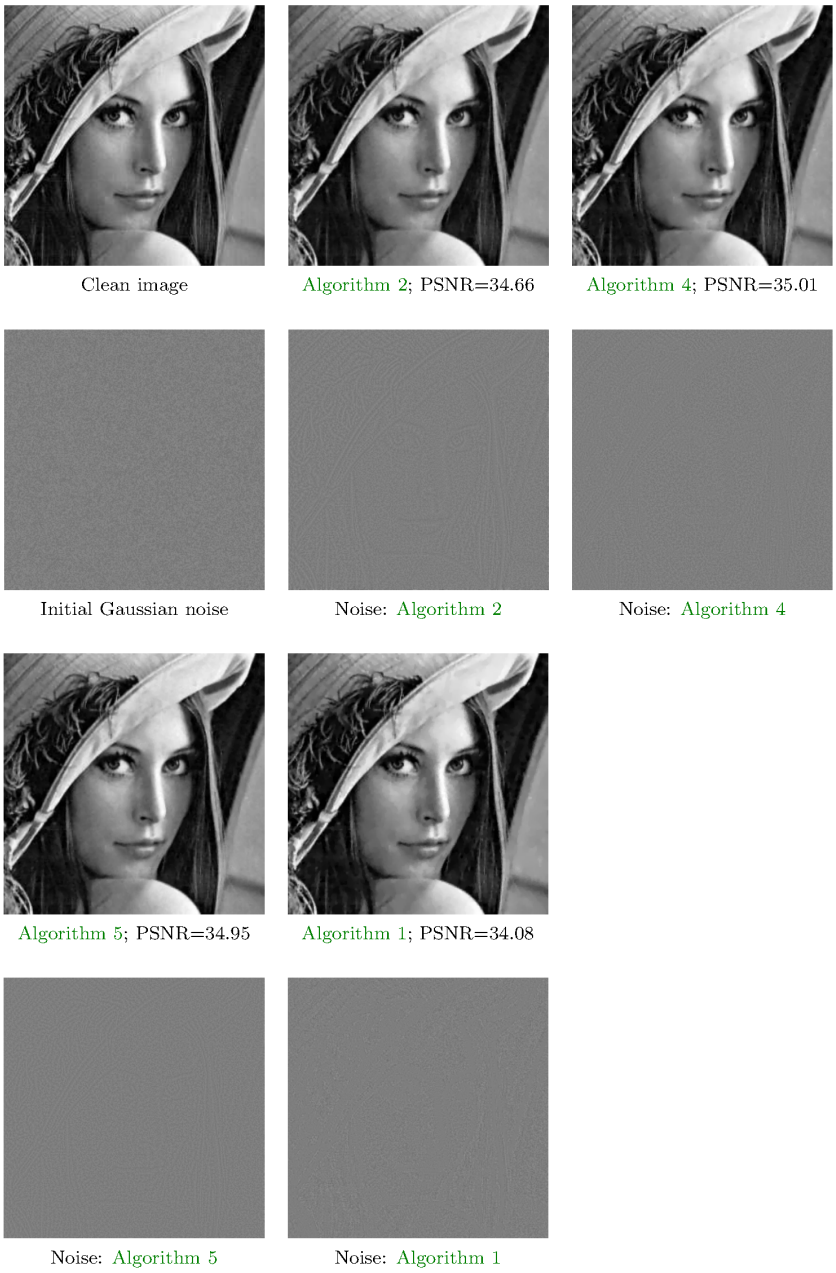

FIG. 5. *Comparing the effect of different iterative algorithms on the noisy Lena image. The initial noise level is* 7.96.

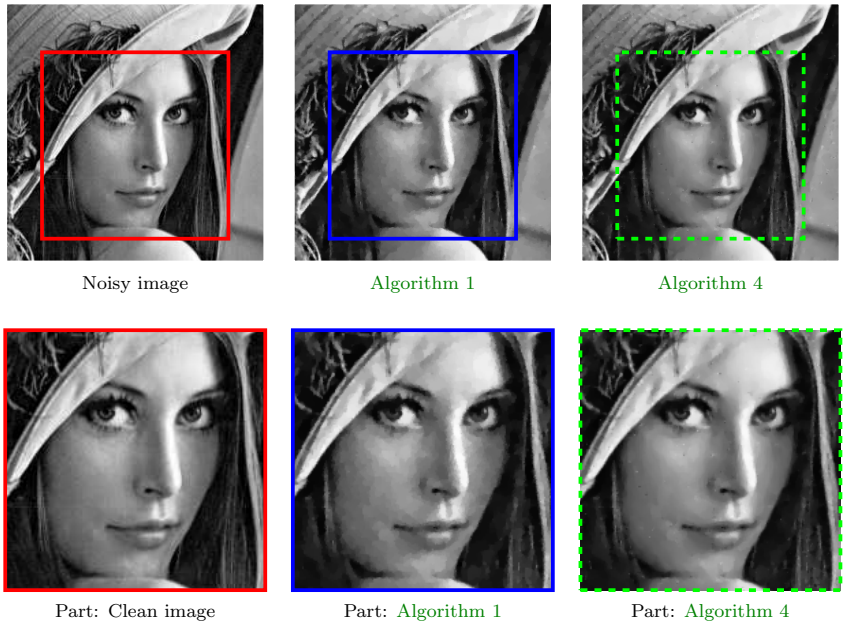

FIG. 6. *Comparing the effect of the proposed Richardson-like iterative Algorithm 4 for the second step of the TV-Stokes model and the Osher-like Algorithm 1 for the ROF model on the noisy Lena image. The initial noise level is* 15.71.

The results shown in FIG. 6 demonstrate that the proposed Algorithm 4 is effective in preserving smooth surfaces like Lena face while the Osher-like iteration, i.e., Algorithm 1, is defective, showing a patch like surface. The PSNR values of the restored image are 30.23 and 30.79 corresponding to Algorithm 1 and Algorithm 4, respectively. The according noise level are 8.50 and 7.97. For a better understanding the two restored images, we also plot the contours of these images, cf. FIG. 7. The contours obtained from Algorithm 4 are of higher parallelity compared to the ones obtained from Algorithm 1. This reflects also that Algorithm 4 can generate smoother surfaces, e.g., Lena face, than Algorithm 1.

The next experiments are taken on a image of a fingerprint, cf. FIG. 8 and FIG. 9, with PSNR 22.65 at noise level 22.11. The results demonstrate the effectiveness of the proposed Algorithm 4 in preserving sharp edges such as fingerprint textures and in restoring the smoothed structures like the surroundings in this image. The contours show a better connectivity of the texture for FIG. 9 resulting longer structures, cf. FIG. 9.

The last experiments are applied on the Barbara image, cf. FIG. 10 and FIG. 11. The noise level of the initial image is 20.06 and initial PSNR is 28.14. The results from the proposed Algorithm 4 show much smoother face and arm compared to the results from Algorithm 1. The results of Algorithm 4 also show a better restoration in preserving pinstripes structures. The final restored images are at noise level 14.15 and 12.91 for Algorithm 1 and Algorithm 4, respectively. The according PSNR are

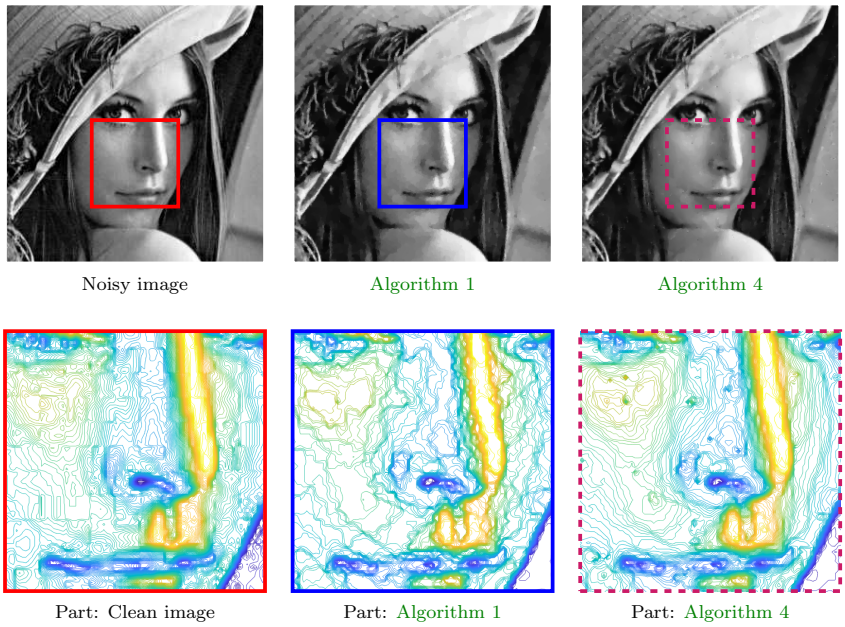

FIG. 7. *Contours show, comparing the effect of Algorithm 4 and Algorithm 1 on the noisy Lena image. The initial noise level is* 15.71.

31.17 and 31.97.

4. Conclusions. In this paper, we have proposed the Richardson-like iterative algorithms applied to the first step, the second step, and both steps of the TV-Stokes model, respectively. We have proven the well-definedness and the convergence of each algorithm. The numerical experiments show a visible improvement compared to the Osher-like iteration on the ROF model in both edge preserving and surface smoothing.

REFERENCES

[1] G. AUBERT AND P. KORNPROBST, *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*, Springer Science + Business Media, LLC, second edi ed., 2006, https://doi.org/10.2307/3615195.

[2] I. BAYRAM AND M. E. KAMASAK, *A Directional Total Variation*, European Signal Processing Conference, 19 (2012), pp. 265–269, https://doi.org/10.1109/LSP.2012.2220349.

[3] K. BREDIES, K. KUNISCH, AND T. POCK, *Total Generalized Variation*, SIAM Journal on Imaging Sciences, (2010), https://doi.org/10.1137/090769521.

[4] A. CHAMBOLLE, V. CASELLES, M. NOVAGA, D. CREMERS, AND T. P. AN, *An Introduction to Total Variation for Image Analysis*, Theoretical Foundations and Numerical Methods for Sparse Recovery, De Gruyter, Radon Series Comp. Appl. Math., 9 (2009), pp. 263–340.

[5] A. CHAMBOLLE AND P.-L. LIONS, *Image Recovery via Total Variation Minimization and Related Problems*, Numerische Mathematik, 76 (1997), pp. 167–188, https://doi.org/10.1007/s002110050258, http://link.springer.com/10.1007/s002110050258.

[6] R. CHAN, T. F. CHAN, AND A. YIP, *Numerical Methods and Applications in Total Variation Image Restoration*, in Handbook of Mathematical Methods in Imaging, Springer, 2011, ch. 24, pp. 1059–1094, https://doi.org/10.1007/978-0-387-92920-0_24.

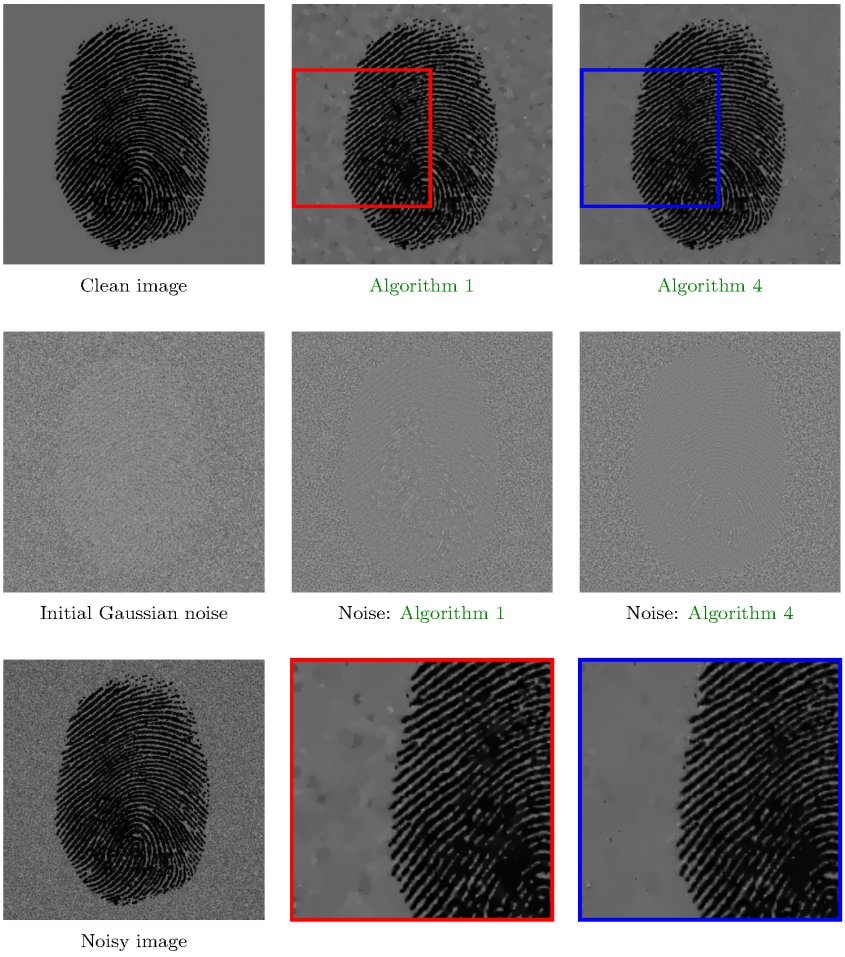

FIG. 8. *Comparing the effect of iterative Algorithm 4 and Osher-like Algorithm 1 for ROF on the noisy fingerprint image. The initial noise level is* 22.11.

[7] T. F. CHAN, G. H. GOLUB, AND P. MULET, *A Nonlinear Primal-Dual Method for Total Variation-based Image Restoration*, SIAM Journal on Scientific Computing, 20 (1999), pp. 1964–1977, https://doi.org/10.1137/S1064827596299767, http://epubs.siam.org/doi/10.1137/S1064827596299767.

[8] T. F. CHAN AND J. SHEN, *Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods*, SIAM, 2005, https://doi.org/10.1137/1.9780898717877.

[9] D. Q. CHEN, L. Z. CHENG, AND F. SU, *A New TV-Stokes Model with Augmented Lagrangian Method for Image Denoising and Deconvolution*, Journal of Scientific Computing, 51 (2012), pp. 505–526, https://doi.org/10.1007/s10915-011-9519-x.

[10] P. P. EGGERMONT AND V. N. LARICCIA, *Maximum Penalized Likelihood Estimation. Volume II: Regression*, Springer, 2009, https://doi.org/10.1007/B12285.

[11] C. A. ELO, *Image Denoising Algorithms Based on the Dual Formulation of Total Variation*, master thesis, University of Bergen, 2009, http://hdl.handle.net/1956/3367.

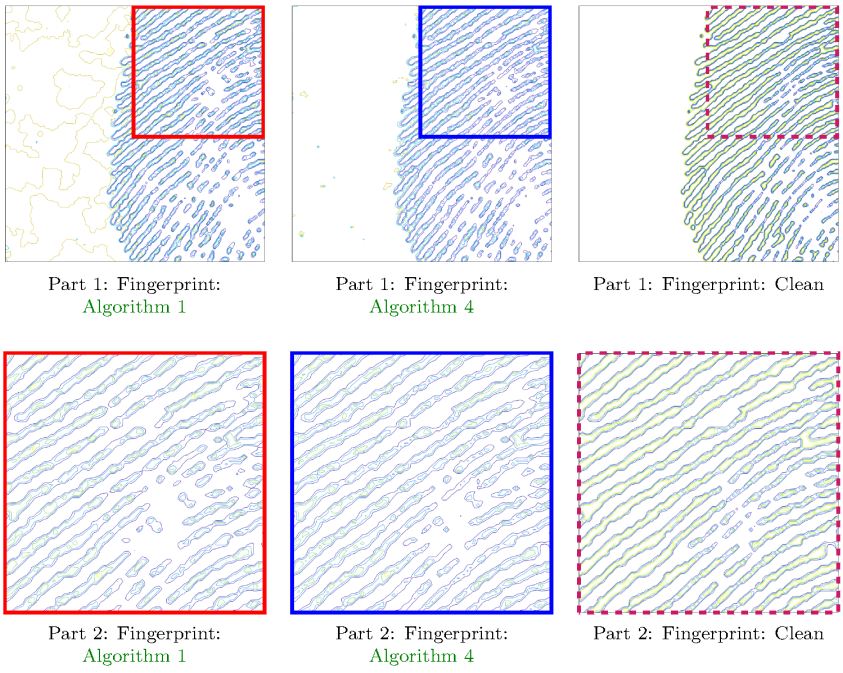

FIG. 9. *Contours show, comparing the effect of iterative Algorithm 4 and Osher-like Algorithm 1 for the ROF model on the noisy fingerprint image. The initial noise level is* 22.11.

[12] C. A. ELO, A. MALYSHEV, AND T. RAHMAN, *A Dual Formulation of the TV-Stokes Algorithm for Image Denoising*, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 5567 LNCS (2009), pp. 307–318, https://doi.org/10.1007/978-3-642-02256-2_26.

[13] J. HAHN, C. WU, AND X.-C. TAI, *Augmented Lagrangian Method for Generalized TV-Stokes Model*, Journal of Scientific Computing, 50 (2012), pp. 235–264, https://doi.org/10.1007/s10915-011-9482-6, http://link.springer.com/10.1007/s10915-011-9482-6.

[14] B. HE AND X. YUAN, *Convergence Analysis of Primal-Dual Algorithms for Total Variation Image Restoration*, Optimization Online, (2010), pp. 1–22, www.optimization-online.org/DB_FILE/2010/11/2790.pdf.

[15] W. G. LITVINOV, T. RAHMAN, AND X.-C. TAI, *A Modified TV-Stokes Model for Image Processing*, SIAM Journal on Scientific Computing, 33 (2011), pp. 1574–1597, https://doi.org/10.1137/080727506, http://epubs.siam.org/doi/10.1137/080727506.

[16] L. MARCINKOWSKI AND T. RAHMAN, *An iterative regularization algorithm for the TV-Stokes in image processing*, in Parallel processing and applied mathematics. Part II, vol. 9574 of Lecture Notes in Comput. Sci., Springer, Cham, 2016, pp. 381–390, https://doi.org/10.1007/978-3-319-32152-3_36.

[17] Y. MEYER, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, vol. 22, American Mathematical Society, 2001.

[18] S. OSHER, M. BURGER, D. GOLDFARB, J. XU, AND W. YIN, *An Iterative Regularization Method for Total Variation-Based Image Restoration*, Multiscale Modeling & Simulation, 4 (2005), pp. 460–489, https://doi.org/10.1137/040605412, http://epubs.siam.org/doi/10.1137/040605412.

[19] S. OSHER, A. SOLÉ, AND L. VESE, *Image Decomposition and Restoration Using Total Variation Minimization and the H1 Norms*, SIAM Multiscale Modeling & Simulation, 1 (2003), pp. 349–370, https://doi.org/10.1137/S1540345902416247, http://epubs.siam.org/

FIG. 10. *Showing the effect of using Algorithm 4 and Algorithm 1 on Barbara image.*

doi/pdf/10.1137/S1540345902416247.

[20] T. RAHMAN, X.-C. TAI, AND S. OSHER, *A TV-Stokes Denoising Algorithm*, in Scale Space and Variational Methods in Computer Vision, Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2007, pp. 473–483, https://doi.org/10.1007/978-3-540-72823-8_41.

[21] L. I. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear Total Variation Noise Removal Algorithm*, Physica D: Nonlinear Phenomena, 60 (1992), pp. 259–268, https://doi.org/10.1016/0167-2789(92)90242-F, http://www.csee.wvu.edu/~xinl/courses/ee565/total_variation.pdf.

[22] O. SCHERZER, M. GRASMAIR, H. GROSSAUER, M. HALTMEIER, AND F. LENZEN, *Variational Methods in Imaging*, vol. 167, Springer Science + Business Media, LLC, 2009, https://doi.org/10.1007/978-0-387-69277-7.

[23] M. SION, *On General Minimax Theorems*, Pacific Journal of Mathematics, 18 (1957), pp. 171–176, https://doi.org/1103040253.

[24] X.-C. TAI AND C. WU, *Augmented Lagrangian Method, Dual Methods and Split Bregman Iteration for ROF Model*, SIAM Journal on Imaging Sciences, 3 (2009), pp. 300–339, https://doi.org/10.1137/090767558.

[25] L. A. VESE AND S. J. OSHER, *Modeling Textures with Total Variation Minimization and Oscillating Patterns in Image Processing*, Journal of Scientific Computing, (2003), https://doi.org/10.1023/A:1025384832106.

[26] C. WU AND X.-C. TAI, *Augmented Lagrangian Method, Dual Methods,and Split Bregman Iteration for ROF,Vectorial TV, and High Order Models*, Methods, 3 (2010), pp. 300–339, https://doi.org/10.1137/090767558.

[27] C. WU, J. ZHANG, Y. DUAN, AND X.-C. TAI, *Augmented Lagrangian Method for Total Variation Based Image Restoration and Segmentation Over Triangulated Surfaces*, Journal of Scientific Computing, 50 (2012), pp. 145–166, https://doi.org/10.1007/s10915-011-9477-3, http://link.springer.com/10.1007/s10915-011-9477-3.

[28] C. WU, J. ZHANG, AND X.-C. TAI, *Augmented Lagrangian Method for Total Variation Restoration with Non-quadratic Fidelity*, Inverse Problems and Imaging, 5 (2011), pp. 237–261, https://doi.org/10.3934/ipi.2011.5.237, http://www.aimsciences.org/

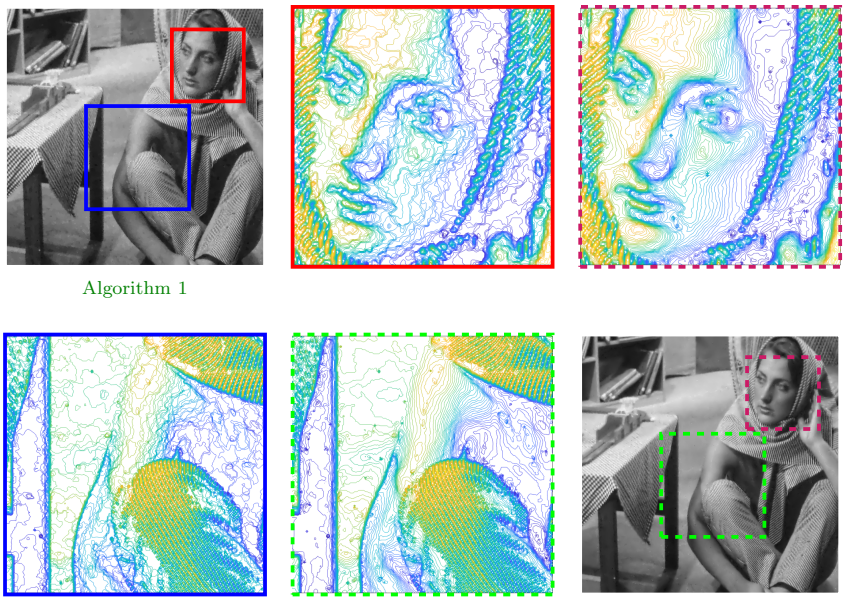

FIG. 11. *The contours showing the effect of using Algorithm 4 and Algorithm 1 on Barbara image.*

journals/displayArticles.jsp?paperID=5934.

[29] J. YANG, Y. ZHANG, AND W. YIN, *An Efficient TVL1 Algorithm for Deblurring multichannel Images Corrupted by Impulsive Noise*, SIAM Journal on Scientific Computing, 31 (2009), pp. 2842–2865, https://doi.org/10.1137/080732894.

[30] M. ZHU AND T. CHAN, *An Efficient Primal-Dual Hybrid Gradient Algorithm for Total Variation Image Restoration*, UCLA CAM Report, 08 (2008), pp. 1–29, http://www.coe.utah.edu/~cs7640/readings/cam08-34.pdf.

Paper E

Alternating Minimization for a Single Step TV-Stokes Model for Image Denoising

Bin Wu, Xue-Cheng Tai and Talal Rahman *Submitted. arXiv:2009.11973*

ALTERNATING MINIMIZATION FOR A SINGLE STEP TV-STOKES MODEL FOR IMAGE DENOISING

BIN WU*, XUE-CHENG TAI†, AND TALAL RAHMAN*

Abstract. The paper presents a fully coupled TV-Stokes model, and propose an algorithm based on an alternating minimization of the objective functional, whose first iteration is exactly the same as the modified TV-Stokes model proposed earlier. The model is also a generalization of the second order Total Generalized Variation model. A convergence analysis is given.

Key words. Total variation, Denoising, Restoration

AMS subject classifications. 68Q25, 68R10, 68U05

1. Introduction. The ROF model or the Rudin-Osher-Fatemi model, cf. [12], is one of the most well-known and successful models in variational image processing. It handles various problems, as for instance, denoising, inpainting, based on a total variation minimization. The model, however, suffers from a so-called staircase effect, which results in block artifacts. Among the models that alleviate the effect, are the TV-Stokes, cf. [11, 4], and its modified variant, cf. [6], henceforth called the LRT model. The modified TV-Stokes model consists of two steps where the object image is reconstructed based on the vector field obtained from its first step. The vector field is achieved by a second-order minimization subject to a divergence-free condition, also known as the Stokes constraint, which implies that the vector field is a tangent field. For simplicity, we consider the gradient field and the condition becomes a curl-free condition, in other words condition for a potential function.

The modified TV-Stokes is defined as follows. Given an observed image f, define its gradient as $\mathbf{n}$, and an orthogonal projector Π as $\Pi(\mathbf{n}) = \nabla\triangle^{\dagger}\nabla\cdot(\mathbf{n})$, as in [4]. The two-step modified TV-Stokes model thus has an equivalent form, cf. [4, 11], as follows.

1. Find the smoothed gradient field with curl-free constraint:

$$\min_{\substack{\mathbf{n}\\ \Pi\mathbf{n}=\mathbf{n}}} |\nabla\mathbf{n}|_{(\Omega)} + \frac{1}{2\delta}\|\mathbf{n}-\nabla f\|^2_{(\Omega)}.$$

2. Restore the object image using the gradient field obtained from the first step:

$$\min_{u} |\nabla u - \mathbf{n}|_{(\Omega)} + \frac{1}{2\theta}\|u-f\|^2_{(\Omega)},$$

where Ω is a bounded open domain in $\mathbb{R}^2$. $|\cdot|_{(\Omega)}$ and $\|\cdot\|_{(\Omega)}$ are the usual L^1 and L^2 norms, respectively. The current paper is about combining the two steps of the modified TV-Stokes into one step, coupling the two variables, that is the image and its normal field in one minimization step. It is then easier to develop an iterative regularization algorithm for the combined model.

We organize the paper as follows: we propose the combined model and an iterative algorithm for the model in section 2, give a convergence analysis in section 3, and finally end the paper with some remarks at the end.

*Department of Computer Science, Electrical Engineering and Mathematical Sciences, Western Norway University of Applied Sciences, Inndalsveien 28, 5063 Bergen, Norway (bin.wu@hvl.no, talal.rahman@hvl.no).

†Department of Mathematics, University of Bergen, Allégaten 41, 5007 Bergen, Norway (xuecheng.tai@uib.no).

2. Proposed Model and Iterative Algorithm. We propose our one-step model as follows.

$$\min_{\substack{\mathbf{n},u \\ \Pi\mathbf{n}=\mathbf{n}}} \alpha|\nabla\mathbf{n}|_{(\Omega)} + \beta|\nabla u - \mathbf{n}|_{(\Omega)} + \frac{\eta_1}{2}\|\mathbf{n} - \nabla f\|^2_{(\Omega)} + \frac{\eta_2}{2}\|u - f\|^2_{(\Omega)}. \tag{1}$$

However, we solve the model by iterations. We split the minimization problem of the model into two sub-problems by freezing one of the two variables which we then solve alternatingly in each iteration. The sub-problems are described in the following sub section.

2.1. Sub-problems. In the section we introduce the two sub-problems, which we solve alternatingly.

1: Find the smoothed gradient field satisfying the potential function constraint, by solving the following minimization problem:

$$\min_{\substack{\mathbf{n} \\ \Pi\mathbf{n}=\mathbf{n}}} \alpha|\nabla\mathbf{n}|_{(\Omega)} + \frac{\eta_1}{2}\|\mathbf{n} - \nabla f\|^2_{(\Omega)} + \beta|\mathbf{n} - \nabla u|_{(\Omega)}. \tag{2}$$

Here u is fixed, which is the solution of the second sub-problem. Initially u is set equal to f.

2: Restore the given image by solving the following minimization problem:

$$\min_{u} \beta|\nabla u - \mathbf{n}|_{(\Omega)} + \frac{\eta_2}{2}\|u - f\|^2_{(\Omega)}. \tag{3}$$

Here $\mathbf{n}$ is fixed, and comes from the first step in the current iteration.

For the first sub-problem, we turn the constrained minimization into an unconstrained minimization. In the following Lemma 1, we prove the equivalency between the two formulations.

Lemma 1 (Equivalency). *For all $\mathbf{n} \in \mathbb{R}^2$, define Π as a projector such that $\Pi(\mathbf{n}) = \nabla\triangle^{\dagger}\nabla\cdot\mathbf{n}$. The constrained problem (2) is equivalent to the following unconstrained problem*

$$\min_{\mathbf{n}} \alpha|\nabla\Pi\mathbf{n}|_{(\Omega)} + \frac{\eta_1}{2}\|\mathbf{n} - \nabla f\|^2_{(\Omega)} + \beta|\Pi\mathbf{n} - \nabla u|_{(\Omega)}. \tag{4}$$

Proof. Let $\mathbf{p}$ and $\mathbf{q}$ be the dual variables such that $\mathbf{p} \in C^1_c(\Omega, \mathbb{R}^4)$ and $\mathbf{q} \in C^1_c(\Omega, \mathbb{R}^2)$, with $\|\mathbf{p}\|_\infty$ and $\|\mathbf{q}\|_\infty \leq 1$, the minimization problem (2) is thus

$$\min_{\mathbf{n}} \max_{\substack{\mathbf{p},\mathbf{q} \\ |\cdot|\leq 1}} \alpha\langle\mathbf{n}, \nabla\cdot\mathbf{p}\rangle + \frac{\eta_1}{2}(\mathbf{n} - \nabla f)^2 + \beta\langle\mathbf{n} - \nabla u, \mathbf{q}\rangle + \lambda\cdot(\Pi\mathbf{n} - \mathbf{n}), \tag{5}$$

where λ is the Lagrange multiplier. By the min-max theorem, cf. [13], we can reformulate the min-max into max-min, above. The corresponding Euler-Lagrange equations become

$$\alpha\nabla\cdot\mathbf{p} + \eta_1(\mathbf{n} - \nabla f) + \beta\mathbf{q} + (\Pi - I)\lambda = 0, \tag{6}$$

and

$$\Pi\mathbf{n} - \mathbf{n} = 0.$$

Applying the Π projection on both sides of (6), we obtain the following relation.

$$\eta_1(\mathbf{n} - \nabla f) = -\alpha\Pi\nabla\cdot\mathbf{p} - \beta\Pi\mathbf{q}.$$

Substituting back in (6), we obtain

$$(\Pi - I)\lambda = -\alpha(I - \Pi)\nabla \cdot \mathbf{p} - \beta(I - \Pi)\mathbf{q}.$$

The problem (5) is accordingly

$$\min_{\mathbf{n}} \max_{\substack{\mathbf{p},\mathbf{q} \\ |\cdot|\leq 1}} \alpha\langle \mathbf{n}, \Pi\nabla \cdot \mathbf{p}\rangle + \frac{\eta_1}{2}(\mathbf{n} - \nabla f)^2 + \beta\langle \mathbf{n}, \Pi\mathbf{q}\rangle - \beta\langle \nabla u, \mathbf{q}\rangle,$$

which is equivalent to the following primal problem:

$$\min_{\mathbf{n}} \alpha|\nabla\Pi\mathbf{n}|_{(\Omega)} + \frac{\eta_1}{2}\|\mathbf{n} - \nabla f\|^2_{(\Omega)} + \beta|\Pi\mathbf{n} - \nabla u|_{(\Omega)}.$$

2.2. Algorithm. We solve the optimization problem by the alternating minimization method, cf. [8, 10], where the method is also known as block-nonlinear GaussSeidel method or the block coordinate descent method, cf. [9] and cf. [1] for a review. The algorithm is given as follows (Algorithm 1).

Algorithm 1 Alternating Minimization TV-Stokes

initialize $k = 0$, $u_0 = f$
repeat
  $k = k + 1$
  $\mathbf{n}_k = \arg\min_{\mathbf{n}}\{\alpha|\nabla\Pi\mathbf{n}|(\Omega) + \frac{\eta_1}{2}\|\mathbf{n} - \nabla f\|^2_{(\Omega)} + \beta|\Pi\mathbf{n} - \nabla u_{k-1}|(\Omega)\}$
  $u_k = \arg\min_u\{\beta|\nabla u - \mathbf{n}_k|(\Omega) + \frac{\eta_2}{2}\|u - f\|^2_{(\Omega)}\}$
until *satisfied*;
return u.

2.3. Solving Sub-problem 1 for n. We adopt Chambolle's semi-implicit algorithm [3] to solve the proposed sub-prob-lems. The dual problem for the sub-problem (4) is as follows:

$$\min_{\mathbf{n}} \max_{\substack{\mathbf{p},\mathbf{q} \\ |\cdot|\leq 1}} \alpha\langle \mathbf{n}, \Pi\nabla \cdot \mathbf{p}\rangle + \frac{\eta_1}{2}(\mathbf{n} - \nabla f)^2 + \beta\langle \mathbf{n}, \Pi\mathbf{q}\rangle - \beta\langle \nabla u, \mathbf{q}\rangle. \tag{7}$$

We first consider the minimization by the min-max theorem, cf. [13], which gives us the optimal condition on $\mathbf{n}$ as follows.

$$\mathbf{n} = \nabla f - \frac{\alpha}{\eta_1}\Pi\nabla \cdot \mathbf{p} - \frac{\beta}{\eta_1}\Pi\mathbf{q}. \tag{8}$$

Accordingly the maximization problem is

$$\max_{\substack{\mathbf{p},\mathbf{q} \\ |\cdot|\leq 1}} \langle \nabla f, \alpha\Pi\nabla \cdot \mathbf{p} + \beta\Pi\mathbf{q}\rangle - \frac{1}{2\eta_1}(\alpha\Pi\nabla \cdot \mathbf{p} + \beta\Pi\mathbf{q})^2 - \beta\langle \nabla u, \mathbf{q}\rangle. \tag{9}$$

The line searching directions $\mathbf{r}$ for $\mathbf{p}$ and $\mathbf{q}$ in Chambolle's semi-implicit algorithm are respectively

$$\mathbf{r}_{\mathbf{p}} = \nabla(\frac{\alpha}{\eta_1}\Pi\nabla \cdot \mathbf{p} + \frac{\beta}{\eta_1}\Pi\mathbf{q} - \nabla f) \tag{10}$$

and

$$\mathbf{r}_{\mathbf{q}} = \nabla f - \nabla u - \frac{\alpha}{\eta_1}\Pi\nabla \cdot \mathbf{p} - \frac{\beta}{\eta_1}\Pi\mathbf{q}. \tag{11}$$

2.4. Solving Sub-problem 2 for u. We use the same approach as for the sub-problem 1, but this time for the scalar function u. Accordingly, the dual problem of the sub-problem (3) is as follows:

$$\min_{u} \max_{\mathbf{s}} \beta\langle \nabla u - \mathbf{n}, -\mathbf{s}\rangle + \frac{\eta_2}{2}(u-f)^2, \tag{12}$$

where $\mathbf{s} \in C_c^1(\Omega, \mathbb{R}^2)$ and $\|\cdot\|_\infty \leq 1$. The minimization problem results in optimal u as follows:

$$u = f - \frac{\beta}{\eta_2}\nabla \cdot \mathbf{s}, \tag{13}$$

The maximization problem becomes

$$\max_{\mathbf{s}} \beta\langle f, \nabla \cdot \mathbf{s}\rangle + \beta\langle \mathbf{n}, \mathbf{s}\rangle - \frac{\beta^2}{2\eta_2}(\nabla \cdot \mathbf{s})^2. \tag{14}$$

The line search direction $\mathbf{r}$ for $\mathbf{s}$ is

$$\mathbf{r}_{\mathbf{s}} = \nabla(\frac{\beta}{\eta_2}\nabla \cdot \mathbf{s} - f) + \mathbf{n}. \tag{15}$$

3. Convergence Analysis. We prove the convergence following the approach in the paper [1].

3.1. Definitions and Assumptions. Define the vector $\mathbf{x} = (\mathbf{n}, u)$ such that $\mathbf{n} \in \mathbb{R}^2(\Omega) \cap \{\mathbf{n} : \Pi\mathbf{n} = \mathbf{n}\}$ and $u \in \mathbb{R}(\Omega)$. Rewrite the problem (1) as:

$$\min_{\mathbf{x}}\{H(\mathbf{x}) := g(\mathbf{x}) + l(\mathbf{x})\}, \tag{16}$$

where $g(\mathbf{x}) := g(\mathbf{n}, u) = \alpha|\nabla\Pi\mathbf{n}|_{(\Omega)} + \beta|\nabla u - \Pi\mathbf{n}|_{(\Omega)}$ and $l(\mathbf{x}) := l(\mathbf{n}, u) = \frac{\eta_1}{2}\|\mathbf{n} - \nabla f\|^2_{(\Omega)} + \frac{\eta_2}{2}\|u - f\|^2_{(\Omega)}$.

Define functions g_1, g_2, l_1 and l_2 respectively as

$$\begin{aligned} g_1(\mathbf{n}) &:= \alpha|\nabla\Pi\mathbf{n}|_{(\Omega)} + \beta|\Pi\mathbf{n} - \nabla u|_{(\Omega)}, \\ g_2(u) &:= \beta|\nabla u - \Pi\mathbf{n}|_{(\Omega)}, \\ l_1(\mathbf{n}) &:= \frac{\eta_1}{2}\|\mathbf{n} - \nabla f\|^2_{(\Omega)}, \\ l_2(u) &:= \frac{\eta_2}{2}\|u - f\|^2_{(\Omega)}. \end{aligned}$$

We assume that g_1, g_2, l_1 and l_2 satisfy the following properties:

(a) Functions $g_1 : \mathbb{R}^2(\Omega) \to (-\infty, \infty)$ and $g_2 : \mathbb{R}(\Omega) \to (-\infty, \infty)$ are closed and proper convex and subdifferentiable.

(b) Functions l_1 and l_2 are continuously differentiable convex functions over domains $\mathrm{dom}(g_1)$ and $\mathrm{dom}(g_2)$ respectively. The partial derivatives of l with respect to $\mathbf{n}$ and u are denoted as $\nabla_1 l(\mathbf{x})$ and $\nabla_2 l(\mathbf{x})$ respectively.

(c) The gradient of l is Lipschitz continuous with respect to $\mathbf{n}$ over domain $\mathrm{dom}(g_1)$ with a constant $L_1 \in (0, \infty)$ and with respect to u over domain $\mathrm{dom}(g_2)$ with a constant $L_2 \in (0, \infty)$ as follows

$$\|\nabla_1 l(\mathbf{n} + \mathbf{d}_1, u) - \nabla_1 l(\mathbf{n}, u)\|_{(\Omega)} \leq L_1\|\mathbf{d}_1\|_{(\Omega)},$$

and

$$\|\nabla_2 l(\mathbf{n}, u + d_2) - \nabla_2 l(\mathbf{n}, u)\|_{(\Omega)} \leq L_2 \|d_2\|_{(\Omega)},$$

where $\mathbf{d}_1 \in \mathbb{R}^2(\Omega)$ and $u \in \mathbb{R}(\Omega)$ such that $\mathbf{n} + \mathbf{d}_1 \in \mathrm{dom}(g_1)$ and $u + d_2 \in \mathrm{dom}(g_2)$.

(d) The optimal solution set of problem (16), denoted by X^*, is nonempty, and the corresponding energy value of $H(\mathbf{x})$ is H^*. In addition, the solution sets for sub-problems are also nonempty.

Using the notations above, we reformulate Algorithm 1 as follows.

Algorithm 2 Alternating Minimization Method (Reformulated)

initialize
  $k = 0$
  $\mathbf{n}_0 \in \mathrm{dom}\ g_1$, $u_0 = f \in \mathrm{dom}\ g_2$, such that $u_0 \in \mathrm{argmin}_{u \in BV(\Omega)}\ g_2(\mathbf{n}_0, u) + l_2(\mathbf{n}_0, u)$
repeat
  $k = k + 1$
  $\mathbf{n}_k = \arg\min_{\mathbf{n} \in BV^2(\Omega)}\ g_1(\mathbf{n}, u_{k-1}) + l_1(\mathbf{n})$
  $u_k = \arg\min_{u \in BV(\Omega)}\ g_2(\mathbf{n}_k, u) + l_2(u)$
until *satisfied*;
return u.

The k-th iteration reads as $\mathbf{x}_k = (\mathbf{n}_k, u_k)$, and the $(k+1)$-th iteration as $\mathbf{x}_{k+1} = (\mathbf{n}_{k+1}, u_{k+1})$. The intermediate iteration is defined as half iteration, i.e., $\mathbf{x}_{k+1/2} = (\mathbf{n}_{k+1}, u_k)$. The sequence generated by the optimal problems has the relation:

$$H(\mathbf{x}_0) \geq H(\mathbf{x}_{1/2}) \geq H(\mathbf{x}_1) \geq H(\mathbf{x}_{3/2}) \geq \dots$$

3.2. Preliminaries. Three concepts are associated with the convergence analysis of the proposed algorithm: the proximal mapping, the gradient mapping, and the block descent lemma.

Define proximal operator $\mathrm{prox}_h(\cdot)$ as

$$\mathrm{prox}_h(\mathbf{z}) = \arg\min_{y} h(\mathbf{y}) + \frac{1}{2}\|\mathbf{y} - \mathbf{z}\|^2,$$

where $h : \mathbb{R}^n \to (-\infty, \infty]$ is a closed, proper convex function. A lemma on proximal mapping, cf. [1], is useful in the our proof of convergence.

LEMMA 2. *Let $h : \mathbb{R}^n \to (-\infty, \infty]$ be a closed, proper and convex function, $M > 0$, and $\mathbf{v} = prox_{\frac{1}{M}h}(\mathbf{z})$. Thus if $\mathbf{y} \in dom(h)$ then*

$$h(\mathbf{y}) \geq h(\mathbf{v}) + M\langle \mathbf{z} - \mathbf{v}, \mathbf{y} - \mathbf{v}\rangle.$$

Proof. The optimal condition for the convex functional $h(\mathbf{y}) + \frac{M}{2}\|\mathbf{y} - \mathbf{z}\|^2$ is $\partial h(\mathbf{v}) + M(\mathbf{v} - \mathbf{z}) \ni 0$, where $\partial h(\mathbf{v})$ is the subgradient at $\mathbf{v}$. By the definition of subgradient, it follows that

$$h(\mathbf{y}) \geq h(\mathbf{v}) + M\langle \mathbf{z} - \mathbf{v}, \mathbf{y} - \mathbf{v}\rangle,$$

and hence the proof.

Now define prox-grad mapping $T_M(\cdot)$ as $T_M(\mathbf{y}) = \text{prox}_{\frac{1}{M}h}(\mathbf{y} - \frac{1}{M}\nabla f(\mathbf{y}))$ for $M > 0$ associated with $h(\mathbf{y}) + f(\mathbf{y})$ where $f(\mathbf{y}) = \frac{M}{2}\|\mathbf{y} - \mathbf{z}\|^2$, cf. [7]. The corresponding gradient mapping is defined as $G_M(\mathbf{y}) = M(\mathbf{y} - T_M(\mathbf{y})) = M(\mathbf{y} - \text{prox}_{\frac{1}{M}h}(\mathbf{y} - \frac{1}{M}\nabla f(\mathbf{y})))$. Note that if $G_M(\mathbf{y}) = 0$ for some $M > 0$, $\mathbf{y}$ is an optimal solution. For the problem (16), the gradient mapping is accordingly

$$G_M(\mathbf{x}) = M(\mathbf{x} - T_M(\mathbf{x})) = M(\mathbf{x} - \text{prox}_{\frac{1}{M}g}(\mathbf{x} - \frac{1}{M}\nabla l(\mathbf{x}))),$$

where $M \in \{\eta_1, \eta_2\}$, is the parameter corresponding to $\mathbf{n}$ and u. The partial gradient mappings are therefore

$$\begin{aligned} G_M^1(\mathbf{x}) &= \eta_1(\mathbf{n} - T_{\eta_1}(\mathbf{n})) = \eta_1(\mathbf{n} - \text{prox}_{\frac{1}{\eta_1}g_1}(\mathbf{n} - \frac{1}{\eta_1}\nabla l_1(\mathbf{n}))), \\ G_M^2(\mathbf{x}) &= \eta_2(u - T_{\eta_2}(u) = \eta_2(u - \text{prox}_{\frac{1}{\eta_1}g_2}(u - \frac{1}{\eta_2}\nabla l_2(u))), \end{aligned}$$

where

$$G_M(\mathbf{x}) = (G_M^1(\mathbf{x}), G_M^2(\mathbf{x})).$$

LEMMA 3 (block descent lemma). *For all $M_1 > L_1$ and $M_2 > L_2$, and the assumption (c), the following relations hold.*

$$\begin{aligned} l(\mathbf{n} + \mathbf{d_1}, u) &\le l(\mathbf{n}, u) + \langle \nabla_1 l(\mathbf{n}, u), \mathbf{d_1} \rangle + \frac{M_1}{2}\|\mathbf{d_1}\|^2, \\ l(\mathbf{n}, u + d_2) &\le l(\mathbf{n}, u) + \langle \nabla_2 l(\mathbf{n}, u), d_2 \rangle + \frac{M_2}{2}\|d_2\|^2. \end{aligned}$$

Proof. By the Taylor expansion, we have

$$\begin{aligned} l(\mathbf{n} + \mathbf{d_1}, u) &= l(\mathbf{n}, u) + \langle \nabla_1 l(\mathbf{n}, u), \mathbf{d_1} \rangle + \mathbf{d_1}^\top \frac{\nabla_1 \nabla_1 l(\tilde{\mathbf{n}}, u)}{2} \mathbf{d_1}, \\ l(\mathbf{n}, u + d_2) &= l(\mathbf{n}, u) + \langle \nabla_2 l(\mathbf{n}, u), d_2 \rangle + \frac{\nabla_2 \nabla_2 l(\mathbf{n}, \tilde{u})}{2}\|d_2\|^2, \end{aligned}$$

where $\tilde{\mathbf{n}}$ is some vector in the box $[\mathbf{n}, \mathbf{n} + \mathbf{d_1}]$, and $\tilde{u}$ is some scalar value in $[u, u + d_2]$. By the Lipschitz assumption, with $M_1 > L_1$ and $M_2 > L_2$, the following relations thus hold:

$$\begin{aligned} l(\mathbf{n} + \mathbf{d_1}, u) &\le l(\mathbf{n}, u) + \langle \nabla_1 l(\mathbf{n}, u), \mathbf{d_1} \rangle + \frac{M_1}{2}\|\mathbf{d_1}\|^2, \\ l(\mathbf{n}, u + d_2) &\le l(\mathbf{n}, u) + \langle \nabla_2 l(\mathbf{n}, u), d_2 \rangle + \frac{M_2}{2}\|d_2\|^2. \end{aligned}$$

The proof follows.

LEMMA 4 (sufficient decrease, cf. [2, 1]).
Let $s \in C^{L,1}(\Omega, \mathbb{R}^p \to \mathbb{R})$, $h : \mathbb{R}^p \to (-\infty, \infty]$ be a closed, proper, and convex sub-differentiable function, and $S(\cdot) = s(\cdot) + h(\cdot)$, then

$$S(x) - S(prox_{\frac{1}{L}h}(x - \frac{\nabla s(x)}{L})) \ge \frac{1}{2L}\|L(x - prox_{\frac{1}{L}h}(x - \frac{\nabla s(x)}{L}))\|^2.$$

Proof. Since h is convex and sub-differentiable, we have

$$h(x) \geq h(\text{prox}_{\frac{h}{L}}(x - \frac{\nabla s(x)}{L})) + (x - \text{prox}_{\frac{h}{L}}(x - \frac{\nabla s(x)}{L}))\partial h(\text{prox}_{\frac{h}{L}}(x - \frac{\nabla s(x)}{L}))).$$

The optimal condition for the minimization of $\text{prox}_{\frac{1}{L}h}(x - \frac{1}{L}\nabla s(x))$ gives

$$\partial h(\text{prox}_{\frac{1}{L}h}(x - \frac{1}{L}\nabla s(x))) \ni L(x - \frac{1}{L}\nabla s(x) - \text{prox}_{\frac{1}{L}h}(x - \frac{1}{L}\nabla s(x))).$$

We therefore have the following inequality

$$\begin{aligned} h(x) \geq h(\text{prox}_{\frac{h}{L}}(x - \frac{\nabla s(x)}{L})) + \frac{1}{L}\|L(x - \text{prox}_{\frac{h}{L}}(x - \frac{\nabla s(x)}{L}))\|^2 \\ -(x - \text{prox}_{\frac{h}{L}}(x - \frac{\nabla s(x)}{L}))\nabla s(x). \end{aligned}$$

Further, we obtain

$$\begin{aligned} S(x) - S(\text{prox}_{\frac{1}{L}h}(x - \frac{1}{L}\nabla s(x))) \geq s(x) - s(\text{prox}_{\frac{1}{L}h}(x - \frac{1}{L}\nabla s(x))) \\ +\frac{1}{L}\|L(x - \text{prox}_{\frac{1}{L}h}(x - \frac{1}{L}\nabla s(x)))\|^2 \\ -(x - \text{prox}_{\frac{1}{L}h}(x - \frac{1}{L}\nabla s(x)))\nabla s(x). \end{aligned}$$

By the second order Taylor expansion and the Lipschitz continuous condition, cf. Lemma 3,

$$\begin{aligned} s(\text{prox}_{\frac{1}{L}h}(x - \frac{1}{L}\nabla s(x))) \leq s(x) + (\text{prox}_{\frac{1}{L}h}(x - \frac{1}{L}\nabla s(x)) - x)\nabla s(x) \\ +\frac{1}{2L}\|L(x - \text{prox}_{\frac{1}{L}h}(x - \frac{1}{L}\nabla s(x)))\|^2. \end{aligned}$$

This leads to

$$S(x) - S(\text{prox}_{\frac{1}{L}h}(x - \frac{1}{L}\nabla s(x))) \geq \frac{1}{2L}\|L(x - \text{prox}_{\frac{1}{L}h}(x - \frac{1}{L}\nabla s(x)))\|^2.$$

The proof thus follows.

Applying this to our problem, we can conclude that

$$H(\mathbf{n}, u) - H(T_{\eta_1}(\mathbf{n}), u) \geq \frac{1}{2\eta_1}\|G^1_{\eta_1}(\mathbf{n}, u)\|^2, \tag{17}$$

and

$$H(\mathbf{n}, u) - H(\mathbf{n}, T_{\eta_2}(u)) \geq \frac{1}{2\eta_2}\|G^2_{\eta_2}(\mathbf{n}, u)\|^2. \tag{18}$$

3.3. Convergence Rate.

Lemma 5. *Let* $\{\mathbf{x}_k\}$ *be the sequence generated by Algorithm 2. Then* $\forall k \geq 0$

$$H(\mathbf{x}_{k+\frac{1}{2}}) - H(\mathbf{x}^*) \leq \|G^1_{\eta_1}(\mathbf{x}_k)\| \cdot \|\mathbf{x}_k - \mathbf{x}^*\|,$$

and

$$H(\mathbf{x}_{k+1}) - H(\mathbf{x}^*) \leq \|G^2_{\eta_2}(\mathbf{x}_{k+\frac{1}{2}})\| \cdot \|\mathbf{x}_{k+\frac{1}{2}} - \mathbf{x}^*\|.$$

Proof. Since $\mathbf{x}_{k+\frac{1}{2}} = (\mathbf{n}_{k+1}, u_k)$ minimizes the energy functional $H(\mathbf{n}, u_k)$, we have

$$H(\mathbf{x}_{k+\frac{1}{2}}) - H(\mathbf{x}^*) \leq H(T_{\eta_1}(\mathbf{x}_k)) - H(\mathbf{x}^*). \tag{19}$$

Meanwhile, we have $T_{\eta_1}(\mathbf{x}_k) = (T^1_{\eta_1}(\mathbf{x}_k), T^2_{\eta_1}(\mathbf{x}_k)) = (T^1_{\eta_1}(\mathbf{x}_k), u_k - G^2_{\eta_1}(\mathbf{x}_k))$ which is the same as $T_{\eta_1}(\mathbf{x}_k) = (T^1_{\eta_1}(\mathbf{x}_k), u_k)$ by the optimal condition $G^2_M(\mathbf{x}_k) = 0$ for all $M > 0$. Inequality (19) is thus rewritten as

$$H(\mathbf{x}_{k+\frac{1}{2}}) - H(\mathbf{x}^*) \leq H(T^1_{\eta_1}(\mathbf{x}_k), u_k) - H(\mathbf{x}^*).$$

By Lemma 3, we have that

$$\begin{aligned} l(T^1_{\eta_1}(\mathbf{x}_k), u_k) - l(\mathbf{x}^*) &\leq l(\mathbf{x}_k) + \langle \nabla_1 l(\mathbf{x}_k), T^1_{\eta_1}(\mathbf{x}_k) - \mathbf{n}_k \rangle \\ &\quad + \frac{\eta_1}{2} \|T^1_{\eta_1}(\mathbf{x}_k) - \mathbf{n}_k\|^2 - l(\mathbf{x}^*) \\ &= l(\mathbf{x}_k) + \langle \nabla_1 l(\mathbf{x}_k), T_{\eta_1}(\mathbf{x}_k) - \mathbf{x}_k \rangle \\ &\quad + \frac{\eta_1}{2} \|T^1_{\eta_1}(\mathbf{x}_k) - \mathbf{n}_k\|^2 - l(\mathbf{x}^*), \end{aligned}$$

and, by the convexity $l(\mathbf{x}_k) - l(\mathbf{x}^*) \leq \langle \nabla l(\mathbf{x}_k), \mathbf{x}_k - \mathbf{x}^* \rangle$, that

$$l(T^1_{\eta_1}(\mathbf{x}_k)) - l(\mathbf{x}^*) \leq \langle \nabla_1 l(\mathbf{x}_k), T_{\eta_1}(\mathbf{x}_k) - \mathbf{x}^* \rangle + \frac{\eta_1}{2} \|T^1_{\eta_1}(\mathbf{x}_k) - \mathbf{n}_k\|^2.$$

Meanwhile, for h, by the convexity and the optimal condition for $T_{\eta_1}(\mathbf{x}_k)$, that is $\partial_1 h(T_{\eta_1}(\mathbf{x}_k)) \ni \eta_1(\mathbf{x}_k - \frac{1}{\eta_1}\nabla_1 l(\mathbf{x}_k) - T_{\eta_1}(\mathbf{x}_k))$, we have

$$\begin{aligned} h(\mathbf{x}^*) &\geq h(T_{\eta_1}(\mathbf{x}_k)) + \langle \partial_1 h(T_{\eta_1}(\mathbf{x}_k)), \mathbf{x}^* - T_{\eta_1}(\mathbf{x}_k) \rangle, \\ h(\mathbf{x}^*) &\geq h(T_{\eta_1}(\mathbf{x}_k)) + \eta_1 \langle \mathbf{x}_k - \frac{1}{\eta_1}\nabla_1 l(\mathbf{x}_k) - T_{\eta_1}(\mathbf{x}_k), \mathbf{x}^* - T_{\eta_1}(\mathbf{x}_k) \rangle. \end{aligned}$$

From Proof 5, we get

$$\begin{aligned} H(\mathbf{x}_{k+\frac{1}{2}}) - H(\mathbf{x}^*) \leq & l(T^1_{\eta_1}(\mathbf{x}_k), u_k) - l(\mathbf{x}^*) + h(T_{\eta_1}(\mathbf{x}_k)) - h(\mathbf{x}^*) \\ \leq & \eta_1 \langle \mathbf{x}_k - T_{\eta_1}(\mathbf{x}_k), T_{\eta_1}(\mathbf{x}_k) - \mathbf{x}^* \rangle + \frac{\eta_1}{2} \|T^1_{\eta_1}(\mathbf{x}_k) - \mathbf{n}_k\|^2 \\ = & -\frac{1}{2\eta_1} \|G_{\eta_1}(\mathbf{x}_k)\|^2 + \langle G_{\eta_1}(\mathbf{x}_k), \mathbf{x}_k - \mathbf{x}^* \rangle \\ \leq & \|G_{\eta_1}(\mathbf{x}_k)\| \cdot \|\mathbf{x}_k - \mathbf{x}^*\| \\ = & \|G^1_{\eta_1}(\mathbf{x}_k)\| \cdot \|\mathbf{x}_k - \mathbf{x}^*\|. \end{aligned}$$

By a similar analysis, we can also show that

$$H(\mathbf{x}_{k+1}) - H(\mathbf{x}^*) \leq \|G^2_{\eta_2}(\mathbf{x}_{k+\frac{1}{2}})\| \cdot \|\mathbf{x}_{k+\frac{1}{2}} - \mathbf{x}^*\|,$$

and the proof follows.

Lemma 6. *Let $\{\mathbf{x}_k\}$ be the sequence generated by Algorithm 2. Then $\forall k \geq 0$*

$$H(\mathbf{x}_k) - H(\mathbf{x}_{k+1}) \geq \frac{1}{2\min\{\eta_1, \eta_2\} M^2} (H(\mathbf{x}_{k+1}) - H^*)^2,$$

where $M := \max \|\mathbf{x} - \mathbf{x}^*\|$.

Proof. By Lemma 4 and Lemma 5, we have

$$\begin{aligned} H(\mathbf{x}_k) - H(\mathbf{x}_{k+1}) &\geq H(\mathbf{x}_k) - H(\mathbf{x}_{k+\frac{1}{2}}) \\ &\geq H(\mathbf{x}_k) - H(T_{\eta_1}(\mathbf{x}_k)) \\ &\geq \frac{1}{2\eta_1}\|G^1_{\eta_1}(\mathbf{x}_k)\|^2 \\ &\geq \frac{(H(\mathbf{x}_{k+\frac{1}{2}}) - H(\mathbf{x}^*))^2}{2\eta_1 M^2} \\ &\geq \frac{(H(\mathbf{x}_{k+1}) - H(\mathbf{x}^*))^2}{2\eta_1 M^2}. \end{aligned}$$

Similarly,

$$\begin{aligned} H(\mathbf{x}_k) - H(\mathbf{x}_{k+1}) &\geq H(\mathbf{x}_{k+\frac{1}{2}}) - H(\mathbf{x}_{k+1}) \\ &\geq H(\mathbf{x}_{k+\frac{1}{2}}) - H(T_{\eta_2}(\mathbf{x}_{k+\frac{1}{2}})) \\ &\geq \frac{1}{2\eta_2}\|G^2_{\eta_2}(\mathbf{x}_{k+\frac{1}{2}})\|^2 \\ &\geq \frac{(H(\mathbf{x}_{k+1}) - H(\mathbf{x}^*))^2}{2\eta_2 M^2}. \end{aligned}$$

Therefore,

$$H(\mathbf{x}_k) - H(\mathbf{x}_{k+1}) \geq \frac{1}{2\min\{\eta_1, \eta_2\}M^2}(H(\mathbf{x}_{k+1}) - H^*)^2.$$

Hence, the lemma is proved.

Lemma 7. *Let $\{A_k\}_{k\geq 0}$ be a nonnegative sequence of real numbers such that $A_k - A_{k+1} \geq \gamma A_{k+1}^2$, $A_1 \leq \frac{1.5}{\gamma}$ and $A_2 \leq \frac{1.5}{2\gamma}$ for some positive γ. Then $A_k \leq \frac{1.5}{\gamma}\frac{1}{k}$ for $k > 0$.*

Proof. We use induction to prove. For $k = 1, 2$, it holds. Assume it also holds for k, that is $A_k \leq \frac{1.5}{\gamma}\frac{1}{k}$. For case $k+1$, we have $\gamma A_{k+1}^2 + A_{k+1} \leq A_k \leq \frac{1.5}{\gamma}\frac{1}{k}$, which is same as saying $(\gamma A_{k+1} + \frac{1}{2})^2 \leq \frac{1.5}{k} + \frac{1}{4}$. Therefore

$$\gamma A_{k+1} \leq \sqrt{\frac{1.5}{k} + \frac{1}{4}} - \frac{1}{2} = \frac{1.5}{\sqrt{0.5k - 1 + (\frac{k}{2}+1)^2} + \frac{k}{2}}.$$

Since $k \geq 2$, we have $\gamma A_{k+1} \leq \frac{1.5}{k+1}$. The proof thus follows.

Theorem 8. *Let $\{\mathbf{x}_k\}_{k\geq 0}$ be the sequence generated by Algorithm 2. Then $\forall k \geq 1$*

$$H(\mathbf{x}_k) - H^* \leq \frac{\max\{2(H(\mathbf{x}_0) - H^*), 3\min\{\eta_1, \eta_2\}M^2\}}{k},$$

where $M := \max\|\mathbf{x} - \mathbf{x}^*\|$.

Proof. Set $A_k = H(\mathbf{x}_k) - H^*$ and $\tilde{\gamma} = 1/(2\min\{\eta_1, \eta_2\}M^2)$. By Lemma 6, we have

$$A_k - A_{k+1} \geq \tilde{\gamma} A_{k+1}^2.$$

Since $A_1 = H(\mathbf{x}_1) - H^* \leq H(\mathbf{x}_0) - H^*$ and $A_2 = H(\mathbf{x}_2) - H^* \leq H(\mathbf{x}_0) - H^*$, by setting $\gamma = \min\{\frac{1.5}{2(H(\mathbf{x}_0)-H^*)}, \frac{1}{2\min\{\eta_1,\eta_2\}M^2}\}$, we have

$$A_k - A_{k+1} \geq \gamma A_{k+1}^2,$$

and $A_1 \leq \frac{1.5}{\gamma}$ and $A_2 \leq \frac{1.5}{2\gamma}$. By Lemma 7, we obtain

$$H(\mathbf{x}_k) - H^* \leq \frac{\max\{2(H(\mathbf{x}_0) - H^*), 3\min\{\eta_1, \eta_2\}M^2\}}{k},$$

Hence the lemma is proved.

Lemma 9. *Let $\{A_k\}_{k\geq 0}$ be a nonnegative sequence of real numbers such that $A_k - A_{k+1} \geq \gamma A_{k+1}^2$. Then $A_k \leq \max\{(\frac{1}{2})^{\frac{k-1}{2}} A_0, \frac{4}{\gamma(k-1)}\}$ for all $k > 1$, and $\forall \varepsilon > 0$, $A_k \leq \varepsilon$ if $k \geq \max\{\frac{2}{ln(2)}(ln(A_0) - ln(\varepsilon)), \frac{4}{\gamma\varepsilon}\} + 1$*

Proof. By the fact that $A_k - A_{k+1} \geq \gamma A_{k+1}^2$, we have

$$\frac{1}{A_{k+1}} - \frac{1}{A_k} = \frac{A_k - A_{k+1}}{A_k A_{k+1}} \geq \gamma \frac{A_{k+1}}{A_k}.$$

For the case $A_{k+1} > \frac{1}{2}A_k$, we have

$$\frac{1}{A_{k+1}} - \frac{1}{A_k} \geq \frac{\gamma}{2}.$$

Suppose that k is even and there are at least $\frac{k}{2}$ elements satisfy $A_{k+1} > \frac{1}{2}A_k$, then $\frac{1}{A_k} \geq \frac{k\gamma}{4}$, i.e. $A_k \leq \frac{4}{k\gamma}$. If there are less than $\frac{k}{2}$ elements satisfying $A_{k+1} > \frac{1}{2}A_k$, which implies that there are at least $\frac{k}{2}$ elements $A_{k+1} \leq \frac{1}{2}A_k$, then $A_k \leq (\frac{1}{2})^{\frac{k}{2}} A_0$. Therefore for both cases, we have $A_k \leq \max\{(\frac{1}{2})^{\frac{k}{2}} A_0, \frac{4}{k\gamma}\}$. If k is odd, then $A_k \leq A_{k-1} \leq \max\{(\frac{1}{2})^{\frac{k-1}{2}} A_0, \frac{4}{(k-1)\gamma}\}$. By comparison, we obtain $A_k \leq \max\{(\frac{1}{2})^{\frac{k-1}{2}} A_0, \frac{4}{(k-1)\gamma}\}$ for all $k > 1$.

In order to guarantee the inequality $A_k \leq \varepsilon$, we must have

$$\max\{(\frac{1}{2})^{\frac{k-1}{2}} A_0, \frac{4}{(k-1)\gamma}\} \leq \varepsilon.$$

It leads to a set of two inequalities $(\frac{1}{2})^{\frac{k-1}{2}} A_0 \leq \varepsilon$, and $\frac{4}{(k-1)\gamma} \leq \varepsilon$, which is the same as

$$k \geq \frac{2}{\ln(2)}(\ln(A_0) - \ln(\varepsilon)) + 1,$$

and

$$k \geq \frac{4}{\varepsilon\gamma} + 1.$$

Therefore,

$$k \geq \max\{\frac{2}{\ln(2)}(\ln(A_0) - \ln(\varepsilon)), \frac{4}{\varepsilon\gamma}\} + 1$$

is sufficient to guarantee $A_k \leq \varepsilon$.

THEOREM 10. *Let* $\{\mathbf{x}_k\}_{k\geq 0}$ *be the sequence generated by Algorithm 2. Then* $\forall k \geq 2$

$$H(\mathbf{x}_k) - H^* \leq \max\{(\frac{1}{2})^{\frac{k-1}{2}}(H(\mathbf{x}_0) - H^*), \frac{8\min\{\eta_1,\eta_2\}M^2}{k-1}\},$$

and $H(\mathbf{x}_k) - H^* \leq \varepsilon$ *is obtained after*

$$\max\{\frac{2}{ln(2)}(ln(H(\mathbf{x}_0) - H^*) - ln(\varepsilon)), \frac{8\min\{\eta_1,\eta_2\}M^2}{\varepsilon}\} + 1,$$

iterations where $M := \max\|\mathbf{x} - \mathbf{x}^*\|$.

Proof. By LEMMA 6, and setting $A_k = H(\mathbf{x}_k) - H^*$ and $\gamma = \frac{1}{2\min\{\eta_1,\eta_2\}M^2}$, we have $A_k - A_{k+1} \geq \gamma A_{k+1}^2$. The proof then follows directly from LEMMA 9.

4. Conclusions. We proposed a model combining the two minimization steps of the original TV-Stokes for image denoising, into a single minimization step. The model optimizes a functional with respect to the intensity of the objective image and its gradient in a single step model. We have applied an alternating minimization algorithm to solve the model. The first iteration of the algorithm is exactly the modified variant of the TV-Stokes, cf. [6]. The model is also equivalent to the second order Total Generalized Variation (TGV), cf. [5], when $\eta_1 = 0$.

REFERENCES

[1] A. BECK, *On the Convergence of Alternating Minimization for Convex Programming with Applications to Iteratively Reweighted Least Squares and Decomposition Schemes*, SIAM Journal on Optimization, 25 (2015), pp. 185–209, https://doi.org/10.1137/13094829X, https://doi.org/10.1137/13094829X.

[2] A. BECK AND M. TEBOULLE, *Gradient-Based Algorithms with Applications to Signal Recovery Problems*, Convex Optimization in Signal Processing and Communications, (2010), pp. 3–51, https://doi.org/10.1017/CBO9780511804458.003, http://www.math.tau.ac.il/{~}teboulle/papers/gradient_chapter.pdf.

[3] A. CHAMBOLLE AND P.-L. LIONS, *Image Recovery via Total Variation Minimization and Related Problems*, Numerische Mathematik, 76 (1997), pp. 167–188, https://doi.org/10.1007/s002110050258, http://link.springer.com/10.1007/s002110050258.

[4] C. A. ELO, A. MALYSHEV, AND T. RAHMAN, *A Dual Formulation of the TV-Stokes Algorithm for Image Denoising*, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 5567 LNCS (2009), pp. 307–318, https://doi.org/10.1007/978-3-642-02256-2_26.

[5] F. KNOLL, K. BREDIES, T. POCK, AND R. STOLLBERGER, *Second Order Total Generalized Variation (TGV) for MRI*, Magnetic Resonance in Medicine, (2011), https://doi.org/10.1002/mrm.22595.

[6] W. G. LITVINOV, T. RAHMAN, AND X.-C. TAI, *A Modified TV-Stokes Model for Image Processing*, SIAM Journal on Scientific Computing, 33 (2011), pp. 1574–1597, https://doi.org/10.1137/080727506, http://epubs.siam.org/doi/10.1137/080727506.

[7] Y. NESTEROV, *Introductory Lectures on Convex Optimization: A Basic Course*, Springer US, 1 ed., 2004, https://www.springer.com/gp/book/9781402075537.

[8] J. M. ORTEGA AND W. C. RHEINBOLDT, *Iterative Solution of Nonlinear Equations in Several Variables*, Society for Industrial and Applied Mathematics, 1987, https://doi.org/10.1137/1.9780898719468.

[9] D. P. BERTSEKAS, *Nonlinear Programming*, Athena Scientific, 2 ed., 1999.

[10] D. P. BERTSEKAS AND J. N. TSITSIKLIS, *Distributed And Parallel Computing*, Athena Scientific, 1997.

[11] T. RAHMAN, X. TAI, AND S. OSHER, *A TV-Stokes Denoising Algorithm*, in Scale Space and Variational Methods in Computer Vision, Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2007, pp. 473–483, https://doi.org/10.1007/978-3-540-72823-8_41, https://doi.org/10.1007/978-3-540-72823-8_41.

[12] L. I. Rudin, S. Osher, and E. Fatemi, *Nonlinear Total Variation Noise Removal Algorithm*, Physica D: Nonlinear Phenomena, 60 (1992), pp. 259–268, https://doi.org/10.1016/0167-2789(92)90242-F, http://www.csee.wvu.edu/{~}xinl/courses/ee565/total_variation.pdf, https://arxiv.org/abs/arXiv:1011.1669v3.

[13] M. Sion, *On General Minimax Theorems*, Pacific Journal of Mathematics, 18 (1957), pp. 171–176, https://doi.org/1103040253, http://projecteuclid.org/DPubS?service=UI{&}version=1.0{&}verb=Display{&}handle=euclid.pjm/1103040253{%}5Cnhttp://projecteuclid.org/DPubS?verb=Display{&}version=1.0{&}service=UI{&}handle=euclid.pjm/1103040253{&}page=record.

Bibliography

[1] Aubert, G., and Kornprobst, P. *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*, second edi ed. Springer Science + Business Media, LLC, 2006.

[2] Bayram, I., and Kamasak, M. E. A Directional Total Variation. *European Signal Processing Conference 19*, 12 (2012), 265–269, doi: 10.1109/LSP.2012.2220349.

[3] Bertalmio, M., Bertozzi, A., and Sapiro, G. Navier-Stokes, Fluid Dynamics, and Image and Video Inpainting. *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001 1* (2001), I–355–I–362, doi: 10.1109/CVPR.2001.990497.

[4] Bredies, K., Kunisch, K., and Pock, T. Total Generalized Variation. *SIAM Journal on Imaging Sciences* (2010), doi: 10.1137/090769521.

[5] Chambolle, A. An Algorithm for Total Variation Minimization and Applications. *Journal of Mathematical Imaging and Vision 20*, 2 (2004), 89–97, doi: 10.1023/B:JMIV.0000011325.36760.1e.

[6] Chan, R., Chan, T. F., and Yip, A. Numerical Methods and Applications in Total Variation Image Restoration. *Handbook of Mathematical Methods in Imaging*, May (2011), 1059–1094, doi: 10.1007/978-0-387-92920-0_24.

[7] Chan, T. F., Golub, G. H., and Mulet, P. A Nonlinear Primal-Dual Method for Total Variation-based Image Restoration. *SIAM Journal on Scientific Computing 20*, 6 (1999), 1964–1977, doi: 10.1137/S1064827596299767.

[8] Chan, T. F., and Shen, J. *Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods*. SIAM, 2005.

[9] Chen, D. Q., Cheng, L. Z., and Su, F. A New TV-Stokes Model with Augmented Lagrangian Method for Image Denoising and Deconvolution. *Journal of Scientific Computing 51*, 3 (2012), 505–526, doi: 10.1007/s10915-011-9519-x.

[10] Elo, C. A. *Image Denoising Algorithms Based on the Dual Formulation of Total Variation*. Master thesis, University of Bergen, 2009.

[11] Elo, C. A., Malyshev, A., and Rahman, T. A Dual Formulation of the TV-Stokes Algorithm for Image Denoising. *Lecture Notes in Computer Science*

(including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 5567 LNCS, 3 (2009), 307–318, doi: 10.1007/978-3-642-02256-2_26.

[12] Evans, L. *Partial Differential Equations*, 2nd ed. American Mathematical Society, 2010.

[13] Goldfarb, D., and Yin, W. Second-order Cone Programming Methods for Total Variation-Based Image Restoration. *SIAM Journal on Scientific Computing 27*, 2 (2005), 622–645, doi: 10.1137/040608982.

[14] Gonzalez, R. C., and Woods, R. E. *Digital Image Processing (3rd Edition)*, 3rd ed. Prentice-Hall, 2007.

[15] Hahn, J., Qiu, J., Sugisaki, E., Jia, L., Tai, X.-C., and Seah, H. S. Stroke-Based Surface Reconstruction. *Numerical Mathematics: Theory, Methods and Applications 6*, 1 (2013), 297–324, doi: 10.1017/S1004897900001240.

[16] Hahn, J., Tai, X.-C., Borok, S., and Bruckstein, A. M. Orientation-Matching Minimization for Image Denoising and Inpainting. *International Journal of Computer Vision* (2011), doi: 10.1007/s11263-010-0371-5.

[17] Hahn, J., Wu, C., and Tai, X.-C. Augmented Lagrangian Method for Generalized TV-Stokes Model. *Journal of Scientific Computing 50*, 2 (2012), 235–264, doi: 10.1007/s10915-011-9482-6.

[18] He, B., and Yuan, X. Convergence Analysis of Primal-Dual Algorithms for Total Variation Image Restoration. *Optimization Online* (2010), 1–22.

[19] Holm, R. *Image Inpainting Using Nonlinear Partial Differential Equations*. Master thesis, University of Bergen, 2005.

[20] Jia, Z. G., and Wei, M. A New TV-Stokes Model for Image Deblurring and Denoising with Fast Algorithms. *Journal of Scientific Computing 72*, 2 (2017), 522–541, doi: 10.1007/s10915-017-0368-0.

[21] Kaltenbacher, B., Neubauer, A., and Scherzer, O. *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*. Walter de Gruyter GmbH & Co. KG, 2008.

[22] Knoll, F., Bredies, K., Pock, T., and Stollberger, R. Second Order Total Generalized Variation (TGV) for MRI. *Magnetic Resonance in Medicine* (2011), doi: 10.1002/mrm.22595.

[23] Litvinov, W. G., Rahman, T., and Tai, X.-C. A Modified TV-Stokes Model for Image Processing. *SIAM Journal on Scientific Computing 33*, 4 (2011), 1574–1597, doi: 10.1137/080727506.

[24] Loan, C. V. *Computational Frameworks for the Fast Fourier Transform*. SIAM, 1992.

[25] Lysaker, M., Osher, S., and Tai, X.-C. Iterative Image Restoration Combining Total Variation Minimization and a Second-Order Functional. *IEEE Trans. Image Process 13*, 10 (2004), 1345–1357.

[26] Marcinkowski, L., and Rahman, T. An Iterative Regularization Algorithm for the TV-Stokes in Image Processing. In *Parallel Processing and Applied Mathematics, Lecture Notes in Computer Science* (2016), vol. 9574, Springer, Cham, pp. 381–390.

[27] Meyer, Y. *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, vol. 22. American Mathematical Society, 2001.

[28] Nocedal, J., and Wright, S. J. *Numerical optimization*, 2nd ed. Springer-Verlag New York, 2006.

[29] Osher, S., Burger, M., Goldfarb, D., Xu, J., and Yin, W. An Iterative Regularization Method for Total Variation-Based Image Restoration. *Multiscale Modeling & Simulation 4*, 2 (2005), 460–489, doi: 10.1137/040605412.

[30] Rahman, T., Tai, X., and Osher, S. A TV-Stokes Denoising Algorithm. In *Scale Space and Variational Methods in Computer Vision, Lecture Notes in Computer Science* (2007), Springer, Berlin, Heidelberg, pp. 473–483.

[31] Rudin, L. I., Osher, S., and Fatemi, E. Nonlinear Total Variation Noise Removal Algorithm. *Physica D: Nonlinear Phenomena 60*, 1-4 (1992), 259–268, doi: 10.1016/0167-2789(92)90242-F.

[32] Scherzer, O. *Handbook of Mathematical Methods in Imaging*, 1st ed. Springer-Verlag New York, 2011.

[33] Scherzer, O., Grasmair, M., Grossauer, H., Haltmeier, M., and Lenzen, F. *Variational Methods in Imaging*, vol. 167. Springer Science + Business Media, LLC, 2009.

[34] Sion, M. On General Minimax Theorems. *Pacific Journal of Mathematics 18*, 4 (1957), 171–176, doi: 1103040253.

[35] Tai, X.-C., Osher, S., and Holm, R. Image Inpainting Using a TV-Stokes Equation. *CAM report 06*, 01 (2006), 1–20.

[36] Tai, X.-C., and Wu, C. Augmented Lagrangian Method, Dual Methods and Split Bregman Iteration for ROF Model. *SIAM Journal on Imaging Sciences 3*, 3 (2009), 300–339, doi: 10.1137/090767558.

[37] Vese, L. A., and Osher, S. J. Modeling Textures with Total Variation Minimization and Oscillating Patterns in Image Processing. *Journal of Scientific Computing* (2003), doi: 10.1023/A:1025384832106.

[38] Wu, B., Rahman, T., and Tai, X. C. Sparse-Data Based 3D Surface Reconstruction for Cartoon and Map. *Mathematics and Visualization 0*, 221219 (2018), 47–64, doi: 10.1007/978-3-319-91274-5_3.

[39] Wu, C., and Tai, X.-C. Augmented Lagrangian Method, Dual Methods,and Split Bregman Iteration for ROF,Vectorial TV, and High Order Models. *Methods 3* (2010), 300–339, doi: 10.1137/090767558.

[40] Wu, C., Zhang, J., Duan, Y., and Tai, X.-C. Augmented Lagrangian Method for Total Variation Based Image Restoration and Segmentation Over Triangulated Surfaces. *Journal of Scientific Computing 50*, 1 (2012), 145–166, doi: 10.1007/s10915-011-9477-3.

[41] Wu, C., Zhang, J., and Tai, X.-C. Augmented Lagrangian Method for Total Variation Restoration with Non-quadratic Fidelity. *Inverse Problems and Imaging 5*, 1 (2011), 237–261, doi: 10.3934/ipi.2011.5.237.

[42] Wu, T., Yang, Y., and Jing, H. Two-step Methods for Image Zooming Using Duality Strategies. *Numerical Algebra, Control and Optimization 4*, 3 (2014), 209–225, doi: 10.3934/naco.2014.4.209.

[43] Yang, J., Zhang, Y., and Yin, W. An Efficient TVL1 Algorithm for Deblurring multichannel Images Corrupted by Impulsive Noise. *SIAM Journal on Scientific Computing 31*, 4 (2009), 2842–2865, doi: 10.1137/080732894.

[44] Zhu, M., and Chan, T. An Efficient Primal-Dual Hybrid Gradient Algorithm for Total Variation Image Restoration. *UCLA CAM Report*, 1 (2008), 1–29.

Graphic design: Communication Division, UiB / Print: Skipnes Kommunikasjon AS

uib.no

ISBN: 9788230843543 (print)
9788230846971 (PDF)