Sparse Dowker nerves

Morten Brun[1] · Nello Blaser[1]

Abstract

We propose sparse versions of filtered simplicial complexes used to compute persistent homology of point clouds and of networks. In particular, we extend the Sparse Čech Complex of Cavanna et al. (A geometric perspective on sparse filtrations. CoRR, arXiv:1506.03797, 2015) from point clouds in convex metric spaces to point clouds in arbitrary metric spaces. Along the way we formulate interleaving in terms of strict 2-categories, and we introduce the concept of Dowker dissimilarities that can be considered as a common generalization of metric spaces and networks.

1 Introduction

This paper is the result of an attempt to obtain the interleaving guarantee for the sparse Čech complex of Cavanna et al. (2015) without using the Nerve Theorem. The rationale for this was to generalize the result from convex metric spaces to arbitrary metric spaces. We have not been able to show that the constructions of Sheehy (2013) or Cavanna et al. (2015) are interleaved with the Čech complex in arbitrary metric spaces. However, changing the construction slightly, we obtain a sub-complex of the Čech complex that is interleaved in a similar way. When applied to point clouds in a convex metric space this sub-complex is homotopic to the construction of Cavanna et al. (2015). The above constructions are two examples of general constructions presented in the body of our paper.

✉ Morten Brun
morten.brun@uib.no

Nello Blaser
nello.blaser@uib.no

[1]   Department of Mathematics, University of Bergen, Bergen, Norway

The search for a more general version of the sparse Čech complex led us to study both different versions of filtered covers and extended metrics. We discovered that these concepts are instances of filtered relations given by functions of the form

$$\Lambda \colon L \times W \to [0, \infty]$$

from the product of two sets L and W to the interval $[0, \infty]$. Here we think of L as a set of landmark points and W as a set of witness points. Given $t \in [0, \infty)$, the relation Λ_t at filtration level t is

$$\Lambda_t = \{(l, w) \in L \times W \mid \Lambda(l, w) < t\}.$$

Dowker (1952) observed that a relation $R \subseteq L \times W$ gives a cover $(R(l))_{l \in L}$ of the set

$$R_W = \{w \in W \mid \text{ there exists } l \in L \text{ with } (l, w) \in R\}$$

with

$$R(l) = \{w \in W \mid (l, w) \in R\}.$$

The *Dowker complex* of the relation R is the Borsuk Nerve of this cover. The *Dowker Homology Duality Theorem* (Dowker 1952, Theorem 1) states that the Dowker complexes of R and the transposed relation

$$R^T = \{(w, l) \mid (l, w) \in R\} \subseteq W \times L$$

have isomorphic homology. In Chowdhury and Mémoli (2018), Chowdhury and Mémoli have sharpened the Dowker Homology Duality Theorem to a Dowker Homotopy Duality Theorem stating that the Dowker complexes of R and R^T are homotopy equivalent after geometric realization. That result is a central ingredient in this paper.

In honor of Dowker we name functions $\Lambda \colon L \times W \to [0, \infty]$ *Dowker dissimilarities*. Forming the Dowker complexes of the relations Λ_t for $t \in [0, \infty)$ we obtain a filtered simplicial complex, the *Dowker Nerve $N\Lambda$ of Λ*, with $N\Lambda_t$ equal to the Dowker complex of Λ_t.

The main result of our work is Theorem 2 on sparsification of Dowker nerves. Here we formulate it in the context of a finite set P contained in a metric space (M, d). Let $p_0, \ldots, p_n$ be a farthest point sampling of P with insertion radii $\lambda_0, \ldots \lambda_n$. That is, $p_0 \in P$ is arbitrary, $\lambda_0 = \infty$ and for each $0 < k \leq n$, the point $p_k \in P$ is of maximal distance to $p_0, \ldots, p_{k-1}$, and this distance is λ_k. Let $\varepsilon > 0$ and let $\Lambda \colon P \times M \to [0, \infty]$ be the Dowker dissimilarity given by the metric d, that is, $\Lambda(p, w) = d(p, w)$. Then the Dowker Nerve $N\Lambda$ is equal to the relative Čech complex $\check{C}(P, M)$ of P in M given by the Borsuk Nerve of all balls in M centred at points in P.

Let $[n] = \{0, \ldots, n\}$ and let $\varphi \colon [n] \to [n]$ be a function with $\varphi(0) = 0$ and $\varphi(k) < k$ for $k > 0$ and

$$d(p_k, p_{\varphi(k)}) + (\varepsilon + 1)\lambda_k/\varepsilon \leq (\varepsilon + 1)\lambda_{\varphi(k)}/\varepsilon \tag{1}$$

for $k = 1, \ldots, n$. The geometric idea behind this formula is that the ball centred at p_k of radius $\frac{(\varepsilon+1)\lambda_k}{\varepsilon}$ is contained in the ball centred at $p_{\varphi(k)}$ of radius $\frac{(\varepsilon+1)\lambda_{\varphi(k)}}{\varepsilon}$, and therefore at filtration levels greater than $\frac{(\varepsilon+1)\lambda_{\varphi(k)}}{\varepsilon}$ this ball can be omitted without changing the homotopy type of the nerve of the cover (Fig. 1). The *Sparse Dowker Nerve* of Λ is the filtered sub-complex $N(\Lambda, \varphi, (\varepsilon + 1)\lambda/\varepsilon)$ of $N\Lambda$ filtered by letting $N(\Lambda, \varphi, (\varepsilon + 1)\lambda/\varepsilon))_t$ consist of subsets $\sigma \subseteq P$ such that there exists $w \in M$ with

$$d(p_k, w) < \min\{t, (\varepsilon + 1)\lambda_k/\varepsilon, (\varepsilon + 1)\lambda_{\varphi(l)}/\varepsilon\}$$

for every $k, l \in \sigma$. The geometric interpretation of this formula is firstly that balls around p_k are truncated at radius $(\varepsilon + 1)\lambda_k/\varepsilon$, that is, when they reach this radius they stop growing. Secondly, the nerve is restricted in the sense that p_k is not contributing to simplices of radius greater than $(\varepsilon + 1)\lambda_{\varphi(k)}/\varepsilon$. In the geometric description of Cavanna et al. (2015), cones at p_k are truncated at width $(\varepsilon + 1)\lambda_k/\varepsilon$ and their tops are cut off at height $(\varepsilon + 1)^2\lambda_k/\varepsilon$.

Theorem 1 *The Sparse Dowker Nerve $N(\Lambda, \varphi, (\varepsilon + 1)\lambda/\varepsilon)$ is multiplicatively $(1, 1 + \varepsilon)$-interleaved with the relative Čech complex $\check{C}(P, M)$ of P in M.*

Explicitly, there are maps $f_t \colon N\Lambda_t \to N(\Lambda, \varphi, (\varepsilon + 1)\lambda/\varepsilon)_{(1+\varepsilon)t}$ so that for the inclusion maps $g_t \colon N(\Lambda, \varphi, (\varepsilon + 1)\lambda/\varepsilon)_t \to N\Lambda_t$, the maps $f_t g_t$ and $g_{(1+\varepsilon)t} f_t$ are homotopic to the inclusion of the space of level t into the space of level $(1 + \varepsilon)t$ of the filtered simplicial complexes $N(\Lambda, \varphi, (\varepsilon + 1)\lambda/\varepsilon)$ and $N\Lambda$ respectively. For $M = \mathbb{R}^d$ the Sparse Dowker Nerve is closely related to the Sparse Čech Complex of Cavanna et al. (2015). We have implemented both constructions and made them available at GitHub (Brun and Blaser 2018). We also provide a tutorial that can be accessed from the same GitHub repository. It turned out that the two constructions are of similar size.

Chazal et al. (2014) witness complexes and Čech complexes are both instances of Dowker dissimilarities. The weighted Čech complex in (Buchet et al. 2016, Definition 5.1) is also an instance of a Dowker complex. Also the filtered clique complex of a finite weighted undirected simple graph $G = (V, w)$, where w is a function $w \colon G \times G \to [0, \infty]$ is an instance of a Dowker nerve: let $\mathcal{P}(V)$ be the set of subsets of V and define

$$\Lambda \colon V \times \mathcal{P}(V) \to [0, \infty], \qquad (x, X) \mapsto \begin{cases} \mathrm{diam}(X) & \text{if } x \in X \\ \infty & \text{otherwise,} \end{cases}$$

where $\mathrm{diam}(X) = \max_{x, x'' \in X} w(x, x')$. Then the Dowker Nerve of Λ is equal to the filtered clique complex of G.

For disjoint sets L and W a Dowker dissimilarity $\Lambda \colon L \times W \to [0, \infty]$ is the same thing as a weighted simple bipartite graph. On the other hand, a Dowker dissimilarity
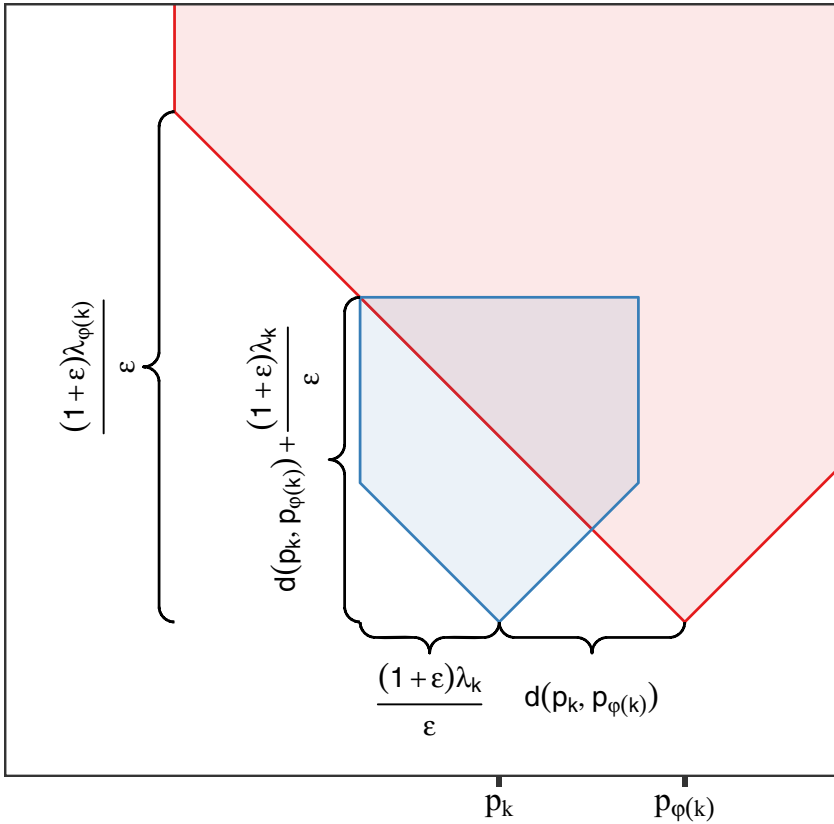
Fig. 1 *Sparse Dowker nerves.* Let $P = \{p_0, \ldots, p_n\}$ be a finite subset of a metric space M, d obtained by farthest point sampling with insertion radii $\lambda_0, \ldots, \lambda_n$ and let $\varphi\colon [n] \to [n]$ be a function satisfying the inequality (1). Around each point $p_k \in X$, we draw a truncated cone with width $\frac{(\varepsilon+1)}{\varepsilon}\lambda_k$ and height $d(p_k, p_{\varphi(k)}) + \frac{(\varepsilon+1)}{\varepsilon}\lambda_k$. The figure shows that the projection of the cone with center p_k onto M is contained in the projection of the cone with center $p_{\varphi(k)}$

of the form $\Lambda\colon X \times X \to [0, \infty]$ is the same thing as a weighted directed graph with no multiple directed edges. Chowdhury and Mémoli (2018), call Dowker dissimilarities of this form weighted networks, and study their Dowker nerves thoroughly under the name Dowker complexes. In particular, they show that the persistent homology of the Dowker Nerve of a network is sensitive to the direction its edges. For example, for the networks A and B in Fig. 2, with self-loops of weight 0, the Dowker Nerve of network A is contractible while the Dowker Nerve of network B is homotopic to a circle at all filtration levels. Chowdhury and Mémoli also formulate a stability result for homology of Dowker nerves (Chowdhury and Mémoli 2018). We formulate interleaving of Dowker dissimilarities in such a way that their network distance is bounded below by our interleaving distance. Together with functoriality for interleaving distance and the Algebraic Stability Theorem (Chazal et al. 2009) this implies the
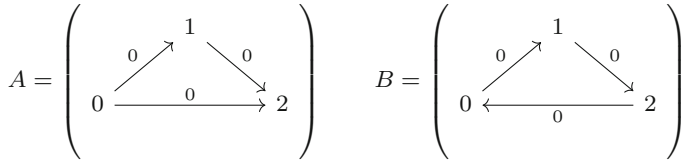
$$A = \begin{pmatrix} & & 1 & & \\ & {}^{0}\nearrow & & \searrow^{0} & \\ 0 & \xrightarrow{0} & & & 2 \end{pmatrix} \qquad B = \begin{pmatrix} & & 1 & & \\ & {}^{0}\nearrow & & \searrow^{0} & \\ 0 & \xleftarrow{0} & & & 2 \end{pmatrix}$$

Fig. 2 The Dowker Nerve of network A is contractible while the Dowker Nerve of network B is homotopic to a circle

stability result of Chowdhury and Mémoli (2018). In the context of metric spaces, this Stability Theorem is contained in Chazal et al. (2014).

Imposing conditions on a Dowker dissimilarity of the form

$$\Lambda \colon X \times X \to [0, \infty],$$

we arrive at concepts of independent interest. Most importantly, (X, Λ) is a metric space if and only if Λ satisfies

Finiteness $\Lambda(x, y) < \infty$ for all $x, y \in X$.
Triangle inequality $\Lambda(x, z) \leq \Lambda(x, y) + \Lambda(y, z)$ for $x, y, x \in X$.
Identity of indiscernibles $d(x, y) = 0$ if and only if $x = y$.
Symmetry $d(x, y) = d(y, x)$ for all $x, y \in X$.

Removing some of the above conditions on Λ leads to various generalizations of metric spaces. In particular, the situation where Λ only is required to satisfy the triangle inequality has been studied by Lawvere (1957). He noticed that $[0, \infty]$ is a closed symmetric monoidal category and that when the triangle inequality holds, then Λ gives X the structure of a category enriched over $[0, \infty]$.

Guided by the Strong Functorial Dowker Theorem (Chowdhury and Mémoli 2018, Theorem 3) we have chosen to work with interleavings in the homotopy category instead of on the level of homology groups. We leave it for further investigation to decide if the Functorial Dowker Theorem can be extended to homotopy interleavings in the sense of Blumberg and Lesnick (2017).

We extend the usual notion of interleaving between $[0, \infty)$-filtered objects in two ways. Firstly, we consider interleavings in 2-categories. We were led to do this because Dowker dissimilarities form a 2-category, and the proof of the Stability Theorem is streamlined by working in this generality. Secondly, following Bubenik et al. (2015), we allow interleaving with respect to order preserving functions of the form $\alpha \colon [0, \infty) \to [0, \infty)$ satisfying $t \leq \alpha(t)$ for all t. In this context, additive interleaving corresponds to functions of the form $\alpha(t) = t + a$ and multiplicative interleaving corresponds to functions of the form $\alpha(t) = ct$.

After setting terminology and notation, the proof of our main result, Theorem 2, is a quite simple application of the functorial Dowker Theorem. It consists of two parts. First, we truncate the Dowker dissimilarity associated to a metric by replacing certain distances by infinity and show that the truncated Dowker dissimilarity is interleaved with the original Dowker dissimilarity. At that point we use the functorial Dowker

Theorem. Second, we give conditions that allow us to sparsify the Dowker Nerve of the truncated Dowker dissimilarity without changing the filtered homotopy type.

The paper is organized as follows: Sect. 2 presents our main result, Theorem 2. We also show how Theorem 1 is a consequence of Theorem 2 and how the Sparse Čech complex (Cavanna et al. 2015) fits into this context. Section 3 shows that, under certain conditions, when some of the values $\Lambda(l, w)$ in a Dowker dissimilarity are set to infinity, the homotopy type of the Dowker Nerve is only changed up to a certain interleaving. This is the first step in our proof of Theorem 2. In Sect. 4, we give a criterion ensuring that a certain sub-complex is homotopy equivalent to the Dowker Nerve of a Dowker dissimilarity. In Sect. 5, we present the homotopy category of simplicial complexes. In Sect. 6, we recollect basic terminology about 2-categories. The main motivation for going to this level of generality is that interleaving distance in the 2-category Dow of Dowker dissimilarities defined in Definition 32 generalizes network distance from Chowdhury and Mémoli (2018). In Sect. 7, we introduce the 2-category of sets and relations. Section 8 uses the Dowker Nerve construction to define a 2-category with relations as objects. Section 9 introduces interleavings in 2-categories. In Sect. 10, we define the 2-category of Dowker dissimilarities. In Sect. 11, we relate interleaving distance of Dowker dissimilarities to Gromov–Hausdorff distance of metric spaces. We finish with some concluding remarks in Sect. 12.

2 Sparse nerves of truncated Dowker dissimilarities

In this section, we state and motivate our main results, postponing proofs to later sections.

Definition 1 A *Dowker dissimilarity* Λ consists of two sets L and W and a function $\Lambda\colon L \times W \to [0, \infty]$. Given $t \in [0, \infty)$, we let

$$\Lambda_t = \{(l, w) \in L \times W \mid \Lambda(l, w) < t\}.$$

The *Dowker Nerve* of Λ is the filtered simplicial complex $N\Lambda = (N\Lambda_t)_{t \geq 0}$ with

$$N\Lambda_t = \{\text{finite } \sigma \subseteq L \mid \text{there exists } w \in W \text{ with } \Lambda(l, w) < t \text{ for all } l \in \sigma\}.$$

The *zero set* of Λ is

$$L_\Lambda = \{l \in L \mid \text{there exists } w \in W \text{ with } \Lambda(l, w) = 0\}.$$

The triangle inequality is central in the work of Cavanna et al. (2015), and it will also play a central role for us. For Dowker dissimilarities we need extra structure in order to formulate the triangle inequality.

Definition 2 A Dowker dissimilarity $\Lambda\colon L \times W \to [0, \infty]$ satisfies the *triangle inequality* if the following holds:

1. For every $w \in W$ there exists $l \in L$ with $\Lambda(l, w) = 0$.
2. For all $(l, w) \in L \times W$ with $\Lambda(l, w) = 0$ and all $(l', w') \in L \times W$, the *triangle inequality*

$$\Lambda(l', w') \leq \Lambda(l', w) + \Lambda(l, w')$$

holds.

Note that if Λ satisfies the triangle inequality and if both $\Lambda(l, w) = 0$ and $\Lambda(l', w) = 0$, then for every $w' \in W$ the triangle inequality implies that

$$\Lambda(l', w') \leq \Lambda(l', w) + \Lambda(l, w') = \Lambda(l, w'),$$

and thus by symmetry $\Lambda(l', w') = \Lambda(l, w')$. Also note that if Λ is actually a metric $\Lambda \colon L \times L \to [0, \infty]$, then it satisfies the triangle inequality in the above sense.

Unfortunately the Sparse Dowker Nerve presented in this paper needs a Dowker dissimilarity satisfying the triangle inequality as input. There are many interesting examples of Dowker dissimilarities where it is not satisfied. For example the triangle inequality is rarely satisfied for weighted networks or Bregman divergences. Typical examples of Dowker dissimilarities satisfying the triangle inequality are quasi-metrics and pseudo-metrics.

The following concept of insertion pairs is the crucial ingredient in Sect. 3 where we treat interleaving of truncated Dowker Nerves.

Definition 3 An *insertion pair for* Λ consists of a pair (f, λ) of functions $f \colon W \times [0, \infty] \to W$ and $\lambda \colon W \to [0, \infty]$ with the property that for every $(l, w) \in L \times W$ with $\Lambda(l, w) = 0$ and every $t \geq 0$ the inequalities

$$\Lambda(l, f(w, t)) \leq t < \lambda(f(w, t))$$

are satisfied.

Insertion pairs generalize farthest point samples. Let $d \colon W \times W \to [0, \infty]$ be a metric and $w_0, \ldots, w_n$ be a farthest point sample with insertion times $\lambda_0, \ldots, \lambda_n$. Then the functions $\lambda(w_i) = \lambda_i$ and $f(w, t) = \text{argmin}_{w_i \in W}\{d(w, w_i) \mid \lambda_i > t\}$ are an insertion pair.

For Dowker dissimilarities of the form $\Lambda \colon L \times [n] \to [0, \infty]$ satisfying the triangle inequality the following definition gives an example of an insertion pair for Λ. Here and throughout this paper we will use the notation $[n] = \{0, \ldots, n\}$.

Definition 4 Let $\Lambda \colon L \times [n] \to [0, \infty]$ be a Dowker dissimilarity satisfying the triangle inequality. The *canonical insertion function for* $\Lambda \colon L \times [n] \to [0, \infty]$ is the function $\lambda_\Lambda \colon [n] \to [0, \infty]$ defined as

$$\lambda_\Lambda(w) = \begin{cases} \infty & \text{if } w = 0 \\ \sup_{l \in L_\Lambda} \inf_{w' \in [w-1]} \Lambda(l, w') & \text{if } w > 0. \end{cases}$$

The *canonical insertion pair for* Λ is the pair $(f_\Lambda, \lambda_\Lambda)$ where the function

$$f_\Lambda \colon [n] \times [0, \infty] \to [n]$$

is defined as follows: Given $w \in [n]$ and $t \geq 0$, pick $l \in L$ with $\Lambda(l, w) = 0$ and define

$$f_\Lambda(w, t) = \min\{w' \in [n] \mid \Lambda(l, w') \leq t\}.$$

The definition of $f_\Lambda(w, t)$ is independent of the choise of l and the set

$$\{w' \in [n] \mid \Lambda(l, w') \leq t\}$$

is non-empty because

$$\sup_{l \in L_\Lambda} \inf_{w' \in [n]} \Lambda(l, w') = 0 \leq t.$$

Also, by construction $\Lambda(l, w) = 0$ implies that

$$\Lambda(l, f_\Lambda(w, t)) \leq t < \lambda_\Lambda(f_\Lambda(w, t)).$$

The following lemma summarizes this discussion.

Lemma 1 *Let $\Lambda \colon L \times [n] \to [0, \infty]$ be a Dowker dissimilarity satisfying the triangle inequality. Then the canonical insertion pair $(f_\Lambda, \lambda_\Lambda)$ is an insertion pair for Λ.*

The Sparse Čech Nerve of Cavanna et al. (2015) is constructed with a positive number ε as a parameter controlling its quality in the sense that the Sparse Čech Nerve is multiplicatively $(1, 1 + \varepsilon)$-interleaved with the Čech Nerve. The following definition allows us to construct Sparse Dowker Nerves that are (id, α)-interleaved with the Dowker Nerve for a large class of translation functions $\alpha \colon [0, \infty) \to [0, \infty)$. Here a *translation function* is an order-preserving function $\alpha \colon [0, \infty) \to [0, \infty)$ with the property that $\alpha(t) \geq t$ for every $t \in [0, \infty)$. In the following definition, we use the generalized inverse of order preserving functions (see e.g. Embrechts and Hofert 2013 for more details).

Definition 5 Let $\alpha \colon [0, \infty] \to [0, \infty]$ be order preserving with

$$\lim_{t \to \infty} \alpha(t) = \infty.$$

The generalized inverse function $\alpha^\leftarrow \colon [0, \infty] \to [0, \infty]$ is the order preserving function

$$\alpha^\leftarrow(s) = \inf\{t \in [0, \infty] \mid \alpha(t) \geq s\}.$$

The above definition can be summarized as follows:

$$\alpha^\leftarrow(s) \leq t \iff s \leq \alpha(t).$$

Definition 6 Let $\beta\colon [0, \infty] \to [0, \infty]$ be an order preserving function. The *translation function associated to β* is the function $\alpha\colon [0, \infty] \to [0, \infty]$ defined by

$$\alpha(t) = t + \beta(t).$$

The first step in the construction of a Sparse Dowker Nerve of a Dowker dissimilarity Λ, is to truncate Λ by replacing some of the values $\Lambda(l, w)$ by ∞. In the Euclidean case, this means that we truncate cones by only allowing them to grow to a certain width, as illustrated in Fig. 1. With the objective of the Sparse Dowker Nerve being (α, id)-interleaved with the Dowker Nerve in the sense of Definition 28. We do this as follows:

Definition 7 Let $\Lambda\colon L \times W \to [0, \infty]$ be a Dowker dissimilarity satisfying the triangle inequality and let (f, λ) be an insertion pair for Λ. Let $\beta\colon [0, \infty] \to [0, \infty]$ be an order preserving function with $\lim_{t\to\infty} \beta(t) = \infty$ and let α be the translation function associated to β. The *β-truncation function $\lambda_\beta\colon W \to W$ associated to λ* is the function

$$\lambda_\beta(w) = \alpha\beta^\leftarrow(\lambda(w)).$$

The *(λ, β)-truncation* of Λ is the Dowker dissimilarity $\Lambda^{(\lambda,\beta)}\colon L \times W \to [0, \infty]$ defined by

$$\Lambda^{(\lambda,\beta)}(l, w) = \begin{cases} \Lambda(l, w) & \text{if } \Lambda(l, w) \leq \lambda_\beta(w) \text{ and } \beta(0) \leq \lambda(w) \\ \infty & \text{otherwise.} \end{cases}$$

Example 1 Let $\Lambda\colon L \times [n] \to [0, \infty]$ be a Dowker dissimilarity satisfying the triangle inequality. Given $a \geq 0$ and $\varepsilon > 0$, we define $\beta(t) = a + \varepsilon t$ for $t \in [0, \infty]$. The translation function α associated to β is the affine function $\alpha(t) = a + (1 + \varepsilon)t$, and the function $\alpha\beta^\leftarrow$ is given by the formula

$$(\alpha\beta^\leftarrow)(t) = \begin{cases} \frac{(1+\varepsilon)t - a}{\varepsilon} & \text{if } t \geq a/(1 + \varepsilon) \\ 0 & \text{otherwise.} \end{cases}$$

In Theorem 2 below we state that persistent homology of the (λ, β)-truncation of Λ is (α, id)-interleaved with the persistent homolgoy of Λ.

The last step in the construction of a Sparse Dowker Nerve of a Dowker dissimilarity Λ, is to restrict the Dowker Nerve of its *transposed Dowker dissimilarity* $\Lambda^T\colon W \times L \to [0, \infty]$ given by $\Lambda^T(w, l) = \Lambda(l, w)$. By restricting we mean that $w \in W$ is not contributing to simplices of radius greater than a certain threshold. The purpose of the following definition is to give such a threshold.

Definition 8 Let $\Lambda\colon L \times [n] \to [0, \infty]$ be a Dowker dissimilarity and let $\lambda\colon [n] \to [0, \infty]$ be a function with $\lambda(w) < \infty$ for $w > 0$ and $\lambda(0) = \infty$. The *parent function for Λ with respect to λ* is the function

$$\varphi\colon [n] \to [n]$$

defined by $\varphi(0) = 0$ and

$$\varphi(w) = \operatorname{argmin}\{\lambda(w') \mid \lambda(w') > \lambda(w) \text{ and for all } l \in L \text{ with}$$
$$\Lambda(l, w) < \lambda(w), \text{ we have } \Lambda(l, w') < \lambda(w')\}$$

for $w > 0$ in $[n]$.

Note that if $w > 0$, then $\lambda(w) < \infty$ and $\lambda(w) < \lambda(\varphi(w))$. Also note that the parent function is not necessarily uniquely defined because there may be more than one w' with $\lambda(w')$ minimal under the condition that both $\lambda(w') > \lambda(w)$ and $\Lambda(l, w) < \lambda(w)$ implies $\Lambda(l, w') < \lambda(w')$. We now give the last ingredient in the construction of the Sparse Dowker Nerve.

Definition 9 Let $\Gamma \colon [n] \times W \to [0, \infty]$ be a Dowker dissimilarity. Given functions $\varphi \colon [n] \to [n]$ and $\lambda \colon [n] \to [0, \infty]$, the *sparse nerve of Γ with respect to λ and φ* is the filtered simplicial complex $N(\Gamma, \varphi, \lambda)$ with $N(\Gamma, \varphi, \lambda)(t)$ consisting of all $\sigma \subseteq [n]$ so that there exists $w \in W$ with

$$\Gamma(l, w) < \min(t, \lambda(l), \lambda(\varphi(l'))) \quad \text{for all } l, l' \in \sigma.$$

In abbreviated form the simplicial complex $N(\Gamma, \varphi, \lambda)(t)$ is

$$\{\sigma \subseteq [n] \mid \exists w \in W \colon \Gamma(l, w) < \min(t, \lambda(l), \lambda(\varphi(l'))) \text{ for all } l, l' \in \sigma\}.$$

We are ready to state our main result about the Sparse Dowker Nerve:

Theorem 2 *Let $\Lambda \colon L \times [n] \to [0, \infty]$ be a Dowker dissimilarity satisfying the triangle inequality and with $\Lambda(l, 0) < \infty$ for every $l \in L$. Let $\beta \colon [0, \infty] \to [0, \infty]$ be an order preserving function with $\beta(t) < \infty$ for $t < \infty$ and $\lim_{t \to \infty} \beta(t) = \infty$ and let λ_β be the β-truncation function associated to the canonical insertion function for Λ. If we let φ be the parent function for Λ with respect to λ_β and write $\Gamma = (\Lambda^{(\lambda_\Lambda, \beta)})^T$, then the Dowker Nerve $N\Lambda^T$ of Λ^T is (α, id)-interleaved with the filtered simplicial complex $N(\Gamma, \varphi, \lambda_\beta)$.*

Proof We show the interleaving in four steps

$$N\Lambda \overset{\mathrm{FDT}}{\simeq} N\Lambda^T \overset{\mathrm{P2}}{\longrightarrow} N(\Lambda^{(\lambda_\Lambda, \beta)}) \overset{\mathrm{FDT}}{\simeq} N\Gamma \overset{\mathrm{P3}}{\simeq} N(\Gamma, \varphi, \lambda_\beta),$$

where Proposition 2 (P2) gives the interleaving and the homotopy equivalences are given by the Strong Functorial Dowker Theorem (Chowdhury and Mémoli 2018, Theorem 3) (FDT) and Proposition 3 (P3).

More precisely, the functorial Dowker Theorem (Chowdhury and Mémoli 2018, Theorem 3) implies that the filtered simplicial complexes $N\Lambda$ and $N\Lambda^T$ are homotopy equivalent. By construction in Definitions 6 and 7, if $\Gamma(k, l) < \infty$ then $\Gamma(k, l) < \lambda_\beta(k)$. Therefore, by Proposition 3 the filtered simplicial complexes $N(\Gamma, \varphi, \lambda_\beta)$ and $N\Gamma$ are homotopy equivalent. The Strong Functorial Dowker Theorem (Chowdhury and Mémoli 2018, Theorem 3) implies that the filtered simplicial complexes $N\Gamma$ and

$N(\Lambda^{(\lambda_\Lambda, \beta)})$ are homotopy equivalent. By Lemma 1 the pair $(f_\Lambda, \lambda_\Lambda)$ is an insertion pair for Λ, so by Proposition 2 the filtered simplicial complexes $N\Lambda$ and $N(\Lambda^{(\lambda_\Lambda, \beta)})$ are (α, id)-interleaved. □

Remark 1 The Dowker Theorem (Dowker 1952, Theorem 1a) implies that the persistent homology of $N\Lambda$ is isomorphic to the persistent homology of $N\Lambda^T$ and that the persistent homology of $N\Gamma$ is isomorphic to the persistent homology of $N(\Lambda^{(\lambda_\Lambda, \beta)})$. Thus, if we are only interested in persistent homology, we do not need the Functorial Dowker Theorem.

Corollary 1 *Let* $\Lambda \colon L \times [n] \to [0, \infty]$ *be a Dowker dissimilarity satisfying the triangle inequality and with* $\Lambda(l, 0) < \infty$ *for every* $l \in L$ *and let* $\beta \colon [0, \infty] \to [0, \infty]$ *be the function*

$$\beta(t) = \varepsilon t$$

and let $\alpha \colon [0, \infty] \to [0, \infty]$ *be the translation function*

$$\alpha(t) = (\varepsilon + 1)t$$

associated to β. *For* λ_β *the* β-truncation function associated to the canonical insertion function of Λ and φ the parent function for Λ with respect to λ_β, the Dowker Nerve $N\Lambda^T$ of Λ^T is (α, id)-interleaved with the filtered simplicial complex*

$$N((\Lambda^{(\lambda_\Lambda, \beta)})^T, \varphi, \lambda_\beta).$$

Specializing even further, we obtain a variation of the Sparse Čech complex of Cavanna et al. (2015). Recall that the Hausdorff distance of a metric space $L = (L, d)$ and a subset P of L is

$$d_H(L, P) = \sup_{l \in L} \inf_{p \in P} d(l, p).$$

In the following result we have in mind the situation where L is a compact submanifold of Euclidean space $\mathbb{R}^d$ with convex hull M and P is a sample of points from L. Note that in this situation the relative Čech complexes $\check{\mathcal{C}}(L, M)$ and $\check{\mathcal{C}}(L, \mathbb{R}^d)$ are homotopy equivalent.

Corollary 2 *Let* (M, d) *be a metric space, let* $L \subseteq M$ *be a compact subset, let* P *be a finite subset of* L *and let* $[n] \xrightarrow{p} P$ *be a bijection. Let* $\Lambda \colon M \times [n] \to [0, \infty]$ *be the function*

$$\Lambda(x, w) = d(x, p_w),$$

where we write $p_w = p(w)$. *Let* $\varepsilon > 0$ *and let* $\beta \colon [0, \infty] \to [0, \infty]$ *be the function* $\beta(t) = \varepsilon t$ *with associated translation function* $\alpha(t) = (\varepsilon + 1)t$. *For* λ_β *the* β-truncation *function associated to the canonical insertion function of* Λ *and* φ *the parent function*

for Λ with respect to λ_β, the Dowker Nerve $N\Lambda^T$ of Λ^T is (α, id)-interleaved with the filtered simplicial complex

$$N((\Lambda^{(\lambda_\Lambda, \beta)})^T, \varphi, \lambda_\beta).$$

Moreover, the Dowker Nerve of $N\Lambda^T$ is additively $(0, d_H(L, P))$-interleaved with the relative Čech complex $\check{C}(L, M)$ consisting of all balls in M with centers in L.

Sometimes we refer to the relative Čech complex $\check{C}(L, M)$ as the ambient Čech complex and to the relative Čech complex $\check{C}(L, L)$ as the intrinsic Čech complex of L.

Proof Corollary 1 gives that $N\Lambda$ is (α, id)-interleaved with

$$N((\Lambda^{(\lambda_\Lambda, \beta)})^T, \varphi, \lambda_\beta).$$

For the second statement, first note that $N\Lambda^T$ is homotopy equivalent to the relative Čech complex $\check{C}(P, M)$. Thus it suffices to show that the Čech complexes $\check{C}(P, M)$ and $\check{C}(L, M)$ are additively $(0, d_H(L, P))$-interleaved. Since $P \subseteq L$, for every $t \in [0, \infty)$, there is an inclusion $\iota_t : \check{C}_t(P, M) \to \check{C}_t(L, M)$. Let $f : L \to P$ be a function with $d(l, f(l)) \leq d_H(L, P)$ for every $l \in L$. The triangle inequality implies that f induces a morphism $f_t : \check{C}_t(L, M) \subseteq \check{C}_{t+d_H(L,P)}(L, M)$. By construction the composites $f_t \iota_t$ and $\iota_{t+d_H(L,P)} f_t$ are contiguous to the respective identity maps, and thus their geometric realizations are homotopic Spanier (1966). □

Finally, we show that the Sparse Čech complex of Cavanna et al. (2015) is an instance of a Sparse Dowker Nerve. We suspect that Cavanna et al.'s construction performed to Dowker dissimilarities is homotopy equivalent to the Sparse Dowker Nerve but we have not been able to prove this. However, the following result, together with the first part of the proof of (Cavanna et al. 2015, Theorem 5) implies that the Cavanna et al.'s Sparse Čech Nerve is homotopy equivalent to the Sparse Dowker Nerve for the Dowker dissimilarities in the following proposition.

Proposition 1 *Let d be a convex metric on $\mathbb{R}^d$ and let P be a finite subset of $\mathbb{R}^d$ together with a greedy order $[n] \xrightarrow{p} P$. Let the function $\Lambda : \mathbb{R}^d \times [n] \to [0, \infty]$ be given by*

$$\Lambda(l, w) = d(l, p_w),$$

where we write $p_w = p(w)$. Let $\varepsilon > 0$ and let $\beta : [0, \infty] \to [0, \infty]$ and $\alpha : [0, \infty] \to [0, \infty]$ be the functions $\beta(t) = \varepsilon t$ and $\alpha(t) = (1 + \varepsilon)t$. Let λ_Λ be the canonical insertion function for Λ and let $\lambda = ((1 + \varepsilon)^2 / \varepsilon) \lambda_\Lambda$. Then the filtered simplicial complex

$$N((\Lambda^{(\lambda_\Lambda, \beta)})^T, \mathrm{id}, \lambda)(t)$$

is isomorphic to the filtered simplicial complex $\bigcap_{\varepsilon' > \varepsilon} C(\varepsilon')$ where

$$C(\varepsilon')_t = \bigcup_{s < t} S^s(\varepsilon')$$

and $S(\varepsilon') = \{S^t(\varepsilon')\}_{t \geq 0}$ is the sparse Čech complex constructed in (Cavanna et al. 2015, Sect. 4) *for the parameter ε'.*

Proof A subset $\sigma \subseteq [n]$ is in

$$N((\Lambda^{(\lambda_\Lambda, \beta)})^T)$$

if and only if there exists $w \in \mathbb{R}^d$ so that for all $l \in \sigma$ we have

$$d(p_l, w) < t \quad \text{and} \quad d(p_l, w) < \lambda_\Lambda(l)(1 + \varepsilon)/\varepsilon.$$

Moreover

$$\sigma \in N(((\Lambda^T)^{(\lambda_\Lambda, \beta)})^T, \mathrm{id}, \lambda)(t)$$

if and only if there exists $w \in \mathbb{R}^d$ so that for all $k, l \in \sigma$ we have $d(p_k, w) < t$ and

$$d(p_k, w) < \lambda_\Lambda(k)(1 + \varepsilon)/\varepsilon \quad \text{and} \quad d(p_k, w) < \lambda_\Lambda(l)(1 + \varepsilon)^2/\varepsilon.$$

On the other hand, $\sigma \in \bigcup_{s < t} S^s(\varepsilon')$ if and only if there exists $s < t$ and $w \in \mathbb{R}^d$ so that $w \in b_l(s)$ for all $l \in \sigma$. By the definition of $b_l(s)$ in Cavanna et al. (2015), Sect. 3, this is the case if and only if $s < t$ and

$$s \leq \lambda_\Lambda(l)(1 + \varepsilon')^2/\varepsilon' \quad \text{and} \quad d(p_l, w) \leq \min(s, \lambda_\Lambda(l)(1 + \varepsilon')/\varepsilon')$$

for every $l \in \sigma$. We conclude that $\sigma \in S^t$ if and only if there exists $w \in \mathbb{R}^d$ satisfying $d(p_l, w) < t$ and

$$d(p_l, w) \leq \lambda_\Lambda(l)(1 + \varepsilon')/\varepsilon' \quad \text{and} \quad d(p_k, w) < \lambda_\Lambda(l)(1 + \varepsilon')^2/\varepsilon'.$$

for all $k, l \in \sigma$. $\qquad \square$

We have not performed any complexity analysis of Sparse Dowker Nerves. Instead we have made proof-of-concept implementations of both the Sparse Čech Complex of Cavanna et al. (2015) described in Proposition 1 and the Sparse Dowker Nerve described in Corollary 2. These implementations come with the same interleaving guarantees, but for practical reasons concerning the miniball algorithm we consider complexes that are slightly bigger than the ones described above. More specifically, we compute filtration values using the miniball algorithm instead of computing radius of simplices in the truncated Dowker nerve. This gives us a filtered simplicial complex that lies between the truncated Dowker nerve and the full Dowker nerve, so it has the same interleaving guarantee as the truncated Dowker nerve.

We have tested these implementations on the following data: the optical patch data sets called $X(300, 30)$ and $X(15, 30)$ in Carlsson et al. (2008), 6040 points from the cyclo-octane conformation space as analysed in Zomorodian (2012), the Clifford data set consisting of 2000 points on a curve on a torus considered in Oudot (2015), Chapter

Table 1 Sizes of sparse Dowker nerves using Sparse Čech Complex of Cavanna et al. (2015) described in Proposition 1 (Sheehy) and the Sparse Dowker Nerve described in Corollary 2 (parent)

Data	Complex	Dim	ε	Unreduced (millions)	Sheehy (millions)	Parent (millions)
Clifford torus	Intrinsic	2	2.0	2593.9	1.3	2.8
Clifford torus	Ambient	2	2.0	2593.9	1.4	1.0
Cyclo-octane	Intrinsic	1	2.0	20.8	1.0	1.0
Cyclo-octane	Ambient	1	2.0	20.8	1.0	1.0
Double torus	Intrinsic	2	2.0	2593.9	1.5	2.9
Double torus	Ambient	2	2.0	2593.9	1.8	1.9
$X(15, 30)$	Intrinsic	1	1.5	20.8	3.3	3.2
$X(15, 30)$	Ambient	1	1.5	20.8	3.3	3.2
$X(300, 30)$	Intrinsic	1	2.5	20.8	4.9	1.8
$X(300, 30)$	Ambient	1	2.5	20.8	1.8	1.7

In order to run these examples, we first used a farthest point sampling to take 500 points and calculated persistent homology of the subsamples. In the table we specify the name of the dataset (data), if persistent homology was computed using intrinsic or ambient Čech complex (complex), the homology dimension (dim), and the sparsification constant (ε). Note that both sparsifications result in substantially smaller nerves than the unreduced nerve. In general the two sparsifications result in nerves of similar sizes, but that for each sparsification there are outliers which are substantially larger

5, and the double torus from Dey et al. (2016) (Table 1). Computing the Sparse Čech complexes and the Sparse Dowker Nerves on these data sets with the same interleaving constant ε the resulting simplicial complexes are almost of the same size, with the size of the Sparse Dowker Nerve slightly smaller than the size of the Sparse Čech Complex. Our implementations, the data sets mentioned above and the scripts used to run compute persistent homology are available (Brun and Blaser 2018).

3 Truncated Dowker dissimilarities

In this section, we provide the interleaving guarantee for the truncated Dowker dissimilarity used in the first step of the construction of the Sparse Dowker Nerve.

Lemma 2 *Let $\Lambda\colon L \times W \to [0, \infty]$ be a Dowker dissimilarity satisfying the triangle inequality and let (f, λ) be an insertion pair for Λ. If $\beta\colon [0, \infty] \to [0, \infty]$ is an order preserving function satisfying $\beta(t) < \infty$ for $t < \infty$ and $\lim_{t \to \infty} \beta(t) = \infty$ with associated translation function $\alpha\colon [0, \infty] \to [0, \infty]$, then for every $t \in [0, \infty)$, the simplicial complex $N\Lambda_t$ is contained in the simplicial complex $N\Lambda_{\alpha(t)}^{(\lambda, \beta)}$.*

Proof Let $t \in [0, \infty)$ and $\sigma \in N\Lambda_t$. We need to show that $\sigma \in N\Lambda_{\alpha t}^{(\lambda, \beta)}$. Pick $w \in W$ with $\Lambda(l, w) < t$ for all $l \in \sigma$. Since Λ satisfies the triangle inequality we can pick $l_0 \in L$ so that $\Lambda(l_0, w) = 0$. Let $w_0 = f(w, \beta(t))$. Since (f, λ) is an insertion pair for Λ we have

$$\Lambda(l_0, w_0) \leq \beta(t) < \lambda(w_0).$$

The triangle inequality for Λ now gives

$$\Lambda(l, w_0) \leq \Lambda(l_0, w_0) + \Lambda(l, w).$$

Let $l \in \sigma$. Since $\Lambda(l, w) < t$ and $\Lambda(l_0, w_0) \leq \beta(t)$, we get that

$$\Lambda(l, w_0) < \beta(t) + t = \alpha(t).$$

The inequality $\beta(t) < \lambda(w_0)$ implies that the set $A = \{s \mid \lambda(w_0) \leq \beta(s)\}$ is contained in (t, ∞). By definition $\beta^{\leftarrow}(\lambda(w_0)) = \inf A$, so we get $t \leq \beta^{\leftarrow}(\lambda(w_0))$. Since α is order preserving this gives $\alpha(t) \leq \alpha\beta^{\leftarrow}(\lambda(w_0))$ and thus $\Lambda(l, w_0) < \alpha\beta^{\leftarrow}(\lambda(w_0))$. We conclude that $\sigma \in N\Lambda_{\alpha(t)}^{(\lambda, \beta)}$. $\qquad\square$

Proposition 2 *Let $\Lambda \colon L \times W \to [0, \infty]$ be a Dowker dissimilarity satisfying the triangle inequality and let (f, λ) be an insertion pair for Λ. If $\beta \colon [0, \infty] \to [0, \infty]$ is an order preserving function satisfying $\beta(t) < \infty$ for $t < \infty$ and $\lim_{t \to \infty} \beta(t) = \infty$ with associated translation function $\alpha \colon [0, \infty] \to [0, \infty]$, then the filtered simplicial complexes $N\Lambda$ and $N\Lambda^{(\lambda, \beta)}$ are (α, id)-interleaved.*

Proof By Lemma 2, for every $t \in [0, \infty)$, the simplicial complex $N\Lambda_t$ is contained in the simplicial complex $N\Lambda_{\alpha(t)}^{(\lambda, \beta)}$. Since $\Lambda(l, w) \leq \Lambda^{(\lambda, \beta)}(l, w)$ for all $(l, w) \in L \times W$, the simplicial complex $N\Lambda_t^{(\lambda, \beta)}$ is contained in the simplicial complex $N\Lambda_t$ for every $t \in [0, \infty)$. $\qquad\square$

4 Sparse Dowker nerves

In this section, we provide the technical result ensuring that the Dowker Nerve of the truncated Dowker dissimilarity can be further sparsified without changing its homotopy type.

Proposition 3 *Let $\Lambda \colon [n] \times W \to [0, \infty]$ be a Dowker dissimilarity and let $\varphi \colon [n] \to [n]$ be a parent function for Λ^T with respect to $\lambda \colon [n] \to [0, \infty]$. If $\Lambda(l, w) < \infty$ implies $\Lambda(l, w) < \lambda(l)$, then for every $t \in [0, \infty)$ the inclusion*

$$\iota \colon N(\Lambda, \varphi, \lambda)(t) \overset{\subseteq}{\to} N\Lambda_t$$

is a homotopy equivalence.

Proof By permuting the elements of $[n]$ we can, without loss of generality, assume that

$$\lambda(\varphi(l)) < \lambda(\varphi(l')) \text{ implies } l > l'.$$

Note that if $\varphi(l) > 0$, then $\lambda(\varphi(l)) < \lambda(\varphi(\varphi(l)))$, and deduce that $l > \varphi(l)$ for every $l > 0$.

Let $t \in [0, \infty)$. We will prove the assertion by induction on n. The statement obviously holds when $\lambda(\varphi(n)) \geq t$. In particular, it holds for $n = 0$. Let $n \geq 1$ and suppose by induction that the statement holds for $n - 1$. We assume that $\lambda(\varphi(n)) < t$ since otherwise the statement obviously holds for n. Let $\Lambda' : [n-1] \times W \to [0, \infty]$ be the restriction of Λ to $[n-1] \times W \subseteq [n] \times W$ and let λ' and φ' the restrictions of λ and φ to $[n-1]$. Then φ' is the parent function for Λ' with respect to λ' so the induction hypothesis is satisfied for Λ', φ' and λ'.

We define $f_t : [n] \to [n-1]$ by

$$f_t(l) = \begin{cases} \varphi(n) & \text{if } l = n \\ l & \text{otherwise.} \end{cases}$$

Given $\sigma \in N\Lambda_t$ we claim that $\sigma \cup f_t(\sigma) \in N\Lambda_t$. If $n \notin \sigma$, then this claim is trivial. In order to justify the claim when $n \in \sigma$, we pick $w \in W$ with $\Lambda(l, w) < t$ for every $l \in \sigma$. By our assumption on Λ we then also have $\Lambda(l, w) < \lambda(l)$ for every $l \in \sigma$. By definition of the parent function, $\Lambda(n, w) < \lambda(n)$ implies $\Lambda(\varphi(n), w) < \lambda(\varphi(n)) < t$. We conclude that $\sigma \cup f_t(\sigma) \in N\Lambda_t$.

Our next claim is that if $\sigma \in N(\Lambda, \varphi, \lambda)(t)$, then $\sigma \cup f_t(\sigma) \in N(\Lambda, \varphi, \lambda)(t)$. Again, we only need to consider the case $n \in \sigma$. We have already shown that $\sigma \cup f_t(\sigma) \in N\Lambda_t$. Since n is maximal in $[n]$ we have $\lambda(\varphi(l)) \geq \lambda(\varphi(n))$ for every $l \in [n]$, so this claim follows since $\Lambda(n, w) < \lambda(n)$ implies $\Lambda(\varphi(n), w) < \lambda(\varphi(n))$.

In particular, we can now conclude that the function $f_t : [n] \to [n-1]$ and defines simplicial maps

$$f_t : N\Lambda_t \to N\Lambda'_t$$

and

$$f_t : N(\Lambda, \varphi, \lambda)(t) \to N(\Lambda', \varphi', \lambda')(t).$$

On the other hand, the inclusion $\iota : [n-1] \to [n]$ defines simplicial maps (see Sect. 5)

$$\iota : N\Lambda'_t \to N\Lambda_t$$

and

$$\iota : N(\Lambda', \varphi', \lambda')(t) \to N(\Lambda, \varphi, \lambda)(t).$$

Moreover the above claims imply that the compositions

$$N\Lambda_t \xrightarrow{f_t} N\Lambda'_t \xrightarrow{\iota} N\Lambda_t$$

and

$$N(\Lambda, \varphi, \lambda)(t) \xrightarrow{f_t} N(\Lambda', \varphi', \lambda')(t) \xrightarrow{\iota} N(\Lambda, \varphi, \lambda)(t)$$

are contiguous to the identity maps. Since $f_t \iota$ is the identity this implies that geometric realizations of the inclusions

$$N \Lambda'_t \xrightarrow{\iota} N \Lambda_t$$

and

$$N(\Lambda', \varphi', \lambda')(t) \xrightarrow{\iota} N(\Lambda, \varphi, \lambda)(t)$$

are homotopy equivalences. Since we have assumed by induction that the geometric realization of the inclusion

$$N(\Lambda', \varphi', \lambda')(t) \xrightarrow{\iota} N \Lambda'_t$$

is a homotopy equivalence, we can conclude that the geometric realization of the inclusion

$$N(\Lambda, \varphi, \lambda)(t) \xrightarrow{\iota} N \Lambda_t$$

is a homotopy equivalence. $\qquad\square$

5 The homotopy category of simplicial complexes

The rest of the paper is aimed at stability results for Dowker Nerves. First we need some prerequisites on the homotopy category.

Recall that a simplicial complex $K = (V, K)$ consists of a vertex set V and a set K of finite subsets of V with the property that if σ is a member of K, then every subset of σ is a member of K. Given a subset $V' \subseteq V$ and a simplicial complex $K = (V, K)$, we write $K_{V'}$ for the simplicial complex $K_{V'} = (V', K_{V'})$ consisting of subsets of V' of the form $\sigma \cap V'$ for $\sigma \in K$. The *geometric realization* of a simplicial complex $K = (V, K)$ is the space $|K|$ consisting of all functions $f : V \to [0, 1]$ satisfying:

1. The support $\{v \in V \mid f(v) \neq 0\}$ of f is a member of K
2. $\sum_{v \in V} f(v) = 1$.

If V is finite, then $|K|$ is given the subspace topology of the Euclidean space $\mathbb{R}^V$. Otherwise $U \subseteq |K|$ is open if and only if for every finite $V' \subseteq V$, the set $U \cap |K_{V'}|$ is open in $|K_{V'}|$.

The *order complex* of a partially ordered set $(X, \leq)$ is the simplicial complex with vertex set X consisting of totally ordered finite subsets of X, that is, subsets of the form $\sigma = \{x_0, x_1, \ldots, x_k\}$ with $x_0 < x_1 < \cdots x_k$.

A *simplicial map* $f : K \to L$ of simplicial complexes $K = (V, K)$ and $L = (W, L)$ consists of a function $f : V \to W$ such that

$$f(\sigma) = \{f(v) \mid v \in \sigma\}$$

is in L for every $\sigma \in K$. Observe that a simplicial map $f : K \to L$ induces a continuous map $|f| : |K| \to |L|$ of geometric realizations and that this promotes the geometric realization to a functor $| \cdot | : \mathrm{Cx} \to \mathrm{Top}$ from the category Cx of simplicial complexes and simplicial maps to the category Top of topological spaces and continuous maps.

Definition 10 The *homotopy category* hCx *of simplicial complexes* has the class of simplicial complexes as objects. Given simplicial complexes K and L, the morphism set $\mathrm{hCx}(K, L)$ is the set of homotopy classes of continuous maps from the geometric realization of K to the geometric realization of L. Composition in hCx is given by composition of functions representing homotopy classes.

We remark in passing that the homotopy category of simplicial complexes is equivalent to the weak homotopy category of topological spaces.

6 Background on 2-categories

We have chosen to reason about stability in the context of 2-categories. The pay-off from this high level of abstraction is that the proofs get relatively easy. In this section, we present some terminology on 2-categories. Most of this material is taken from Leinster (1998).

Recall that a 2-category $\mathcal{C}$ consists of

1. A class of objects $A, B, \ldots$
2. For all objects A, B a category $\mathcal{C}(A, B)$. The objects of $\mathcal{C}(A, B)$ are the morphisms in $\mathcal{C}$ and the morphisms $\alpha : f \Rightarrow g$ of $\mathcal{C}(A, B)$ are the 2-cells in $\mathcal{C}$.
3. For every object A of $\mathcal{C}$ there is an identity morphism $\mathrm{id}_A : A \to A$ and an identity 2-cell $\mathrm{id}_{\mathrm{id}_A} : \mathrm{id}_A \Rightarrow \mathrm{id}_A$.
4. For all objects A, B and C of $\mathcal{C}$ there is a functor

$$\mathcal{C}(A, B) \times \mathcal{C}(B, C) \to \mathcal{C}(A, C)$$
$$(f, g) \mapsto g \cdot f,$$

which is associative and admits the identity morphisms and identity 2-cells of $\mathcal{C}$ as identities.

Definition 11 Given 2-categories $\mathcal{C}$ and $\mathcal{D}$, a *functor* $F : \mathcal{C} \to \mathcal{D}$ consists of

1. A function $F : \mathrm{ob}\,\mathcal{C} \to \mathrm{ob}\,\mathcal{D}$
2. Functors $F : \mathcal{C}(A, B) \to \mathcal{D}(FA, FB)$

such that $F(\mathrm{id}_A) = \mathrm{id}_{FA}$ and $Fg \circ Ff = F(g \circ f)$ for A an object of $\mathcal{C}$ and $f : A \to B$ and $g : B \to C$ morphisms of $\mathcal{C}$.

Definition 12 Given two functors $F, G : \mathcal{C} \to \mathcal{D}$ of 2-categories, a *transformation* $\alpha : F \to G$ consists of

1. A morphism $\alpha_A : FA \to GA$ in $\mathcal{D}$ for every $A \in \mathrm{ob}\,\mathcal{C}$
2. A 2-cell $\alpha_f : Gf \circ \alpha_A \to \alpha_B \circ Ff$ for every morphism $f : A \to B$ in $\mathcal{C}$.

This structure is subject to the axioms given by commutativity of the following two diagrams:

$$
\begin{array}{ccc}
 & Gg \circ \alpha_B \circ Ff & \\
{\scriptstyle \mathrm{id}_{Gg} \circ \alpha_f} \nearrow & & \searrow {\scriptstyle \alpha_g \circ \mathrm{id}_{Ff}} \\
Gg \circ Gf \circ \alpha_A \xrightarrow{\quad \alpha_{g \cdot f} \quad} & & \alpha_C \circ Fg \circ Ff
\end{array}
$$

$$
\begin{array}{ccc}
 & \alpha_A & \\
{\scriptstyle \mathrm{id}} \nearrow & & \searrow {\scriptstyle \mathrm{id}} \\
G(\mathrm{id}_A) \circ \alpha_A \xrightarrow{\quad \alpha_{\mathrm{id}_A} \quad} & & \alpha_A \circ F(\mathrm{id}\, A).
\end{array}
$$

Definition 13 Given two functors $F, G \colon \mathcal{C} \to \mathcal{D}$ of 2-categories, and transformations $\alpha, \beta \colon F \to G$, a *modification* $M \colon \alpha \to \beta$ consists of a 2-cell

$$
M_A \colon \alpha_A \to \beta_A
$$

for every object A of $\mathcal{C}$ such that for every morphism $f \colon A \to B$ of $\mathcal{C}$ the following diagram commutes:

$$
\begin{array}{ccc}
Gf \circ \alpha_A & \xrightarrow{\mathrm{id}_{Gf} \circ M_A} & Gf \circ \beta_A \\
{\scriptstyle \alpha_f} \downarrow & & \downarrow {\scriptstyle \beta_f} \\
\alpha_B \circ Ff & \xrightarrow{M_B \circ \mathrm{id}_{Ff}} & \beta_B \circ Ff.
\end{array}
$$

Definition 14 Given 2-categories $\mathcal{C}$ and $\mathcal{D}$, the *functor* 2-*category* $[\mathcal{C}, \mathcal{D}]$ is the 2-category with functors $F \colon \mathcal{C} \to \mathcal{D}$ as objects, transformations of such functors as morphisms and with 2-cells given by modifications.

Given a category $\mathcal{C}$ we will consider it as a 2-category with only identity 2-cells. Thus, if $\mathcal{C}$ is a category and $\mathcal{D}$ is a 2-category we have defined the functor 2-categories $[\mathcal{C}, \mathcal{D}]$ and $[\mathcal{D}, \mathcal{C}]$.

Definition 15 The opposite of a 2-category $\mathcal{C}$ is the 2-category $\mathcal{C}^{\mathrm{op}}$ with the same objects as $\mathcal{C}$, with

$$
\mathcal{C}^{\mathrm{op}}(A, B) = \mathcal{C}(B, A)
$$

and with composition obtained from composition in $\mathcal{C}$.

7 Relations

Dowker dissimilarities can be considered as filtered relations. In order to formulate this precisely, we need some background information on relations.

Definition 16 Let X and Y be sets. A *relation* $R\colon X \leftrightarrows Y$ is a subset $R \subseteq X \times Y$.

Definition 17 We define a partial order on the set of relations between X and Y by set inclusion. That is, for relations $R\colon X \leftrightarrows Y$ and $R'\colon X \leftrightarrows Y$, we have $R \leq R'$ if and only if R is contained in the subset R' of $X \times Y$.

Definition 18 Given two relations $R\colon X \leftrightarrows Y$ and $S\colon Y \leftrightarrows Z$, their composition

$$S \circ R\colon X \leftrightarrows Z$$

is

$$S \circ R = \{(x, z) \in X \times Z \mid \exists\, y \in Y : (x, y) \in R \text{ and } (y, z) \in S, \}.$$

Definition 19 The 2-category $\mathcal{S}$ of sets and relations has as objects the class of sets and as morphisms the class of relations. The 2-cells are given by the inclusion partial order on the class of relations. Composition of morphisms is composition of relations and composition of 2-cells is given by composition of inclusions. The identity morphism on the set X is the diagonal

$$\Delta_X = \{(x, x) \mid x \in X\}.$$

The identity 2-cell on a relation R is the identity inclusion $R \leq R$.

Definition 20 The transposition functor $T\colon \mathcal{S} \to \mathcal{S}^{\mathrm{op}}$ is defined by $T(X) = X$,

$$T(R) = R^T = \{(y, x) \mid (x, y) \in R\}$$

and $T(i) = i^T$, where $i^T\colon R^T \to S^T$ takes (y, x) to (z, w) when $(w, z) = i(x, y)$.

Definition 21 A *correspondence* $C\colon X \leftrightarrows Y$ is a relation such that:

1. For every $x \in X$ there exists $y \in Y$ so that $(x, y) \in C$ and
2. For every $y \in Y$ there exists $x \in X$ so that $(x, y) \in C$.

Lemma 3 *A relation $C\colon X \leftrightarrows Y$ is a correspondence if and only if there exists a relation $D\colon Y \leftrightarrows X$ so that $\Delta_X \leq D \circ C$ and $\Delta_Y \leq C \circ D$.*

Proof By definition of a correspondence, for every $x \in X$, there exists $y \in Y$ so that $(x, y) \in C$. This means that $\Delta_X \subseteq C^T \circ C$, where

$$C^T \circ C = \{(x, z) \in X \times X \mid \exists\, y \in Y : (x, y) \in C \text{ and } (y, x) \in C^T\}.$$

Reversing the roles of C and C^T we get the inclusion $\Delta_Y \subseteq C \circ C^T$. Conversely, if C and D are relations with $\Delta_Y \subseteq C \circ D$, then for every $y \in Y$, the element (y, y) is contained in $C \circ D$. This means that there exists $x \in X$ so that $(x, y) \in C$, and $(y, x) \in D$. In particular for every $y \in Y$, there exists $x \in X$ so that $(x, y) \in C$. Reversing the roles of C and D we get that for every $x \in X$ there exists $y \in Y$ so that $(x, y) \in C$. $\qquad\square$

8 The category of relations

In this section, we introduce a novel 2-category of relations. We start by recalling Dowker's definition of the nerve of a relation (called the complex K in Dowker 1952, Sect. 1).

Definition 22 Let $R \subseteq X \times Y$ be a relation. The *nerve* of R is the simplicial complex

$$NR = \{\text{finite } \sigma \subseteq X \mid \exists y \in Y \text{ with } (x, y) \in R \text{ for all } x \in \sigma\}.$$

Example 2 Let X be a space, and let Y be a cover of X. In particular every element $y \in Y$ is a subset of X. Let R be the relation $R \subseteq X \times Y$ consisting of pairs (x, y) with $x \in y$. A direct inspection reveals that the nerve of R is equal to the Borsuk Nerve of the cover Y.

Definition 23 The 2-category $\mathcal{R}$ of relations has as objects the class of relations. A morphism $C \colon R \to R'$ in $\mathcal{R}$ between relations $R \subseteq X \times Y$ and $R' \subseteq X' \times Y'$ consists of a relation $C \subseteq X \times X'$ such that for every $\sigma \in NR$, the set

$$(NC)(\sigma) = \{x' \in X' \mid \text{ there exists } x \in \sigma \text{ with } (x, x') \in C\}$$

is an element $(NC)(\sigma) \in NR'$ of the nerve of R'. In particular, $(NC)(\sigma)$ is finite and non-empty. The class of 2-cells in $\mathcal{R}$ is the class of inclusions $C_1 \subseteq C_2$ for morphisms $C_1, C_2 \subseteq X \times X'$. Composition in $\mathcal{R}$ is given by composition of relations.

Note that if $C \colon R \to R'$ is a morphism in $\mathcal{R}$, then NC is an order preserving function $NC \colon NR \to NR'$, and thus it induces a simplicial map $NC \colon \Delta(NR) \to \Delta(NR')$ of order complexes. Given a simplicial complex K, the simplicial complex $\Delta(K)$ is the barycentric subdivision of K. The geometric realizations of K and $\Delta(K)$ are homeomorphic. In particular, they represent isomorphic objects in the homotopy category of simplicial complexes.

Lemma 4 *Let $C_0, C_1 \colon R \to R'$ be morphisms in $\mathcal{R}$. If there exists a 2-cell $\alpha \colon C_0 \to C_1$, then the geometric realizations of the simplicial maps*

$$NC_0, NC_1 \colon \Delta(NR) \to \Delta(NR')$$

are homotopic.

Proof Let I be the partially ordered set

$$I = \{0 \to 1\}.$$

Since $C_0 \subseteq C_1$, we can define an order preserving map

$$C \colon I \times NR \to NR'$$

by

$$C(i, \sigma) = NC_i(\sigma) \qquad \text{for } i = 0, 1.$$

On order complexes C induces a simplicial map

$$\Delta(C) \colon \Delta(I \times NR) \to \Delta(NR').$$

As a consequence of Milnor (1981), Theorem 1 and Milnor (1981), Theorem 2 the geometric realization of $\Delta(I \times NR)$ is homeomorphic to the product of the geometric realizations of $\Delta(I)$ and $\Delta(NR)$. The geometric realization of $\Delta(I)$ is homeomorphic to the closed unit interval. Thus we have constructed a homotopy between the geoemtric realizations of NC_1 and NC_2.   □

Definition 24 The *nerve functor* $N \colon \mathcal{R} \to \text{hCx}$ is the functor taking a relation R the order complex $\Delta(NR)$ of its nerve and taking a morphism $C \colon R \to R'$ in $\mathcal{R}$ to the morphism $|NC| \colon |\Delta(NR)| \to |\Delta(NR')|$ in hCx.

Let us emphasize that if $\alpha \colon C_1 \to C_2$ is a 2-cell in $\mathcal{R}$, then $|NC_1| = |NC_2|$ in hCx.

9 Interleavings

In this section, we set the concept of interleaving into the context of 2-categories. This material is well-known in the applied topology community.

We write $[0, \infty)$ for the set of non-negative real numbers and consider it as a partially ordered set. We also consider $[0, \infty)$ as a category with object set $[0, \infty)$ and with a unique morphism $s \to t$ if and only if $s \leq t$.

Definition 25 Let $\mathcal{C}$ be a 2-category. The *category of filtered objects* in $\mathcal{C}$ is the functor 2-category $[[0, \infty), \mathcal{C}]$. A *filtered object* in $\mathcal{C}$ is an object $C \colon [0, \infty) \to \mathcal{C}$ of $[[0, \infty), \mathcal{C}]$, that is, C is a functor from $[0, \infty)$ to $\mathcal{C}$. A *morphism* $f \colon C \to C'$ of filtered objects in $\mathcal{C}$ is a transformation.

We will be so interested in the category of filtered objects in the 2-category of relations that we give a name to this particular 2-category.

Definition 26 A *filtered relation* is a functor from $[0, \infty)$ to $\mathcal{R}$. We define the 2-category of filtered relations to be the 2-category $[[0, \infty), \mathcal{R}]$ of functors from $[0, \infty)$ to $\mathcal{R}$.

Definition 27 Let $\mathcal{C}$ be a 2-category and let $\alpha \colon [0, \infty) \to [0, \infty)$ be a translation function, that is, an order preserving function satisfying $t \leq \alpha(t)$ for all $t \in [0, \infty)$.

1. The pull-back functor $\alpha^* \colon [[0, \infty), \mathcal{C}] \to [[0, \infty), \mathcal{C}]$ is the functor taking a filtered object $C \colon [0, \infty) \to \mathcal{C}$ in $\mathcal{C}$ to the filtered object $\alpha^* C = C \circ \alpha$.
2. The *unit of the functor* $\alpha^* \colon [[0, \infty), \mathcal{C}] \to [[0, \infty), \mathcal{C}]$ is the natural transformation $\alpha_* \colon \text{id} \to \alpha^*$ defined by

$$\alpha_{*C}(t) = C(t \leq \alpha(t)) \colon C(t) \to \alpha^*(C)(t).$$

Definition 28 Let C and C' be filtered objects in a 2-category $\mathcal{C}$ and let $\alpha, \alpha' : [0, \infty) \to [0, \infty)$ be functors under the identity.

1. An (α, α')-interleaving between C and C' is a pair (F, F') of morphisms $F \colon C \to \alpha^* C'$ and $F' \colon C' \to \alpha'^* C$ in $[[0, \infty), \mathcal{C}]$ such that there exist 2-cells

$$(\alpha' \circ \alpha)_* \to (\alpha^* F') \circ F \quad \text{and} \quad (\alpha \circ \alpha')_* \to (\alpha'^* F) \circ F'.$$

2. We say that C and C' are (α, α')-*interleaved* if there exists an (α, α')-interleaving between C and C'.

The following results appear in Bubenik et al. (2015), Propositions 2.2.11 and 2.2.13.

Lemma 5 (Functoriality) *Let C and C' be filtered objects in a 2-category $\mathcal{C}$, let $\alpha, \alpha' : [0, \infty) \to [0, \infty)$ be functors under the identity and let $H \colon \mathcal{C} \to \mathcal{D}$ be a functor of 2-categories. If C and C' are (α, α')-interleaved, then the filtered objects HC and HC' in $\mathcal{D}$ are (α, α')-interleaved.*

Lemma 6 (Triangle inequality) *Let C, C' and C'' be filtered objects in a 2-category $\mathcal{C}$. If C and C' are (α, α')-interleaved and C' and C'' are (β, β')-interleaved, then C and C'' are $(\beta\alpha, \alpha'\beta')$-interleaved.*

10 Filtered relations and Dowker dissimilarities

In this section, we introduce a novel 2-category of Dowker dissimilarities as a certain sub-2-category of the category of filtered relations.

Definition 29 The *filtered nerve functor* is the functor

$$N \colon [[0, \infty), \mathcal{R}] \to [[0, \infty), \mathrm{hTop}]$$

from the 2-category of filtered relations to the category of homotopy filtered spaces taking $X \colon [0, \infty) \to \mathcal{R}$ to the composition

$$[0, \infty) \xrightarrow{X} \mathcal{R} \xrightarrow{N} \mathrm{hTop}.$$

From Lemma 5 we get:

Corollary 3 *If R and R' are (α, α')-interleaved filtered relations, then NR and NR' are (α, α')-interleaved filtered simplicial complexes.*

Given a Dowker dissimilarity $\Lambda \colon L \times W \to [0, \infty]$ and $t \in [0, \infty)$ we consider

$$\Lambda_t = \{(l, w) \in L \times W \mid \Lambda(l, w) < t\}$$

as an object of the category $\mathcal{R}$ of relations. Given $s \leq t$ in $[0, \infty)$ we let

$$\Lambda_{s \leq t} = \Delta_L = \{(l, l) \, l \in L\} \subseteq L \times L$$

considered as a morphism $\Lambda_{s \leq t} \colon \Lambda_s \to \Lambda_t$ in $\mathcal{R}$.

Definition 30 The *filtered relation associated to* a Dowker dissimilarity $\Lambda \colon L \times W \to [0, \infty]$ is the functor

$$\Lambda \colon [0, \infty) \to \mathcal{R}$$

taking $t \in [0, \infty)$ to the relation Λ_t and taking $s \leq t$ in $[0, \infty)$ to the morphism $\Lambda_{s \leq t}$ in $\mathcal{R}$.

Definition 31 Let $\Lambda \colon L \times W \to [0, \infty]$ and $\Lambda' \colon L' \times W' \to [0, \infty]$ be Dowker dissimilarities. A morphism $C \colon \Lambda \to \Lambda'$ of filtered relations is a *morphism of Dowker dissimilarities* if there exists a relation $C \subseteq L \times L'$ so that $C_t = C \colon \Lambda_t \to \Lambda'_t$ for every $t \in [0, \infty)$.

Definition 32 The 2-*category* Dow *of Dowker dissimilarities* is the 2-category with Dowker dissimilarities as objects and morphisms of Dowker dissimilarities as morphisms. Given morphisms $C_1, C_2 \colon \Lambda \to \Lambda'$ of Dowker dissimilarities, we define the set of 2-cells $\alpha \colon C_1 \to C_2$ in Dow by letting $\mathrm{Dow}(C_1, C_2) = [[0, \infty), \mathcal{R}](C_1, C_2)$.

Definition 33 Let $\Lambda \colon L \times W \to [0, \infty]$ be a Dowker dissimilarity. The *Dowker Nerve* $N\Lambda$ of Λ is the filtered nerve of the underlying filtered relation.

Note that the Dowker Nerve is filtered by inclusion of sub-complexes, that is, if $s \leq t$, then $N\Lambda_{s \leq t} \colon N\Lambda_s \to N\Lambda_t$ is an inclusion of simplicial complexes.

Definition 34 Let $\Lambda \colon L \times W \to [0, \infty]$ be a Dowker dissimilarity. Given $l \in L$ and $t > 0$, the Λ-*ball of radius t centred at l* is

$$B_\Lambda(l, t) = \{ w \in W \mid \Lambda(l, w) < t \}.$$

Example 3 Let (M, d) be a metric space and L and W be subsets of M. Then the restriction $\Lambda \colon L \times W \to [0, \infty]$ of d to $L \times W$ is a Dowker dissimilarity. The Dowker Nerve of Λ is the composite

$$[0, \infty) \xrightarrow{\Lambda} \mathcal{R} \xrightarrow{N} \mathrm{Cx}$$

taking $t \in [0, \infty)$ to

$$\{ \text{finite } \sigma \subseteq L \mid \text{ there exists } w \in W \text{ with } d(l, w) < t \text{ for all } l \in \sigma \}.$$

If $L = W = M$, then the Λ-ball of radius t centred at l is the usual open ball in M of radius t centred at l and the Dowker Nerve of Λ is equal to the Čech complex $\check{\mathcal{C}}(M)$.

Lemma 7 *Let $\Lambda \colon L \times W \to [0, \infty]$ be a Dowker dissimilarity. Given $t > 0$, the nerve $N\Lambda_t$ is isomorphic to the Borsuk Nerve of the cover of the set*

$$\bigcup_{l \in L} B_\Lambda(l, t)$$

by balls $B_\Lambda(l, s)$ of radius $s \leq t$ centred at points in L.

It is well-known that, in the above situation, the geometric realization of $N\Lambda_t$ is homotopic to the geometric realization of the Borsuk Nerve of the cover of the set

$$\bigcup_{l \in L} B_\Lambda(l, t)$$

by balls $B_\Lambda(l, t)$ of radius t centred at points in L. See e.g. Blaser and Brun (2018), Lemma 2.3. Corollary 3 gives:

Corollary 4 *If $\Lambda \colon L \times W \to [0, \infty]$ and $\Lambda' \colon L' \times W' \to [0, \infty]$ are (α, α')-interleaved Dowker dissimilarities, then $N\Lambda$ and $N\Lambda'$ are (α, α')-interleaved filtered simplicial complexes.*

Definition 35 Let $\Lambda \colon L \times W \to [0, \infty]$ be a Dowker dissimilarity. The *Rips complex* of Λ is the filtered simplicial complex $R\Lambda$ defined by

$$(R\Lambda)(t) = \{\text{finite } \sigma \subseteq L \mid \text{every } \tau \subseteq \sigma \text{ with } |\tau| \leq 2 \text{ is in } (N\Lambda)(t)\}.$$

Corollary 5 *If $\Lambda \colon L \times W \to [0, \infty]$ and $\Lambda' \colon L' \times W' \to [0, \infty]$ are (α, α')-interleaved Dowker dissimilarities, then $R\Lambda$ and $R\Lambda'$ are (α, α')-interleaved filtered simplicial complexes.*

Proof Use Corollary 4 and the fact that the Rips complex depends functorially on the one skeleton of the Dowker Nerve. □

11 Stability and interleaving distance

The functoriality of interleaving implies that all functorial constructions are stable with respect to interleaving. In this section we relate interleaving distance of Dowker dissimilarities to Gromov–Hausdorff distance of Edwards (1975), Gromov (1981) and to the network distance defined in Chowdhury and Mémoli (2018).

Definition 36 Let C and C' be filtered objects in a 2-category $\mathcal{C}$.

1. Given $a, a' \in [0, \infty]$ we say that the filtered objects C and C' are *additively* (a, a')-*interleaved* if they are (α, α')-interleaved for the functions $\alpha(t) = a + t$ and $\alpha'(t) = a' + t$.
2. The *interleaving distance* of C and C' is

$$d_{\text{int}}(C, C') = \inf\{a \in [0, \infty] \mid C \text{ and } C' \text{ are additively}$$
$$(a, a) - \text{interleaved}\},$$

with the convention that $\inf \emptyset = \infty$.

Definition 37 A *non-negatively weighted network* is a pair (X, ω_X) of a set X and a weight function $\omega_X \colon X \times X \to [0, \infty)$.

Definition 38 Let $\omega_X \colon X \times X \to [0, \infty)$ and $\omega_{X'} \colon X' \times X' \to [0, \infty)$ be non-negatively weighted networks and let $C \subseteq X \times X'$. The *distortion of C* is

$$\operatorname{dis}(C) = \sup_{(x,x'),\ (y,y')\in C} |\omega_X(x, y) - \omega_{X'}(x', y')|.$$

Recall from Definition 21 that $C \subseteq X \times X'$ is a correspondence if the projections of C on both X and X' are surjective.

Definition 39 Let $\omega_X \colon X \times X \to [0, \infty)$ and $\omega_{X'} \colon X' \times X' \to [0, \infty)$ be non-negatively weighted networks and let $\mathcal{R}$ be the set of correspondences $C \subseteq X \times X'$. The *network distance* between X and X' is

$$d_{\mathcal{N}}(X, X') = \frac{1}{2} \inf_{C \in \mathcal{R}} \operatorname{dis}(C).$$

The Stability Theorem (Chowdhury and Mémoli 2018, Proposition 15) for networks is a consequence of functoriality of interleaving distance, the Algebraic Stability Theorem for bottleneck distance (Chazal et al. 2009, Theorem 4.4) and the following result:

Proposition 4 *Let $\omega_X \colon X \times X \to [0, \infty)$ and $\omega_{X'} \colon X' \times X' \to [0, \infty)$ be networks, and write*

$$\Lambda \colon X \times X \to [0, \infty] \quad and \quad \Lambda' \colon X' \times X' \to [0, \infty]$$

for the corresponding Dowker dissimilarities with $\Lambda(x, y) = \omega_X(x, y)$ and $\Lambda'(x', y') = \omega_{X'}(x', y')$. Then

$$d_{\operatorname{int}}(\Lambda, \Lambda') \leq 2 \, d_{\mathcal{N}}(X, X').$$

Proof We have to show that $d_{\operatorname{int}}(\Lambda, \Lambda') \leq \operatorname{dis}(C)$ for every correspondence $C \subseteq X \times X'$. So let $C \subseteq X \times X'$ be a correspondence and let $a > \operatorname{dis}(C)$. By definition of $\operatorname{dis}(C)$, for all (l, l') and (w, w') in C we have

$$|\omega_X(l, w) - \omega_{X'}(l', w')| < a.$$

Defining $\alpha \colon [0, \infty) \to [0, \infty)$ by $\alpha(t) = t + a$, by symmetry, it suffices to show that C defines a morphism

$$C \colon \Lambda \to \alpha^* \Lambda'.$$

That is, we have to show that if $\sigma \in \Lambda_t$, then $(NC)(\sigma) \in \Lambda'_{\alpha(t)}$. So suppose that $w \in X$ satisfies $\Lambda(l, w) < t$ for all $l \in \sigma$. Since C is a correspondence we can pick $w' \in X'$ so that $(w, w') \in C$. By definition of NC, for every $l' \in (NC)(\sigma)$, there exists $l \in \sigma$ so that $(l, l') \in C$. By definition of distortion distance this gives

$$\Lambda'(l', w') = \omega_{X'}(l', w') < a + \omega_X(l, w) = a + \Lambda(l, w) < a + t = \alpha(t).$$

We conclude that $\sigma \in N\Lambda_t$ implies $(NC)(\sigma) \in N\Lambda'_{\alpha(t)}$ as desired. □

The Stability Theorem (Chazal et al. 2014, Theorem 5.2) for metric spaces is a consequence of functoriality of interleaving distance, the Algebraic Stability Theorem for bottleneck distance (Chazal et al. 2009, Theorem 4.4) and the following result:

Corollary 6 *Let (M, d) and (M', d') be metric spaces, and write*

$$\Lambda \colon M \times M \to [0, \infty] \quad and \quad \Lambda' \colon M' \times M' \to [0, \infty]$$

for the corresponding Dowker dissimilarities with $\Lambda(p, q) = d(p, q)$ and $\Lambda'(p', q') = d'(p', q')$. Then

$$\mathrm{d}_{\mathrm{int}}(\Lambda, \Lambda') \leq 2d_{GH}(M, M').$$

Proof By Burago et al. (2001), Theorem 7.3.25 the Gromov–Hausdorff distance of the metric spaces (M, d) and (M', d') agrees with their network distance when we consider them as non-negatively weighted networks. That is, $d_{GH}(M, M') = \mathrm{d}_{\mathcal{N}}(M, M')$. The result now follows from Proposition 4. □

12 Conclusion

We generalize the Sparse Čech construction of Cavanna et al. (2015) to arbitrary metric spaces and to a large class of Dowker dissimilarities. The abstract context of Dowker dissimilarities is well-suited for sparse nerve constructions. The concepts of filtered relations and strict 2-categories enable us to easily formulate and prove basic stability results. An implementation of the Sparse Dowker Nerve most similar to the Sparse Čech complex is available at GitHub Brun and Blaser (2018) together with a user tutorial. This implementation is not practical for analysis of high dimensional data.

Other recent work on sparse complexes has focused on Euclidean space (Choudhary et al. 2018) or the use of general simplicial maps (Dey et al. 2016; Boissonnat et al. 2018). It is possible that these considerations can be combined to obtain even smaller filtered nerves. In further work, we will improve this construction and we will make Sparse Dowker Nerves for Dowker dissimilarities that do not satisfy the triangle inequality.

References

Blumberg, A.J., Lesnick, M.: Universality of the homotopy interleaving distance (2017). https://arxiv.org/abs/1705.01690

Boissonnat, J.-D., Pritam, S., Pareek, D. Strong Collapse for Persistence (2018). ArXiv e-prints. arXiv:1809.10945

Blaser, N., Brun, M.: Divisive cover. Math. Comput. Sci. (2018). https://doi.org/10.1007/s11786-018-0352-6

Brun, M., Blaser, N.: Sparse-Dowker-nerves (2018). https://github.com/mbr085/Sparse-Dowker-Nerves

Bubenik, P., de Silva, V., Scott, J.: Metrics for generalized persistence modules. Found. Comput. Math. **15**(6), 1501–1531 (2015)

Buchet, M., Chazal, F., Oudot, S.Y., Sheehy, D.: Efficient and robust persistent homology for measures. Comput. Geom.: Theory Appl. **58**, 70–96 (2016)

Burago, D., Burago, Y., Ivanov, S.: A Course in Metric Geometry. Graduate Studies in Mathematics, vol. 33. American Mathematical Society, Providence (2001)

Carlsson, G., Ishkhanov, T., de Silva, V., Zomorodian, A.: On the local behavior of spaces of natural images. Int. J. Comput. Vis. **76**(1), 1–12 (2008)

Cavanna, N.J., Jahanseir, M., Sheehy, D.R.: A geometric perspective on sparse filtrations. *CoRR*, arXiv:1506.03797 (2015)

Chazal, F., Cohen-Steiner, D., Glisse, M., Guibas, L., Oudot, S.: Proximity of persistence modules and their diagrams. In: SoCG, pp. 7–246 (2009)

Chazal, F., Cohen-Steiner, D., Guibas, L.J., Mémoli, F., Oudot, S.Y.: Gromov–hausdorff stable signatures for shapes using persistence. In: Proceedings of the Symposium on Geometry Processing, SGP '09, pp, 1393–1403. Eurographics Association, Aire-la-Ville (2009)

Chazal, F., de Silva, V., Oudot, S.: Persistence stability for geometric complexes. Geom. Dedicata **173**, 193–214 (2014)

Choudhary, A., Kerber, M., Raghvendra, S.: Improved topological approximations by digitization. CoRR, arXiv:1812.04966 (2018)

Chowdhury, S., Mémoli, F.: A functorial Dowker theorem and persistent homology of asymmetric networks (2018). ArXiv e-prints. arXiv:1608.05432

Dey, T.K., Shi, D., Wang, Y.: SimBa: an efficient tool for approximating Rips-filtration persistence via *sim*plicial *ba*tch-collapse. In: 24th Annual European Symposium on Algorithms, vol. 57 of LIPIcs. Leibniz Int. Proc. Inform, Art. No. 35, 16 (2016)

Dowker, C.H.: Homology groups of relations. Ann. Math. **2**(56), 84–95 (1952)

Edwards, D.A.: The structure of superspace. In: Studies in Topology. Proceeding Conference, University of North Carolina, Charlotte, N. C., 1974; dedicated to Mathematics Section Polish Academy of Sciences, pp. 121–133. Academic Press, New York (1975)

Embrechts, P., Hofert, M.: A note on generalized inverses. Math. Methods Oper. Res. **77**(3), 423–432 (2013)

Gromov, M.: Groups of polynomial growth and expanding maps. Inst. Hautes Études Sci. Publ. Math. **53**, 53–73 (1981)

Lawvere, F.W.: Metric spaces, generalized logic, and closed categories. Ann. Math. (2) **65**, 357–362 (1957)

Leinster, T.: Basic bicategories (1998)

Milnor, J.: The geometric realization of a semi-simplicial complex. Inst. Hautes Études Sci. Publ. Math. **53**, 53–73 (1981)

Oudot, S.Y.: Persistence Theory: From Quiver Representations to Data Analysis. Mathematical Surveys and Monographs, vol. 209. American Mathematical Society, Providence (2015)

Sheehy, D.R.: Linear-size approximations to the Vietoris–Rips filtration. Discrete Comput. Geom. **49**(4), 778–796 (2013)

Spanier, E.H.: Algebraic Topology. Springer, New York (1966). Corrected reprint of the 1966 original

Zomorodian, A.: Topological data analysis. In Advances in Applied and Computational Topology. Proceedings of Symposia in Applied Mathematics, vol. 70, pp. 1–39. American Mathematical Society, Providence (2012)