Supplementary Material of The decomposition of the higher-order homology embedding constructed from the k-Laplacian

Table of Contents

A Proof of subspace perturbations (Theorem 1)

A.1 A formal version of Assumption 3

Assumption S1. *Let* $\tilde{\mathbf{w}}_k = |\mathbf{B}_{k+1}[\mathfrak{N}_k, \mathfrak{N}_{k+1}]|\mathbf{w}_{k+1}$, $\tilde{\mathbf{w}}_{k-1} = |\mathbf{B}_k[:, \mathfrak{N}_k]|\tilde{\mathbf{w}}_k$, *with* $\mathbf{w}_k$ *and* $\widehat{\mathbf{w}}_k$ *defined in Section 3. Additionally, write*

$$
\begin{aligned}
\mathbf{W}_{k+1} &= \tilde{\mathbf{W}}_{k+1} + \boldsymbol{\mathcal{E}}_{k+1,+},\\
\widehat{\mathbf{W}}_{k+1} &= \tilde{\mathbf{W}}_{k+1} + \boldsymbol{\mathcal{E}}_{k+1,-},\\
\mathbf{W}_k^{1/2} &= \tilde{\mathbf{W}}_k^{1/2}(\mathbf{I} + \mathbf{E}_{k,+}^{+}) + \boldsymbol{\mathcal{E}}_{k,+}^{1/2},\\
\mathbf{W}_k^{-1/2} &= \tilde{\mathbf{W}}_k^{-1/2}(\mathbf{I} - \mathbf{E}_{k,+}^{-}) + \boldsymbol{\mathcal{E}}_{k,+}^{-1/2},\\
\widehat{\mathbf{W}}_k^{1/2} &= \tilde{\mathbf{W}}_k^{1/2}(\mathbf{I} + \mathbf{E}_{k,-}^{+}) + \boldsymbol{\mathcal{E}}_{k,-}^{1/2},\\
\widehat{\mathbf{W}}_k^{-1/2} &= \tilde{\mathbf{W}}_k^{-1/2}(\mathbf{I} - \mathbf{E}_{k,-}^{-}) + \boldsymbol{\mathcal{E}}_{k,-}^{-1/2},\\
\mathbf{W}_{k-1}^{-1/2} &= \tilde{\mathbf{W}}_{k-1}^{-1/2}(\mathbf{I} - \mathbf{E}_{k-1,+}),\\
\widehat{\mathbf{W}}_{k-1}^{-1/2} &= \tilde{\mathbf{W}}_{k-1}^{-1/2}(\mathbf{I} - \mathbf{E}_{k-1,-}).
\end{aligned}
$$

There exists $\epsilon_\ell > 0$ *and* $\epsilon'_\ell > 0$ *for* $\ell = k, k-1$ *such that the following conditions hold*

1. *Not too many* $(k+1)$*-simplices are created (small* $|\mathfrak{C}_{k+1}|$*)*

$$
\|\mathbf{E}_{k,+}^{+}\| = \max_{\sigma\in\mathfrak{N}_k}\left\{\left[\mathbf{E}_{k,+}^{+}\right]_{\sigma,\sigma}\right\} = \max_{\sigma\in\mathfrak{N}_k}\left\{\frac{w_k^{1/2}(\sigma)}{\tilde{w}_k^{1/2}(\sigma)} - 1\right\} \le \sqrt{\epsilon_k}; \tag{S1a}
$$

$$
\|\mathbf{E}_{k,+}^{-}\| = \max_{\sigma\in\mathfrak{N}_k}\left\{\left[\mathbf{E}_{k,+}^{-}\right]_{\sigma,\sigma}\right\} = \max_{\sigma\in\mathfrak{N}_k}\left\{\frac{\tilde{w}_k^{-1/2}(\sigma)}{w_k^{-1/2}(\sigma)} - 1\right\} \le \sqrt{\epsilon_k}; \tag{S1b}
$$

$$
\max_{\sigma\in\mathfrak{N}_k}\left\{\frac{w_k(\sigma)}{\tilde{w}_k(\sigma)} - 1\right\} \le \epsilon_k; \tag{S1c}
$$

$$
\|\mathbf{E}_{k-1,+}\| = \max_{\nu\in\Sigma_{k-1}}\left\{[\mathbf{E}_{k-1,+}]_{\nu,\nu}\right\} = \max_{\nu\in\Sigma_{k-1}}\left\{\frac{\tilde{w}_{k-1}^{-1}(\nu)}{w_{k-1}^{-1}(\nu)} - 1\right\} \le \sqrt{\epsilon_{k-1}}; \tag{S1d}
$$

$$
\max_{\nu\in\Sigma_{k-1}}\left\{\frac{w_{k-1}(\nu)}{\tilde{w}_{k-1}(\nu)} - 1\right\} \le \epsilon_{k-1}. \tag{S1e}
$$

2. *Not too many* $(k+1)$*-simplices are destroyed (small* $|\mathfrak{D}_{k+1}|$*)*

$$
\|\mathbf{E}_{k,-}^{+}\| = \max_{\sigma\in\mathfrak{N}_k}\left\{\left[\mathbf{E}_{k,-}^{+}\right]_{\sigma,\sigma}\right\} = \max_{\sigma\in\mathfrak{N}_k}\left\{\frac{\widehat{w}_k^{1/2}(\sigma)}{\tilde{w}_k^{1/2}(\sigma)} - 1\right\} \le \sqrt{\epsilon_k}; \tag{S2a}
$$

$$
\|\mathbf{E}_{k,-}^{-}\| = \max_{\sigma\in\mathfrak{N}_k}\left\{\left[\mathbf{E}_{k,-}^{-}\right]_{\sigma,\sigma}\right\} = \max_{\sigma\in\mathfrak{N}_k}\left\{\frac{\tilde{w}_k^{-1/2}(\sigma)}{\widehat{w}_k^{-1/2}(\sigma)} - 1\right\} \le \sqrt{\epsilon_k}; \tag{S2b}
$$

$$
\max_{\sigma\in\mathfrak{N}_k}\left\{\frac{\widehat{w}_k(\sigma)}{\tilde{w}_k(\sigma)} - 1\right\} \le \epsilon_k; \tag{S2c}
$$

$$
\|\mathbf{E}_{k-1,-}\| = \max_{\nu\in\Sigma_{k-1}}\left\{[\mathbf{E}_{k-1,-}]_{\nu,\nu}\right\} = \max_{\nu\in\Sigma_{k-1}}\left\{\frac{\tilde{w}_{k-1}^{-1}(\nu)}{\widehat{w}_{k-1}^{-1}(\nu)} - 1\right\} \le \sqrt{\epsilon_{k-1}}; \tag{S2d}
$$

$$
\max_{\nu\in\Sigma_{k-1}}\left\{\frac{\widehat{w}_{k-1}(\nu)}{\tilde{w}_{k-1}(\nu)} - 1\right\} \le \epsilon_{k-1}. \tag{S2e}
$$

3. The net *changes on* $\mathbf{w}_k$ *and* $\mathbf{w}_{k-1}$ *are small*

$$\|\mathbf{E}^+_{k,+} - \mathbf{E}^+_{k,-}\| = \max_{\sigma\in\mathfrak{N}_k}\left\{\left|\frac{\widehat{w}_k^{1/2}(\sigma)}{w_k^{1/2}(\sigma)} - 1\right|\right\} \le \sqrt{\epsilon'_k}; \tag{S3a}$$

$$\|\mathbf{E}^-_{k,+} - \mathbf{E}^-_{k,-}\| = \max_{\sigma\in\mathfrak{N}_k}\left\{\left|\frac{w_k^{1/2}(\sigma)}{\widehat{w}_k^{1/2}(\sigma)} - 1\right|\right\} \le \sqrt{\epsilon'_k}; \tag{S3b}$$

$$\max_{\sigma\in\mathfrak{N}_k}\left\{\left|\frac{\widehat{w}_k(\sigma)}{w_k(\sigma)} - 1\right|\right\} \le \epsilon'_k; \tag{S3c}$$

$$\|\mathbf{E}_{k-1,+} - \mathbf{E}_{k-1,-}\| = \max_{\nu\in\Sigma_{k-1}}\left\{\left|\frac{w_{k-1}^{-1}(\nu)}{\widehat{w}_{k-1}^{-1}(\nu)} - 1\right|\right\} \le \sqrt{\epsilon'_{k-1}}; \tag{S3d}$$

$$\max_{\nu\in\Sigma_{k-1}}\left\{\left|\frac{\widehat{w}_{k-1}(\nu)}{w_{k-1}(\nu)} - 1\right|\right\} \le \epsilon'_{k-1}. \tag{S3e}$$

A.2 Definitions of $\mathcal{L}_k$ and $\widehat{\mathcal{L}}_k$

Given a manifold $\mathcal{M}$ which is constructed by a series of connected sum, i.e., $\mathcal{M} = \mathcal{M}_1\sharp\cdots\sharp\mathcal{M}_\kappa$. Let the Simplicial complex corresponding to $\mathcal{M}$ be $\mathrm{SC}_\ell = (\Sigma_0, \cdots, \Sigma_\ell)$, with the disjoint simplicial complex (of $\cup_{i=1}^\kappa \mathcal{M}_i$) being $\widehat{\mathrm{SC}}_\ell = (\widehat{\Sigma}_0, \cdots, \widehat{\Sigma}_\ell)$. For each k, the simplex sets can be decomposed into the following

$$\Sigma_k = \underbrace{\bigcup_{i=1}^{\kappa}\Sigma_k^{(i)}}_{\text{non-intersecting set:}\mathfrak{N}_k} \cup \underbrace{\bigcup_{j>i}^{\kappa}\Sigma_k^{(ij)+}}_{\text{created set:}\mathfrak{C}_k}.$$

Similarly,

$$\widehat{\Sigma}_k = \underbrace{\bigcup_{i=1}^{\kappa}\Sigma_k^{(i)}}_{\text{non-intersecting set:}\mathfrak{N}_k} \cup \underbrace{\bigcup_{j>i}^{\kappa}\Sigma_k^{(ij)-}}_{\text{destroyed set:}\mathfrak{D}_k}.$$

W.l.o.g., one can assume that the $(k-1)$-simplices set can be perfectly separated, i.e., $\mathfrak{C}_{k-1} = \mathfrak{D}_{k-1} = \emptyset$ (when analyzing the k-Laplacian). The above construction matches our intuition; by definition, a connected sum is a process of carving out a d-disk ($\mathfrak{D}_k$) and gluing two manifolds together ($\mathfrak{C}_k$).

We are interested in the perturbation of the k-Laplacian $\boldsymbol{\mathcal{L}}_k$ w.r.t. the ideal (disjoint) Laplacian $\widehat{\boldsymbol{\mathcal{L}}}_k$. Without carefully define both $\boldsymbol{\mathcal{L}}_k$ and $\widehat{\boldsymbol{\mathcal{L}}}_k$, the perturbation on the subspaces might be unbounded. With slight abuse of notation, we let $\mathbf{L} \leftarrow \boldsymbol{\mathcal{L}}_k$, $\mathbf{L}_d \leftarrow \boldsymbol{\mathcal{L}}_k^{\mathrm{down}}$, and $\mathbf{L}_u \leftarrow \boldsymbol{\mathcal{L}}_k^{\mathrm{up}}$ (similar definitions for $\widehat{\boldsymbol{\mathcal{L}}}$'s). The k is omitted and can be inferred from the context. The $\widehat{\mathbf{L}}$ and $\mathbf{L}$ are defined as follows. $\widehat{\mathbf{L}}$ is a block diagonal matrix, with the i-th (diagonal) block $\mathbf{L}^{(i)}$ described by $\mathcal{M}_i$ constructed from the sub-complex $\widehat{\mathrm{SC}}^{(i)}$ ($\widehat{\Sigma}_{k-1}^{(i)}$, $\widehat{\Sigma}_k^{(i)}$, and $\widehat{\Sigma}_{k+1}^{(i)}$). Due to manifolds being disjoint (i.e., $\cup_{i=1}^\kappa \mathcal{M}_i$), the Laplacian corresponding to such block, denoted $\widehat{\mathbf{L}}^{(i,i),(i,i)}$, will be a valid Laplacian. As for the intersecting k-simplices $\mathfrak{C}_k \cup \mathfrak{D}_k$, we let $\widehat{\mathbf{L}}^{(i,j),(k,l)} = \mathbf{L}^{(i,j),(k,l)}$ for all $ij, kl \in \binom{[k]}{2}$ so that the corresponding blocks of $\widehat{\mathbf{L}} - \mathbf{L}$ will be zero. Under this scenario, the unbounded increase of $(k+1)$-simplices caused by the intersecting k-simplices can be removed. Lastly, the off-diagonal blocks of $\widehat{\mathbf{L}}$ are set to zero. Specifically, $\widehat{\mathbf{L}}$ is,

$$\widehat{\mathbf{L}} = \left[\begin{array}{ccc|ccc|ccc} \widehat{\mathbf{L}}^{(1,1),(1,1)} & & & & & & \widehat{\mathbf{L}}^{(1,1),(1,2)-} & \cdots & \widehat{\mathbf{L}}^{(1,1),(k-1,k)-} \\ & \ddots & & & \mathbf{0} & & \vdots & \ddots & \vdots \\ & & \widehat{\mathbf{L}}^{(k,k),(k,k)} & & & & \widehat{\mathbf{L}}^{(k,k),(1,2)-} & \cdots & \widehat{\mathbf{L}}^{(k,k),(k-1,k)-} \\ \hline & & & \mathbf{L}^{(1,2)+,(1,2)+} & \cdots & \mathbf{L}^{(1,2)+,(k-1,k)+} & & & \\ & \mathbf{0} & & \vdots & \ddots & \vdots & & \mathbf{0} & \\ & & & \mathbf{L}^{(k-1,k)+,(1,2)+} & \cdots & \mathbf{L}^{(k-1,k)+,(k-1,k)+} & & & \\ \hline \widehat{\mathbf{L}}^{(1,2)-,(1,1)} & \cdots & \widehat{\mathbf{L}}^{(1,2)-,(k,k)} & & & & \widehat{\mathbf{L}}^{(1,2)-,(1,2)-} & \cdots & \widehat{\mathbf{L}}^{(1,2)-,(k-1,k)-} \\ \vdots & \ddots & \vdots & & \mathbf{0} & & \vdots & \ddots & \vdots \\ \widehat{\mathbf{L}}^{(k-1,k)-,(1,1)} & \cdots & \widehat{\mathbf{L}}^{(k,k)-,(k,k)} & & & & \widehat{\mathbf{L}}^{(k-1,k)-,(1,2)-} & \cdots & \widehat{\mathbf{L}}^{(k-1,k)-,(k-1,k)-} \end{array}\right].$$

Similarly, one can define $\mathbf{L}$ to be

$$\mathbf{L} = \left[\begin{array}{ccc|ccc|ccc} \mathbf{L}^{(1,1),(1,1)} & & & \mathbf{L}^{(1,1),(1,2)+} & \cdots & \mathbf{L}^{(1,1),(k-1,k)+} & & & \\ & \ddots & & \vdots & \ddots & \vdots & & \mathbf{0} & \\ & & \mathbf{L}^{(k,k),(k,k)} & \mathbf{L}^{(k,k),(1,2)+} & \cdots & \mathbf{L}^{(k,k),(k-1,k)+} & & & \\ \hline \mathbf{L}^{(1,2)+,(1,1)} & \cdots & \mathbf{L}^{(1,2)+,(k,k)} & \mathbf{L}^{(1,2)+,(1,2)+} & \cdots & \mathbf{L}^{(1,2)+,(k-1,k)+} & & & \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & & \mathbf{0} & \\ \mathbf{L}^{(k-1,k)+,(1,1)} & \cdots & \mathbf{L}^{(k,k)+,(k,k)} & \mathbf{L}^{(k-1,k)+,(1,2)+} & \cdots & \mathbf{L}^{(k-1,k)+,(k-1,k)+} & & & \\ \hline & & & & & & \widehat{\mathbf{L}}^{(1,2)-,(1,2)-} & \cdots & \widehat{\mathbf{L}}^{(1,2)-,(k-1,k)-} \\ & \mathbf{0} & & & \mathbf{0} & & \vdots & \ddots & \vdots \\ & & & & & & \widehat{\mathbf{L}}^{(k-1,k)-,(1,2)-} & \cdots & \widehat{\mathbf{L}}^{(k-1,k)-,(k-1,k)-} \end{array}\right].$$

Under this construction, the four lower right blocks, which correspond to the k-simplices in $\mathfrak{C}_k \cup \mathfrak{D}_k$, will be zero. If no new homology class is created/destroyed (Assumption 1) and the minimum eigenvalues of the last two diagonal blocks are bounded away from zero (Assumption 2), then the eigengap of $\mathbf{L}$ will simply be the minimum eigengap of each $\widehat{\mathbf{L}}^{(i)}$, i.e., $\text{eigengap}(\mathbf{L}) = \min\{\delta_1, \cdots, \delta_\kappa\}$.

Now we formally define our formulation. Following the notations introduced in Section 3, and let $\mathcal{I}_\sigma$ be the index set of the k-simplex $\sigma \in \mathfrak{N}_k$ sampled from $\mathcal{M}_i$. Note that $\mathcal{I}_\sigma$ is defined only for $\sigma \in \mathfrak{N}_k$, which can be extended from the index set $\mathcal{I}_v$ for $v \in V$ introduced in Section 3 by $\mathcal{I}_\sigma = \{\sigma \in \mathfrak{N}_k : v \in \mathcal{I}_v \text{ for } v \in \sigma\}$. Note also that similar to $\mathcal{I}_v$ for V, $\mathcal{S}_\sigma$ can be larger than 1. For instance, if the manifold is constructed by gluing a torus (indexed by 1) and a circle (indexed by 2), then $\mathcal{S}_1 = \{1, 2\}$ and $\mathcal{S}_2 = \{3\}$; for an edge e belongs to the torus, we have $\mathcal{S}_{\mathcal{I}_e} = \{1, 2\}$. For every $\sigma \in \mathfrak{N}_k$, we write,

$$\sum_{\sigma \in \mathfrak{N}_k} \sum_{i \notin S_{\mathcal{I}_\sigma}} \mathbf{Y}_{\sigma,i}^2 \leq \sum_{\sigma \in \mathfrak{N}_k} \sum_{i=1}^{\beta_1} (\mathbf{Y}_{\sigma,i} - \widehat{\mathbf{Y}}_{\sigma,i})^2 \leq \sum_{\sigma \in \Sigma_k \cup \widehat{\Sigma}_k} \sum_{i=1}^{\beta_1} (\mathbf{Y}_{\sigma,i} - \widehat{\mathbf{Y}}_{\sigma,i})^2 = \|\mathbf{YO} - \widehat{\mathbf{Y}}\|_F^2.$$

Let $\mathsf{DiffL}_k^{\text{down}} = \mathbf{L}_d - \widehat{\mathbf{L}}_d$ and $\mathsf{DiffL}_k^{\text{up}} = \mathbf{L}_u - \widehat{\mathbf{L}}_u$, from [61] and the triangular inequality,

$$\begin{aligned}
\left\|\mathbf{Y}_{\mathfrak{N}_k,:} - \widehat{\mathbf{Y}}_{\mathfrak{N}_k,:}\right\|_F^2 &= \sum_{\sigma\in\mathfrak{N}_k}\sum_{i\notin\mathcal{S}_{\mathcal{I}_\sigma}} \mathbf{Y}_{\sigma,i}^2 \le \|\mathbf{Y} - \widehat{\mathbf{Y}}\mathbf{O}\|_F^2 \\
&\le \frac{8\cdot\min\left\{\beta_k\left\|\mathbf{L}-\widehat{\mathbf{L}}\right\|^2, \|\mathbf{L}-\widehat{\mathbf{L}}\|_F^2\right\}}{\min\{\delta_1,\cdots,\delta_\kappa\}} \\
&\le \frac{8\cdot\min\left\{\beta_k\left\|\mathsf{DiffL}_k^{\text{down}}\right\|^2 + \beta_k\left\|\mathsf{DiffL}_k^{\text{up}}\right\|^2, \|\mathsf{DiffL}_k^{\text{down}}\|_F^2 + \|\mathsf{DiffL}_k^{\text{up}}\|_F^2\right\}}{\min\{\delta_1,\cdots,\delta_\kappa\}} \\
&\overset{\dagger}{\le} \frac{8\beta_k\left(\left\|\mathsf{DiffL}_k^{\text{down}}\right\|^2 + \left\|\mathsf{DiffL}_k^{\text{up}}\right\|^2\right)}{\min\{\delta_1,\cdots,\delta_\kappa\}}.
\end{aligned}$$

Remark. The bound w.r.t. the Frobenius norm is omitted (the last inequality †) based on two reasons: (i) $\boldsymbol{\mathcal{L}}_k$ has complicated forms for large k, therefore, it is hard to derive a concise expression; and (ii) $\|\cdot\|_F$ is usually larger than $\beta_k\|\cdot\|$.

A.3 Useful lemmas

Here we omit the k for $\mathfrak{N}$, $\mathfrak{C}$, and $\mathfrak{D}$ for simplicity. Let $\lambda_k = \|\boldsymbol{\mathcal{L}}_k\|$ be the bound on the spectral norm of k-Laplacian. Here, $\lambda_k = k+2$ for $\boldsymbol{\mathcal{L}}$'s built from simplicial complexes; $\lambda_k = 2k+2$ for those built from cubical complexes (see also Proposition S3). The following two lemmas bound the effects of $\boldsymbol{\mathcal{E}}_{k,+}$, $\boldsymbol{\mathcal{E}}_{k,-}$, $\boldsymbol{\mathcal{E}}_{k+1,+}$, and $\boldsymbol{\mathcal{E}}_{k+1,-}$ in their changes to the weights ($\mathbf{W}_k$ and $\mathbf{W}_{k-1}$) of the k and $(k-1)$-simplices; we will find them useful in proving Theorem 1.

Lemma S1. *Let $\mathbf{W}_k$, $\widehat{\mathbf{W}}_k$, $\boldsymbol{\mathcal{E}}_{k,+}$, and $\boldsymbol{\mathcal{E}}_{k,-}$ defined in Assumption S1, we have*

$$\begin{aligned}
\left\|\boldsymbol{\mathcal{E}}_{k,+}\mathbf{B}_k^\top\mathbf{W}_{k-1}^{-1}\mathbf{B}_k\boldsymbol{\mathcal{E}}_{k,+}\right\| &\le \lambda_{k-1}\epsilon_{k-1}, \\
\left\|\boldsymbol{\mathcal{E}}_{k,-}\mathbf{B}_k^\top\widehat{\mathbf{W}}_{k-1}^{-1}\mathbf{B}_k\boldsymbol{\mathcal{E}}_{k,-}\right\| &\le \lambda_{k-1}\epsilon_{k-1}.
\end{aligned}$$

Proof. We first inspect the case of $\mathfrak{C}$, i.e., the first equation involving $\boldsymbol{\mathcal{E}}_{k,+}$,

$$[\boldsymbol{\mathcal{E}}_{k,+}]_{\sigma,\sigma} = \begin{cases} w_k^{1/2}(\sigma) & \text{if } \sigma\in\mathfrak{C}; \\ 0 & \text{otherwise.}\end{cases}$$

for any $\nu\in\Sigma_{k-1}$, we have,

$$\begin{aligned}
w_{k-1}(\nu) &= |\mathbf{B}_k(\nu)|\mathbf{w}_k; \\
\tilde{w}_{k-1}(\nu) &= |\mathbf{B}_k(\nu)|\tilde{\mathbf{w}}_k.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\epsilon_{k-1}w_{k-1}(\nu) &\ge \epsilon_{k-1}\tilde{w}_{k-1}(\nu) \ge w_{k-1}(\nu) - \tilde{w}_{k-1}(\nu) = |\mathbf{B}_k(\nu)|(\mathbf{w}_k - \tilde{\mathbf{w}}_k) \\
&= |\mathbf{B}_k(\nu)|\left[\tilde{\mathbf{w}}_k\mathbf{E}_k + \boldsymbol{\mathcal{E}}_{k,+}\right] \ge |\mathbf{B}_k(\nu)|\boldsymbol{\mathcal{E}}_{k,+} = \deg(\nu).
\end{aligned}$$

Let f_m be the k-eigencochain corresponding to the largest eigenvalue of $\boldsymbol{\mathcal{E}}_{k,+}\mathbf{B}_k^\top\mathbf{W}_{k-1}^{-1}\mathbf{B}_k\boldsymbol{\mathcal{E}}_{k,+}$. From Eq. (3.6) of [26], we have,

$$\begin{aligned}
\|\boldsymbol{\mathcal{E}}_{k,+}\mathbf{B}_k^\top\mathbf{W}_{k-1}^{-1}\mathbf{B}_k\boldsymbol{\mathcal{E}}_{k,+}\|_2 &\le \|\mathbf{L}_d\|\cdot\frac{\sum_{\nu\in\Sigma_{k-1}} f_m^2(\nu)\deg(\nu)}{\sum_{\nu\in\Sigma_{k-1}} f_m^2(\nu)w_{k-1}(\nu)} \\
&\le \lambda_{k-1}\epsilon_{k-1}\cdot\frac{\sum_{\nu\in\Sigma_{k-1}} f_m^2(\nu)w_{k-1}(\nu)}{\sum_{\nu\in\Sigma_{k-1}} f_m^2(\nu)w_{k-1}(\nu)} = \lambda_{k-1}\epsilon_{k-1}.
\end{aligned}$$

The case of $\mathfrak{D}$ follows similarly. ■

The following lemma bounds the changes in $(k+1)$-simplices with ϵ_k.

Lemma S2. *Let $\mathbf{W}$ be either $\mathbf{W}_k$ or $\widehat{\mathbf{W}}_k$, and $\boldsymbol{\mathcal{E}}$ be either $\boldsymbol{\mathcal{E}}_{k+1,+}$ or $\boldsymbol{\mathcal{E}}_{k+1,-}$ defined in Assumption S1, we have*

$$\left\|\mathbf{W}\mathbf{B}_{k+1}\boldsymbol{\mathcal{E}}\mathbf{B}_{k+1}^\top\mathbf{W}\right\| \leq \lambda_k\epsilon_k.$$

Proof. Consider the case of $\mathbf{W}_k$ and $\boldsymbol{\mathcal{E}}_{k+1,+}$. For any $\sigma \in \Sigma_k$,

$$\begin{aligned} w_k(\sigma) &= |\mathbf{B}_{k+1}(\sigma)|\mathbf{w}_{k+1}; \\ \tilde{w}_k(\sigma) &= |\mathbf{B}_{k+1}(\sigma)|\tilde{\mathbf{w}}_{k+1}. \end{aligned}$$

Therefore, for any $\sigma \in \mathfrak{N}$ (do not count the one in $\boldsymbol{\mathcal{E}}_{k,\pm}$) we have,

$$\epsilon_k w_k(\sigma) \geq \epsilon_k\tilde{w}_k(\sigma) \geq w_k(\sigma) - \tilde{w}(\sigma) = |\mathbf{B}_{k+1}(\sigma)|(\mathbf{w}_{k+1} - \tilde{\mathbf{w}}_{k+1}) = |\mathbf{B}_{k+1}(\sigma)|\boldsymbol{\mathcal{E}}_{k+1,+}.$$

Let f_m be the k-eigencochain corresponding to the largest eigenvalue of the matrix $\mathbf{W}_k^{-1/2}\mathbf{B}_{k+1}\boldsymbol{\mathcal{E}}_{k+1,+}\mathbf{B}_{k+1}^\top\mathbf{W}_k^{-1/2}$. From Eq. (3.6) of [26],

$$\begin{aligned} \left\|\mathbf{W}_k^{-1/2}\mathbf{B}_{k+1}\boldsymbol{\mathcal{E}}_{k+1,+}\mathbf{B}_{k+1}^\top\mathbf{W}_k^{-1/2}\right\| &\leq (k+2)\cdot\frac{\sum_{\sigma\in\mathfrak{N}} f_m^2(\sigma)\deg(\sigma)}{\sum_{\sigma\in\mathfrak{N}} f_m^2(\sigma)w_k(\sigma)} \\ &\leq \lambda_k\epsilon_k\frac{\sum_{\sigma\in\mathfrak{N}} f_m^2(e)w_k(e)}{\sum_{\sigma\in\mathfrak{N}} f_m^2(e)w_k(e)} = \lambda_k\epsilon_k \end{aligned}$$

Here $\deg(\sigma) = |\mathbf{B}_{k+1}(\sigma)\mathrm{diag}(\boldsymbol{\mathcal{E}}_{k+1,+})|$. Consider the case when $\mathbf{W} \leftarrow \widehat{\mathbf{W}}_k$ and $\boldsymbol{\mathcal{E}} \leftarrow \boldsymbol{\mathcal{E}}_{k+1,+}$, we have,

$$\epsilon_k\widehat{w}_k(\sigma) \geq \epsilon_k\tilde{w}_k(\sigma) \geq w_k(\sigma) - \tilde{w}(\sigma) = |\mathbf{B}_{k+1}(\sigma)|(\mathbf{w}_{k+1} - \tilde{\mathbf{w}}_{k+1}) = |\mathbf{B}_{k+1}(\sigma)|\boldsymbol{\mathcal{E}}_{k+1,+}.$$

The result follows similarly for $\boldsymbol{\mathcal{E}} \leftarrow \boldsymbol{\mathcal{E}}_{k+1,-}$; this completes the proof. ■

A.4 Proof of Theorem 1

Now we start the formal proof of Theorem 1. We will break the proof into two parts, i.e., the *down* and *up* parts involving $\mathsf{DiffL}_k^{\mathrm{down}}$ and $\mathsf{DiffL}_k^{\mathrm{up}}$, respectively.

Proof of the $\mathsf{DiffL}_k^{\mathrm{down}}$ term in Theorem 1. The explicit form of the *down* Laplacian can be written as

$$\widehat{\mathbf{L}}_d = \left[\begin{array}{c|c|c} \mathbf{M}_{\mathfrak{N}}\widehat{\mathbf{W}}_k^{1/2}\mathbf{B}_k^\top\widehat{\mathbf{W}}_{k-1}^{-1}\mathbf{B}_k\widehat{\mathbf{W}}_k^{1/2}\mathbf{M}_{\mathfrak{N}} & \mathbf{0} & \mathbf{M}_{\mathfrak{N}}\widehat{\mathbf{W}}_k^{1/2}\mathbf{B}_k^\top\widehat{\mathbf{W}}_{k-1}^{-1}\mathbf{B}_k\boldsymbol{\mathcal{E}}_{k,-}^{1/2}\mathbf{M}_{\mathfrak{D}} \\ \hline \mathbf{0} & \mathbf{M}_{\mathfrak{C}}\boldsymbol{\mathcal{E}}_{k,+}^{1/2}\mathbf{B}_k^\top\mathbf{W}_{k-1}^{-1}\mathbf{B}_k\boldsymbol{\mathcal{E}}_{k,+}^{1/2}\mathbf{M}_{\mathfrak{C}} & \mathbf{0} \\ \hline \mathbf{M}_{\mathfrak{D}}\boldsymbol{\mathcal{E}}_{k,-}^{1/2}\mathbf{B}_k^\top\widehat{\mathbf{W}}_{k-1}^{-1}\mathbf{B}_k\widehat{\mathbf{W}}_k^{1/2}\mathbf{M}_{\mathfrak{N}} & \mathbf{0} & \mathbf{M}_{\mathfrak{D}}\boldsymbol{\mathcal{E}}_{k,-}^{1/2}\mathbf{B}_k^\top\widehat{\mathbf{W}}_{k-1}^{-1}\mathbf{B}_k\boldsymbol{\mathcal{E}}_{k,-}^{1/2}\mathbf{M}_{\mathfrak{D}} \end{array}\right].$$

And,

$$\mathbf{L}_d = \left[\begin{array}{c|c|c} \mathbf{M}_{\mathfrak{N}}\mathbf{W}_k^{1/2}\mathbf{B}_k^\top\mathbf{W}_{k-1}^{-1}\mathbf{B}_k\mathbf{W}_k^{1/2}\mathbf{M}_{\mathfrak{N}} & \mathbf{M}_{\mathfrak{N}}\mathbf{W}_k^{1/2}\mathbf{B}_k^\top\mathbf{W}_{k-1}^{-1}\mathbf{B}_k\boldsymbol{\mathcal{E}}_{k,+}^{1/2}\mathbf{M}_{\mathfrak{C}} & \mathbf{0} \\ \hline \mathbf{M}_{\mathfrak{C}}\boldsymbol{\mathcal{E}}_{k,+}^{1/2}\mathbf{B}_k^\top\mathbf{W}_{k-1}^{-1}\mathbf{B}_k\mathbf{W}_k^{1/2}\mathbf{M}_{\mathfrak{N}} & \mathbf{M}_{\mathfrak{C}}\boldsymbol{\mathcal{E}}_{k,+}^{1/2}\mathbf{B}_k^\top\mathbf{W}_{k-1}^{-1}\mathbf{B}_k\boldsymbol{\mathcal{E}}_{k,+}^{1/2}\mathbf{M}_{\mathfrak{C}} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{M}_{\mathfrak{D}}\boldsymbol{\mathcal{E}}_{k,-}^{1/2}\mathbf{B}_k^\top\widehat{\mathbf{W}}_{k-1}^{-1}\mathbf{B}_k\boldsymbol{\mathcal{E}}_{k,-}^{1/2}\mathbf{M}_{\mathfrak{D}} \end{array}\right].$$

Here, $\mathbf{M}_{\mathfrak{N}}$, $\mathbf{M}_{\mathfrak{C}}$, and $\mathbf{M}_{\mathfrak{D}}$ are diagonal masks for k-simplex sets $\mathfrak{N}$, $\mathfrak{C}$, and $\mathfrak{D}$, respectively. By triangular inequality,

$$\left\|\mathbf{L}_d - \widehat{\mathbf{L}}_d\right\| \leq \underbrace{\left\|\mathbf{M}_{\mathfrak{N}}\mathbf{W}_k^{1/2}\mathbf{B}_k^\top\mathbf{W}_{k-1}^{-1}\mathbf{B}_k\mathbf{W}_k^{1/2}\mathbf{M}_{\mathfrak{N}} - \mathbf{M}_{\mathfrak{N}}\widehat{\mathbf{W}}_k^{1/2}\mathbf{B}_k^\top\widehat{\mathbf{W}}_{k-1}^{-1}\mathbf{B}_k\widehat{\mathbf{W}}_k^{1/2}\mathbf{M}_{\mathfrak{N}}\right\|}_{(*)} + 2\left[\underbrace{\left\|\mathbf{M}_{\mathfrak{N}}\mathbf{W}_k^{1/2}\mathbf{B}_k^\top\mathbf{W}_{k-1}^{-1}\mathbf{B}_k\boldsymbol{\mathcal{E}}_{k,+}^{1/2}\mathbf{M}_{\mathfrak{C}}\right\|}_{(\dagger)} + \left\|\mathbf{M}_{\mathfrak{N}}\widehat{\mathbf{W}}_k^{1/2}\mathbf{B}_k^\top\widehat{\mathbf{W}}_{k-1}^{-1}\mathbf{B}_k\boldsymbol{\mathcal{E}}_{k,-}^{1/2}\mathbf{M}_{\mathfrak{D}}\right\|\right]. \tag{S4}$$

Expand the $\mathbf{W}_k$ with $\widehat{\mathbf{W}}_k$ and omit $\mathbf{M}_{\mathfrak{N}}$ for simplicity, the first term of (S4) can be bounded by

$$\begin{aligned}
(*) \leq & \left\|\widehat{\mathbf{W}}_k^{1/2}\left(\mathbf{I} + (\mathbf{E}_{k,+}^+ - \mathbf{E}_{k,-}^+)\right)\mathbf{B}_k^\top\widehat{\mathbf{W}}_{k-1}^{-1}\left(\mathbf{I} - (\mathbf{E}_{k-1,+} - \mathbf{E}_{k-1,-})\right)\mathbf{B}_k\widehat{\mathbf{W}}_k^{1/2}\left(\mathbf{I} + (\mathbf{E}_{k,+}^+ - \mathbf{E}_{k,-}^+)\right)\right. \\
& \left. -\widehat{\mathbf{W}}_k^{1/2}\mathbf{B}_k^\top\widehat{\mathbf{W}}_{k-1}^{-1}\mathbf{B}_k\widehat{\mathbf{W}}_k^{1/2}\right\| \\
\leq & \left[\left(\left\|2\cdot\left(\mathbf{E}_{k,+}^+ - \mathbf{E}_{k,-}^+\right)\right\| + \left\|\left(\mathbf{E}_{k,+}^+ - \mathbf{E}_{k,-}^+\right)^2\right\|\right)\cdot\left\|\widehat{\mathbf{L}}_d\right\| + \right. \\
& \left.\left(1+\sqrt{\epsilon_k'}\right)^2\left\|\widehat{\mathbf{W}}_k^{1/2}\mathbf{B}_k^\top\widehat{\mathbf{W}}_{k-1}^{-1}\left(\mathbf{E}_{k-1,+} - \mathbf{E}_{k-1,-}\right)\mathbf{B}_k\widehat{\mathbf{W}}_k^{1/2}\right\|\right] \\
\leq & \left[\left\|2\cdot\left(\mathbf{E}_{k,+}^+ - \mathbf{E}_{k,-}^+\right)\right\| + \left\|\left(\mathbf{E}_{k,+}^+ - \mathbf{E}_{k,-}^+\right)^2\right\| + \left(1+\sqrt{\epsilon_k'}\right)^2\|\mathbf{E}_{k-1,+} - \mathbf{E}_{k-1,-}\|\right]\cdot\|\widehat{\mathbf{L}}_d\| \\
\overset{*}{\leq} & \left[2\sqrt{\epsilon_k'} + \epsilon_k' + \left(1+\sqrt{\epsilon_k'}\right)^2\sqrt{\epsilon_{k-1}'}\right]\cdot\left\|\tilde{\mathbf{L}}_d\right\|.
\end{aligned}$$

The last two terms of (S4) can be bounded using Lemma S1, i.e.,

$$\begin{aligned}
(\dagger) = \left\|\mathbf{M}_{\mathfrak{N}}\mathbf{W}_k^{1/2}\mathbf{B}_k^\top\mathbf{W}_{k-1}^{-1}\mathbf{B}_k\boldsymbol{\mathcal{E}}_{k,+}^{1/2}\mathbf{M}_{\mathfrak{C}}\right\| &\leq \left\|\mathbf{M}_{\mathfrak{N}}\mathbf{W}_k^{1/2}\mathbf{B}_k^\top\mathbf{W}_{k-1}^{-1/2}\right\|\cdot\left\|\mathbf{W}_{k-1}^{-1/2}\mathbf{B}_k\boldsymbol{\mathcal{E}}_{k,+}^{1/2}\mathbf{M}_{\mathcal{I}}\right\| \\
&\leq \|\mathbf{L}_d\|\sqrt{\epsilon_{k-1}}.
\end{aligned}$$

The last term of (S4) can also be bounded by $\|\tilde{\mathbf{L}}_d\|\sqrt{\epsilon_{k-1}}$ using Lemma S1. Since $\|\mathbf{L}_d\|$, $\left\|\widehat{\mathbf{L}}_d\right\|$, and $\left\|\tilde{\mathbf{L}}_d\right\|$ have the same upper bound λ_{k-1}, we have

$$\left\|\mathbf{L}_d - \widehat{\mathbf{L}}_d\right\|^2 \leq \left[2\sqrt{\epsilon_k'} + \epsilon_k' + \left(1+\sqrt{\epsilon_k'}\right)^2\sqrt{\epsilon_{k-1}'} + 4\sqrt{\epsilon_{k-1}}\right]^2\lambda_{k-1}^2.$$

■

Proof of the $\mathsf{DiffL}_k^{\mathrm{up}}$ term in Theorem 1. The explicit form of $\widehat{\mathbf{L}}_u$ is,

$$\widehat{\mathbf{L}}_u = \left[\begin{array}{c|c|c}
\mathbf{M}_{\mathfrak{N}}\widehat{\mathbf{W}}_k^{-1/2}\mathbf{B}_{k+1}\widehat{\mathbf{W}}_{k+1}\mathbf{B}_{k+1}^\top\widehat{\mathbf{W}}_k^{-1/2}\mathbf{M}_{\mathfrak{N}} & \mathbf{0} & \mathbf{M}_{\mathfrak{N}}\widehat{\mathbf{W}}_k^{-1/2}\mathbf{B}_{k+1}\widehat{\mathbf{W}}_{k+1}\mathbf{B}_{k+1}^\top\boldsymbol{\mathcal{E}}_{k,-}^{1/2}\mathbf{M}_{\mathfrak{D}} \\
\hline
\mathbf{0} & \mathbf{M}_{\mathfrak{C}}\boldsymbol{\mathcal{E}}_{k,+}^{-1/2}\mathbf{B}_{k+1}\mathbf{W}_{k+1}\mathbf{B}_{k+1}^\top\boldsymbol{\mathcal{E}}_{k,+}^{-1/2}\mathbf{M}_{\mathfrak{C}} & \mathbf{0} \\
\hline
\mathbf{M}_{\mathfrak{D}}\boldsymbol{\mathcal{E}}_{k,-}^{-1/2}\mathbf{B}_{k+1}\widehat{\mathbf{W}}_{k+1}\mathbf{B}_{k+1}^\top\widehat{\mathbf{W}}_k^{-1/2}\mathbf{M}_{\mathfrak{N}} & \mathbf{0} & \mathbf{M}_{\mathfrak{D}}\boldsymbol{\mathcal{E}}_{k,-}^{-1/2}\mathbf{B}_{k+1}\widehat{\mathbf{W}}_{k+1}\mathbf{B}_{k+1}^\top\boldsymbol{\mathcal{E}}_{k,+}^{-1/2}\mathbf{M}_{\mathfrak{D}}
\end{array}\right].$$

And,

$$\mathbf{L}_u = \left[\begin{array}{c|c|c} \mathbf{M}_{\mathfrak{N}}\mathbf{W}_k^{-1/2}\mathbf{B}_{k+1}\mathbf{W}_{k+1}\mathbf{B}_{k+1}^\top\mathbf{W}_k^{-1/2}\mathbf{M}_{\mathfrak{N}} & \mathbf{M}_{\mathfrak{N}}\mathbf{W}_k^{-1/2}\mathbf{B}_{k+1}\mathbf{W}_{k+1}\mathbf{B}_{k+1}^\top\boldsymbol{\mathcal{E}}_{k,+}^{1/2}\mathbf{M}_{\mathfrak{C}} & \mathbf{0} \\ \hline \mathbf{M}_{\mathfrak{C}}\boldsymbol{\mathcal{E}}_{k,+}^{-1/2}\mathbf{B}_{k+1}\mathbf{W}_{k+1}\mathbf{B}_{k+1}^\top\mathbf{W}_k^{-1/2}\mathbf{M}_{\mathfrak{N}} & \mathbf{M}_{\mathfrak{C}}\boldsymbol{\mathcal{E}}_{k,+}^{-1/2}\mathbf{B}_{k+1}\mathbf{W}_{k+1}\mathbf{B}_{k+1}^\top\boldsymbol{\mathcal{E}}_{k,+}^{-1/2}\mathbf{M}_{\mathfrak{C}} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{M}_{\mathfrak{D}}\boldsymbol{\mathcal{E}}_{k,-}^{-1/2}\mathbf{B}_{k+1}\widehat{\mathbf{W}}_{k+1}\mathbf{B}_{k+1}^\top\boldsymbol{\mathcal{E}}_{k,+}^{-1/2}\mathbf{M}_{\mathfrak{D}} \end{array}\right].$$

The perturbation is,

$$\left\|\mathbf{L}_u - \widehat{\mathbf{L}}_u\right\| \leq \underbrace{\left\|\mathbf{M}_{\mathfrak{N}}\mathbf{W}_k^{-1/2}\mathbf{B}_{k+1}\mathbf{W}_{k+1}\mathbf{B}_{k+1}^\top\mathbf{W}_k^{-1/2}\mathbf{M}_{\mathfrak{N}} - \mathbf{M}_{\mathfrak{N}}\widehat{\mathbf{W}}_k^{-1/2}\mathbf{B}_{k+1}\widehat{\mathbf{W}}_{k+1}\mathbf{B}_{k+1}^\top\widehat{\mathbf{W}}_k^{-1/2}\mathbf{M}_{\mathfrak{N}}\right\|}_{(*)} + 2\left[\underbrace{\left\|\mathbf{M}_{\mathfrak{N}}\mathbf{W}_k^{-1/2}\mathbf{B}_{k+1}\mathbf{W}_{k+1}\mathbf{B}_{k+1}^\top\boldsymbol{\mathcal{E}}_{k,+}^{1/2}\mathbf{M}_{\mathfrak{C}}\right\|}_{(\dagger)} + \left\|\mathbf{M}_{\mathfrak{N}}\widehat{\mathbf{W}}_k^{-1/2}\mathbf{B}_{k+1}\widehat{\mathbf{W}}_{k+1}\mathbf{B}_{k+1}^\top\boldsymbol{\mathcal{E}}_{k,-}^{1/2}\mathbf{M}_{\mathfrak{D}}\right\|\right]. \tag{S5}$$

The first term of (S5) can be bounded by expanding $\mathbf{W}_{k+1}$ w.r.t. $\widehat{\mathbf{W}}_{k+1}$, i.e., $\mathbf{W}_{k+1} = \widehat{\mathbf{W}}_{k+1} + (\boldsymbol{\mathcal{E}}_{k+1,+} - \boldsymbol{\mathcal{E}}_{k+1,-})$. As slight abuse of notation, we let $\mathbf{W}_k \leftarrow \mathbf{W}_k[\mathfrak{N},\mathfrak{N}]$, $\mathbf{B}_{k+1} \leftarrow \mathbf{B}_{k+1}[\mathfrak{N},:]$. The first term $(*)$ of (S5) becomes

$$\begin{aligned}
(*) &\leq \left\|\mathbf{W}_k^{-1/2}\mathbf{B}_{k+1}\widehat{\mathbf{W}}_{k+1}\mathbf{B}_{k+1}^\top\mathbf{W}_k^{-1/2} - \widehat{\mathbf{W}}_k^{-1/2}\mathbf{B}_{k+1}\widehat{\mathbf{W}}_{k+1}\mathbf{B}_{k+1}^\top\widehat{\mathbf{W}}_k^{-1/2}\right\| \\
&\leq \left\|\mathbf{W}_k^{-1/2}\mathbf{B}_{k+1}\widehat{\mathbf{W}}_{k+1}\mathbf{B}_{k+1}^\top\mathbf{W}_k^{-1/2} - \widehat{\mathbf{W}}_k^{-1/2}\mathbf{B}_{k+1}\widehat{\mathbf{W}}_{k+1}\mathbf{B}_{k+1}^\top\widehat{\mathbf{W}}_k^{-1/2}\right\| + \\
&\quad \left\|\mathbf{W}_k^{-1/2}\mathbf{B}_{k+1}\boldsymbol{\mathcal{E}}_{k+1,+}\mathbf{B}_{k+1}^\top\mathbf{W}_k^{-1/2} - \widehat{\mathbf{W}}_k^{-1/2}\mathbf{B}_{k+1}\boldsymbol{\mathcal{E}}_{k+1,-}\mathbf{B}_{k+1}^\top\widehat{\mathbf{W}}_k^{-1/2}\right\| \\
&\overset{\ddagger}{\leq} \left(2\|\mathbf{E}_{k,+}^- - \mathbf{E}_{k,-}^-\| + \left\|\left(\mathbf{E}_{k,+}^- - \mathbf{E}_{k,-}^-\right)^2\right\|\right)\cdot\left\|\widehat{\mathbf{L}}_u\right\| + \\
&\quad \left\|\mathbf{W}_k^{-1/2}\mathbf{B}_{k+1}\boldsymbol{\mathcal{E}}_{k+1,+}\mathbf{B}_{k+1}^\top\mathbf{W}_k^{-1/2} - \widehat{\mathbf{W}}_k^{-1/2}\mathbf{B}_{k+1}\boldsymbol{\mathcal{E}}_{k+1,-}\mathbf{B}_{k+1}^\top\widehat{\mathbf{W}}_k^{-1/2}\right\| \\
&\overset{\S}{\leq} \left[2\sqrt{\epsilon_k'} + \epsilon_k' + 2\epsilon_k\right]\lambda_k
\end{aligned}$$

The $\ddagger$ term holds by expanding $\mathbf{W}_k^{-1/2} = \widehat{\mathbf{W}}_1^{-1/2}\left(\mathbf{I} - (\mathbf{E}_{k,+}^- - \mathbf{E}_{k,-}^-)\right)$ and following a similar approach of the *down* Laplacian. The $\S$ term holds by bounding $\mathbf{E}_{k,+}^- - \mathbf{E}_{k,-}^-$ with Assumption S1 (ϵ_k') and using Lemma S2 (ϵ_k).

The $(\dagger)$ term in (S5) can be bounded by ϵ_k using Lemma S2, i.e.,

$$\begin{aligned}
(\dagger) &\overset{\ddagger}{=} \left\|\mathbf{M}_{\mathfrak{N}}\mathbf{W}_k^{-1/2}\mathbf{B}_{k+1}\boldsymbol{\mathcal{E}}_{k+1,+}\mathbf{B}_{k+1}^\top\boldsymbol{\mathcal{E}}_{k,+}^{-1/2}\mathbf{M}_{\mathfrak{C}}\right\| \\
&\leq \left\|\mathbf{M}_{\mathfrak{N}}\mathbf{W}_k^{-1/2}\mathbf{B}_{k+1}\boldsymbol{\mathcal{E}}_{k+1,+}^{1/2}\right\|\cdot\left\|\boldsymbol{\mathcal{E}}_{k+1,+}^{1/2}\mathbf{B}_{k+1}^\top\boldsymbol{\mathcal{E}}_{k,+}^{-1/2}\right\| \\
&\leq \sqrt{\lambda_k\epsilon_k}\cdot\left\|\boldsymbol{\mathcal{E}}_{k+1,+}^{1/2}\mathbf{B}_{k+1}^\top\boldsymbol{\mathcal{E}}_{k,+}^{-1/2}\right\| \\
&\overset{\S}{\leq} \sqrt{\epsilon_k}\lambda_k.
\end{aligned}$$

$\ddagger$ holds because the intersection of triangles of $\boldsymbol{\mathcal{E}}_{k,+}$, and $\mathbf{W}_k$ is the triangles with non-zero entries in $\boldsymbol{\mathcal{E}}_{k+1,+}$. $\S$ holds (the $\sqrt{\lambda_k}$ term) because $\boldsymbol{\mathcal{E}}_{k+1,+}^{1/2}\mathbf{B}_{k+1}^\top\boldsymbol{\mathcal{E}}_{k,+}^{-1/2}$ is a submatrix of $\mathbf{W}_{k+1}^{1/2}\mathbf{B}_{k+1}^\top\mathbf{W}_k^{-1/2}$; hence, the spectral norm will be upper bounded by the *up* Laplacian $\|\mathbf{L}_u\| \leq \lambda_k$.

Therefore, we have

$$\left\|\mathbf{L}_u - \widehat{\mathbf{L}}_u\right\|^2 \leq \left[2\sqrt{\epsilon'_k} + \epsilon'_k + 2\epsilon_k + 4\sqrt{\epsilon_k}\right]^2 \lambda_k^2.$$

Combining the bound involving $\mathsf{DiffL}_k^{\mathrm{down}}$ completes the proof of Theorem 1. ■

B Proofs of propositions in Applications (Section 5)

B.1 Proof of Proposition 3: the properties of the induced digraph

The proof is based on the convenient properties of the harmonic flow (the basis of the homology vector space), i.e., they are both *divergence-free* and *curl-free* [12, 35, 48].

Proof of Proposition 3. Reachable: the harmonic flow is divergence free, indicating that the incoming flow must be equal to the outgoing flow. If there exists a vertex that is not *reachable* to itself, then this vertex will either be a *source* or *sink* in the digraph. It violates the assumption that the flow is divergence free. Therefore such vertex will not exist.

No short-circuiting: the harmonic flow is curl free; from Stoke's theorem (or Poincaré Lemma [33]), we have that any path-integral travel along any homology class will be a constant. If there exists a loop such that it does not traverse along with any homology class, the loop integral along this cycle will be zero (by Stoke's theorem). By assumption, the path-integral will always be positive. To generate a loop whose integral is zero, one has to travel "upward" in the digraph; this violates the assumption that we are finding a cycle in the digraph, implying that every loop will traverse along at least one homology class. ■

B.2 Proof of Proposition 4: $\mathcal{H}_1$ embedding of $\mathbb{T}^m$

The proof is based on the fact that each harmonic 1-form of the flat m-(flat) torus can be expressed as the m-dimensional standard basis multiplied with some intensities in the intrinsic coordinate. The closed-form of the upper bound of the embedding distribution in any direction can be derived using the (high-dimensional) polar coordinate system, indicating that the envelope is an m-dimensional ellipsoid. The detailed proof is provided below.

Proof of Proposition 4. The harmonic vector field in an m-flat torus $\mathbb{T}^m$ is a constant in each coordinate, i.e., $\mathbf{v} = [v_1, \cdots, v_m] \in \mathbb{R}^m$. The manifold $\mathbb{T}^m$ is an m-dimensional cube with the periodic boundary condition, i.e., $0 = 2\pi$. From [12, 59], the edge flow $\boldsymbol{\omega}_e$ for an edge $e = (i, j) \in E$ can be written exactly as a linear map, i.e.,

$$\begin{aligned}\omega_e &= \int_0^1 \mathbf{v}^\top(\boldsymbol{\gamma}(t))\gamma'(t)\mathrm{d}t = \int_0^1 [\mathbf{v}(\mathbf{x}_i) + (\mathbf{v}(\mathbf{x}_j) - \mathbf{v}(\mathbf{x}_i))t]^\top(\mathbf{x}_j - \mathbf{x}_i)\mathrm{d}t \\ &= \frac{1}{2}(\mathbf{v}(\mathbf{x}_i) + \mathbf{v}(\mathbf{x}_j))^\top(\mathbf{x}_j - \mathbf{x}_i)\end{aligned}$$

Where $\boldsymbol{\gamma}(t)$ is the geodesic on $\mathcal{M}$ connecting $\mathbf{x}_i$ and $\mathbf{x}_j$ with $\boldsymbol{\gamma}(0) = \mathbf{x}_i$ and $\boldsymbol{\gamma}(1) = \mathbf{x}_j$. Any point $\mathbf{x} \in \mathbb{R}^m$, with $r = \|\mathbf{x}\|$, can be written as $\mathbf{x} = [rf_1(\boldsymbol{\Phi}), rf_2(\boldsymbol{\Phi}), \cdots, rf_m(\boldsymbol{\Phi})]$, where $\boldsymbol{\Phi} \in \mathbb{R}^{m-1}$ is the high-dimensional polar coordinate; for instance, a point in 2D is $[r\cos(\theta), r\sin(\theta)]$ with $\boldsymbol{\Phi} = [\theta]$, while a point in 3D having $\boldsymbol{\Phi} = [\theta, \varphi]$ is $[r\cos\varphi\sin\theta, r\sin\varphi\sin\theta, r\cos\theta]$. The conditional distribution given a fixed $\boldsymbol{\Phi}$ is simply the distribution of edge lengths, i.e., $p(rv_1f_1, \cdots, rv_mf_m|\boldsymbol{\Phi}) = p(r)$. Since $p(r)$ is bounded by some constant δ representing the maximum edge length, the envelope of the distribution is bounded by $[\delta v_1 f_1(\boldsymbol{\Phi}), \cdots, \delta v_m f_m(\boldsymbol{\Phi})]$, indicating that it is an m-ellipsoid with the length of the i-th semi-axes being δv_i. ■

C The maximum eigenvalue of $\mathcal{L}_k$ constructed from a cubical complex

In this section, we would like to show the bound on the spectral norm of $\boldsymbol{\mathcal{L}}_k$ built from a cubical complex. The property is found useful in extending Theorem 1 to Corollary 2; namely, the goal is

to show that $\|\boldsymbol{\mathcal{L}}_k\|_2 \leq \lambda_k = (2k+2)$. Note that $\|\boldsymbol{\mathcal{L}}_k^{\text{down}}\| = \|\mathbf{A}_k^\top \mathbf{A}_k\| = \|\mathbf{A}_k \mathbf{A}_k^\top\| = \|\boldsymbol{\mathcal{L}}_{k-1}^{\text{up}}\|$. W.l.o.g., one can inspect only the up-Laplacian. We provide the following proposition that is largely based on the similar analysis [26] of $\|\boldsymbol{\mathcal{L}}_k\|$ for SC.

Proposition S3. *Given an* up *k-Laplacian $\boldsymbol{\mathcal{L}}_k^{\text{up}} = \mathbf{A}_{k+1}\mathbf{A}_{k+1}^\top$ with $\mathbf{A}_{k+1} = \mathbf{W}_k^{-1/2}\mathbf{B}_{k+1}\mathbf{W}_{k+1}^{1/2}$ built from a cubical complex, we have*

$$\|\boldsymbol{\mathcal{L}}_k^{\text{up}}\|_2 \leq \lambda_k = 2k+2.$$

Proof. From [48], the eigenvalues of the k-th renormalized *up*-Laplacian $\boldsymbol{\mathcal{L}}_k^{\text{up}}$ are identical to those of the k-th *random-walk up*-Laplacian $\boldsymbol{\mathcal{L}}_k^{\text{rw}} = \mathbf{W}_k^{-1}\mathbf{B}_{k+1}\mathbf{W}_{k+1}\mathbf{B}_{k+1}$. Further, let $\mathbf{L}_k^{\text{up}} = \mathbf{B}_{k+1}\mathbf{W}_{k+1}\mathbf{B}_{k+1}^\top$, following the analysis of [26], we have

$$\begin{aligned}
\mathbf{f}^\top \mathbf{L}_k^{\text{up}} \mathbf{f} &= \left(\mathbf{W}_{k+1}^{1/2}\mathbf{B}_{k+1}^\top \mathbf{f}\right)^\top \left(\mathbf{W}_{k+1}^{1/2}\mathbf{B}_{k+1}^\top \mathbf{f}\right) \\
&= \sum_{\sigma \in K_k} \sum_{\tau \in \text{coface}(\sigma)} f^2(\sigma) w_{k+1}(\tau) \\
&\overset{\dagger}{\leq} (2k+2) \sum_{\sigma \in K_k} f^2(\sigma) \sum_{\tau \in \text{coface}(\sigma)} w_{k+1}(\tau) \\
&= (2k+2) \sum_{\sigma \in K_k} f^2(\sigma) \deg(\sigma).
\end{aligned}$$

The inequality † holds using the Cauchy-Schwarz inequality; the $2k+2$ term comes from the fact that a $(k+1)$-cube has $(2k+2)$ faces. Following the rest of the proof in [26], we have

$$\|\boldsymbol{\mathcal{L}}_k^{\text{up}}\| = \|\boldsymbol{\mathcal{L}}_k^{\text{rw,up}}\| = \frac{\|\mathbf{L}_k^{\text{up}}\|}{\mathbf{f}^\top \mathbf{W}_k \mathbf{f}} \leq (2k+2) \frac{\sum_{\sigma \in K_k} f^2(\sigma) \deg(\sigma)}{\sum_{\sigma \in K_k} f^2(\sigma) w_k(\sigma)} = 2k+2.$$

The first equality holds due to the identical eigenvalues of $\boldsymbol{\mathcal{L}}_k$ and $\boldsymbol{\mathcal{L}}_k^{\text{rw}}$; the last inequality holds because we have $\mathbf{w}_k(\sigma) = |\mathbf{B}_{k+1}(\sigma)|\mathbf{w}_{k+1} = \deg(\sigma)$ for all $\sigma \in K_k$. ■

D Datasets, experiment details, and discussions

The edge set E of the neighborhood graph constructed using the CkNN kernel [8] is

$$E = \left\{ i, j \in V : \frac{\|\mathbf{x}_i - \mathbf{x}_j\|}{\sqrt{\rho_k(\mathbf{x}_i)\rho_k(\mathbf{x}_j)}} \leq \delta \right\}.$$

Here, $\rho_k(\mathbf{x})$ is the distance from $\mathbf{x}$ to its k-th nearest neighbor; throughout the experiment, we fix $k = 30$. The δ parameter can be chosen by a variant of the geometric consistent (GC) algorithm [29] suitable for CkNN graphs; for real datasets (except for the ocean drifter data whose geometric property is known), we use the modified GC to choose this parameter. For the rest of the datasets (synthetic manifolds and the ocean drifter), δ's are chosen manually since the topologies are known. The weights on the triangles are selected by a modification to the kernel in [12], with a similar choice of $\varepsilon = \delta^{\frac{2}{3}}/3$,

$$w_2(i,j,\ell) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\varepsilon \rho_k(\mathbf{x}_i)\rho_k(\mathbf{x}_j)}\right) \cdot \exp\left(-\frac{\|\mathbf{x}_j - \mathbf{x}_\ell\|^2}{\varepsilon \rho_k(\mathbf{x}_j)\rho_k(\mathbf{x}_\ell)}\right) \cdot \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_\ell\|^2}{\varepsilon \rho_k(\mathbf{x}_i)\rho_k(\mathbf{x}_\ell)}\right).$$

With this choice of parameters, the corresponding $\boldsymbol{\mathcal{L}}_1$ has a large sample size limit (in terms of Δ_1) w.r.t. the metrics normalized by the k-nearest neighbor distance ρ_k.

D.1 Synthetic manifolds

PUNCTPLANE. PUNCTPLANE is a manifold generated by connected summing two punctured planes, with a (sparsely connected) bridge in between. Each punctured plane has a rectangular hole with width/height being $1/3$ of the width of each manifold.

TORUS. This data is a two-dimensional torus and is generated from the parameterization below,

$$\begin{aligned} x_1 &= (1 + 0.5\cos\theta_1)\cos\theta_2; \\ x_2 &= (1 + 0.5\cos\theta_1)\sin\theta_2; \\ x_3 &= 1 + 0.5\sin\theta_1. \end{aligned}$$

The sample size is $n = 1,156$. Random Gaussian noise is added on the first three dimensions as well as the additional 10 (noise) dimensions.

3-TORUS. The parameterization of 3-TORUS, a three torus with $d = 3$ and $D = 4$, is

$$\begin{aligned} x_1 &= (4 + (2 + \cos\theta_1)\cos\theta_2)\cos\theta_3; \\ x_2 &= (4 + (2 + \cos\theta_1)\cos\theta_2)\sin\theta_3; \\ x_3 &= (2 + \cos\theta_1)\sin\theta_2; \\ x_4 &= \sin\theta_1. \end{aligned}$$

We first sample $n' = 100,000$ points from this manifold; Algorithm S4 is used to generate $\mathbf{X}$ with $n = 2,000$.

GENUS-2. GENUS-2 is a two-dimensional (*genus-2*) surface generated by gluing two tori together. The implicit equation of the surface is

$$\left(\left(x_1^2 + x_2^2\right)^2 - 0.75x_1^2 + 0.75x_2^2\right)^2 + x_3^2 = 0.01.$$

To sample from this surface, we create a $1,000 \times 1,000$ grid in the first two coordinates (x_1, x_2) and solve for the corresponding x_3 from the above implicit equation. The aforementioned procedure generates a point cloud $\tilde{\mathbf{X}}$ ($n' \approx 551$k) having a non-uniform sampling density on the genus-2 surface; we subsample $\tilde{\mathbf{X}}$ by Algorithm S4 and obtain the final dataset $\mathbf{X}$ with $n = 1,500$.

TORI-CONCAT. TORI-CONCAT is generated by concatenating four tori together. Four tori are generated by similar procedures as TORUS with horizontal movements (in x_1) being $a = -3, 0, 3, 6$, i.e., $x_1 = (1 + 0.5\cos\theta_1)\cos\theta_2 - a$, respectively. The sample size of TORI-CONCAT is $n = 4,624$.

D.2 Real datasets

Small molecule datasets (ETH and MDA). The database [3] [13] contains several molecular dynamics (MD) trajectories, with each for a single (small) molecule, e.g., ethanol CH_3CH_2OH (ETH) and malondialdehyde $CH_2(CHO)_2$ (MDA). If a molecule has N atoms, then a point (molecular configuration) in the dataset is specified by an $N \times 3$ matrix representing the Euclidean coordinate of the configuration. To generate a point cloud from a trajectory of configurations, we first preprocess the data by calculating two angles of every triplet of atoms. Secondly, we remove the linear subspaces by keeping the top *principal components* (PCs) such that the unexplained variance ratio is less than 10^{-4}. The ambient dimensions of ETH and MDA are $D = 102$ and $D = 98$, respectively. We subsample furthest $n = 1,500$ points using Algorithm S4 for both datasets. The bond torsions (insets of Figures 3a and 3d) are calculated by the dihedral angles of the corresponding chemical bonds for each molecular configuration. For instance, the green torsion of ethanol (Figure 3a) for every point is

[3] Data from http://quantum-machine.org/datasets/

computed by the angle of the planes spanned by OCC and CCH in the configuration (3D Euclidean) space. One can think of the bond torsions as intrinsic coordinates of TORUS, i.e., θ_1 and θ_2; note that the correct bond torsions parametrizing the manifold (or $\mathbf{X}$) are usually unknown beforehand. In this work, this information is provided based on our knowledge to validate our framework.

Single-cell RNA sequencing data PANCREAS. PANCREAS [7] is a single-cell RNA sequencing data with cell cycles. The data and preprocessing codes can be found in https://github.com/theislab/scvelo_notebooks/blob/master/Pancreas.ipynb. The original data has sample size $n' = 3,696$; we subsample $n = 2,000$ furthest points (Algorithm S4) to remove the non-uniform sampling density on the original manifold.

3D graphics 3D-GRAPH. The 3D model of a Buddha statue, which can be downloaded from https://www.cc.gatech.edu/projects/large_models/, provides a triangulation computed by [15]; in other words, the simplicial complex $\mathrm{SC}_2' = (V', E', T')$ is available beforehand, with $n' \approx 500$k and $n_1' \approx 2$M. To illustrate the efficacy of our framework (and Theorem 1), we treat 3D-GRAPH as a point cloud and build SC from the subsampled $n = 3,000$ furthest points by Algorithm S4.

Ocean buoys dataset ISLAND. The global Lagrangian drifter data (available in http://www.aoml.noaa.gov/envids/gld/) was collected by NOAA's Atlantic Oceanographic and Meteorological Laboratory and analyzed by Froyland and Padberg-Gehle [23] on the coherent flow structures of the ocean current. The dataset contains multiple trajectories of buoys dated between 2010–2019, with the location, velocity, and water temperature of each buoy recorded. The dataset itself is a 3D point cloud by converting the location (in latitude and longitude coordinates) to the *earth-centered, earth-fixed* (ECEF) coordinate system. We subsample $n = 5,000$ furthest points/buoys (Algorithm S4) with longitudes within $142°$E–$179°$E and latitudes between $48°$S–$33°$S; namely, we sampled buoys that were located around the Tasman sea.

Medical imaging data RETINA. RETINA is one of the medical images of the STARE project [25], a retinal imaging data collection. The database consists of around 400 raw images of human retinas, with diagnosis codes, the segmented blood vessel, and the detected optic nerve available in http://cecas.clemson.edu/~ahoover/stare/. We use the retinal image with ID being 179, which has numerous bright (circular) spots visible. We construct the cubical complex by intensity thresholding and morphological closing, resulting in $n = 25,237$, $n_1 = 49,793$, and $n_2 = 24,548$.

D.3 Pairwise scatter plots

In this section, we show the pairwise scatter plots for $\mathbf{Z}$ (blue) and $\mathbf{Y}$ (red); specifically, we would like to show that the independent homology embedding $\mathbf{Z}$ obtained by Algorithm 1 is (approximately) factorizable. The blue embeddings (lower diagonal) in Figures S1–S6 confirm this. By contrast, most coordinate of the red embeddings $\mathbf{Y}$ do not correspond to a subspace, except for PANCREAS and 3D-GRAPH in Figures S5 and S3, respectively.

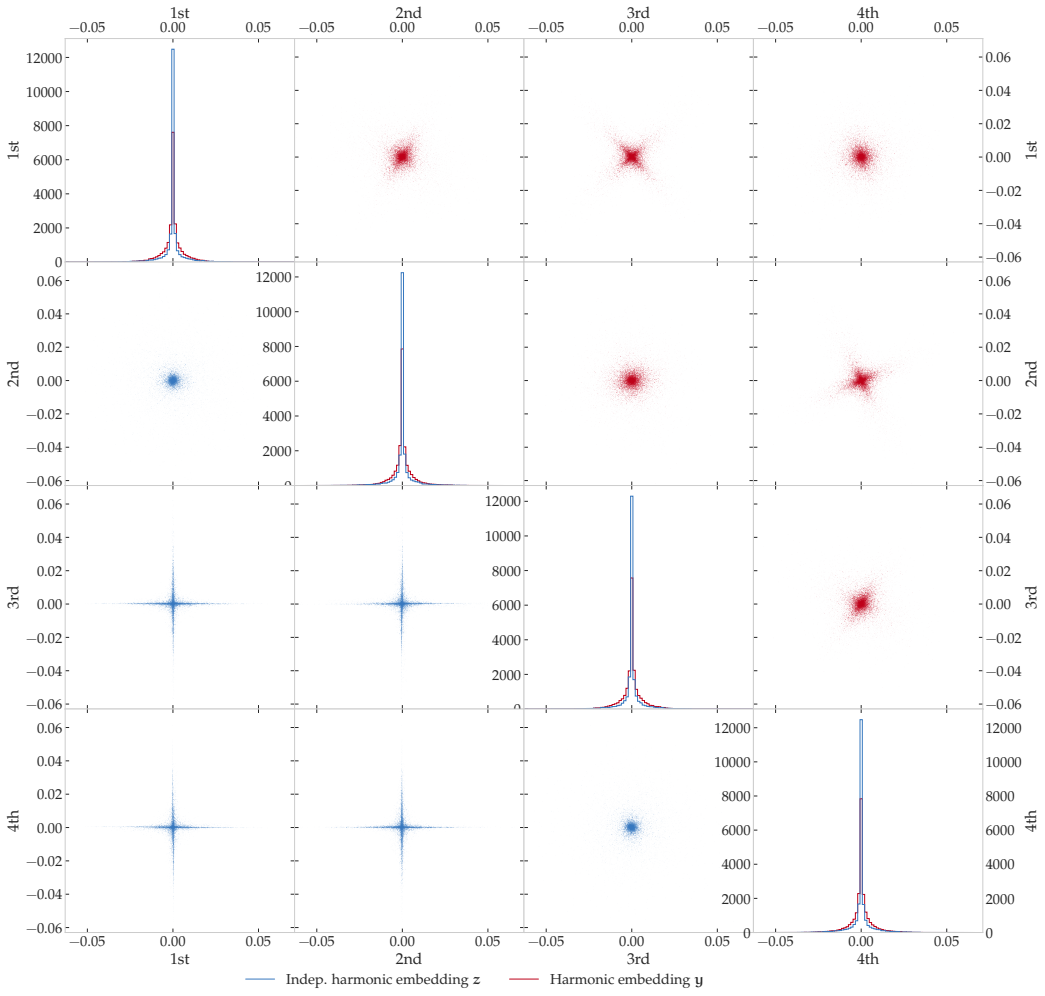

Figure S1: The independent ($\mathbf{z}$, in blue) and the coupled ($\mathbf{y}$, in red) homology embeddings of `GENUS-2`. The (i, j)-th (off-diagonal) subplot represents the two-dimensional scatter plot with the i-th and j-th coordinates of the embedding; the i-th diagonal term is the histograms of the i-th coordinate of the corresponding embedding.

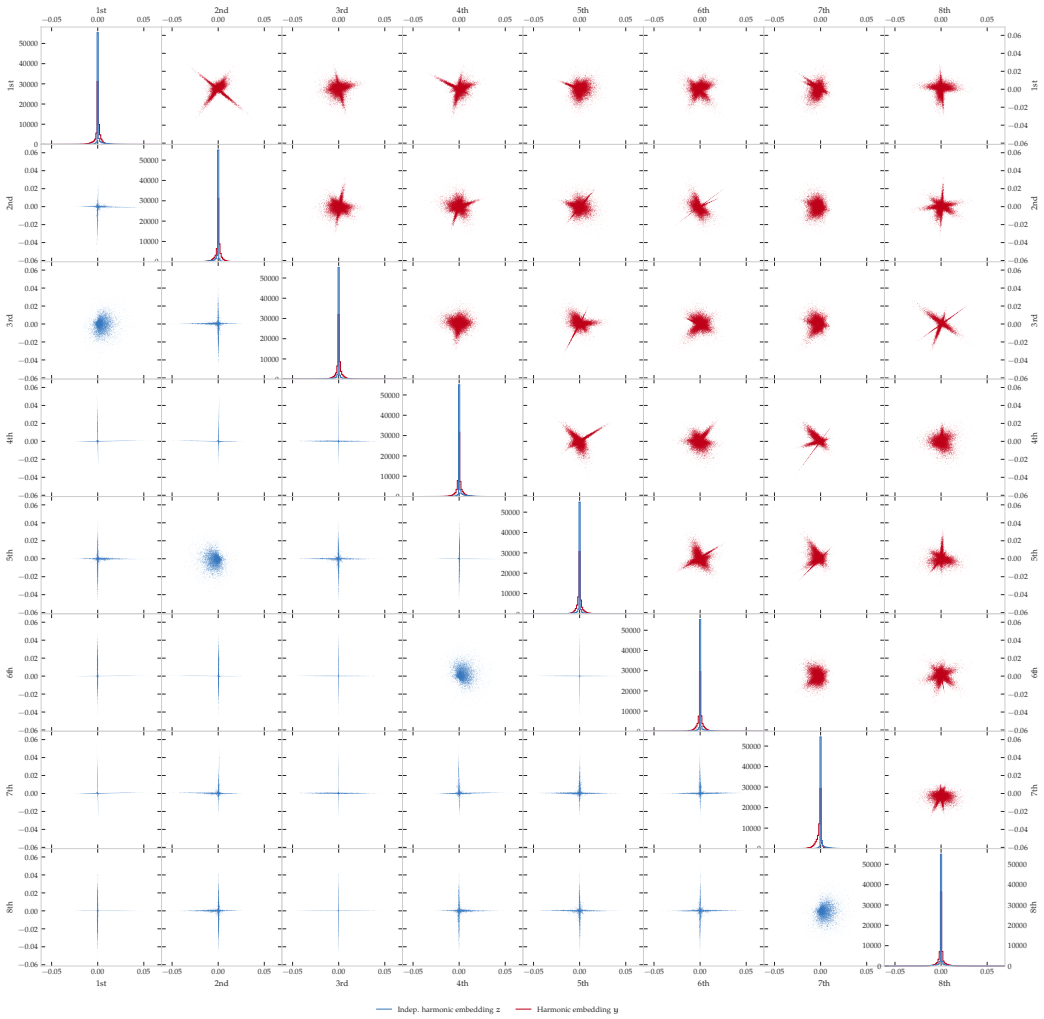

Figure S2: The independent ($\mathbf{z}$, in blue) and the coupled ($\mathbf{y}$, in red) homology embeddings of TORI-CONCAT.

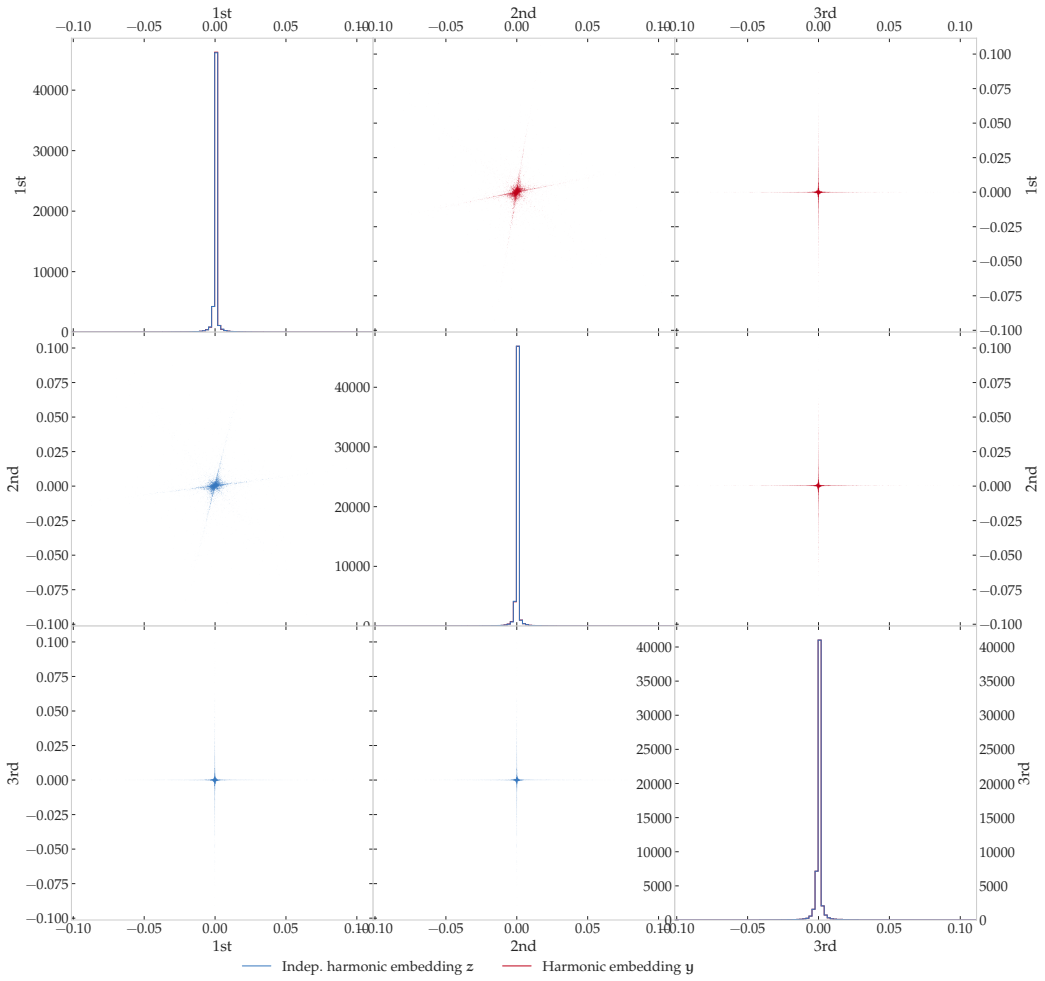

Figure S3: The independent ($\mathbf{z}$, in blue) and the coupled ($\mathbf{y}$, in red) homology embeddings of `3D-GRAPH`.

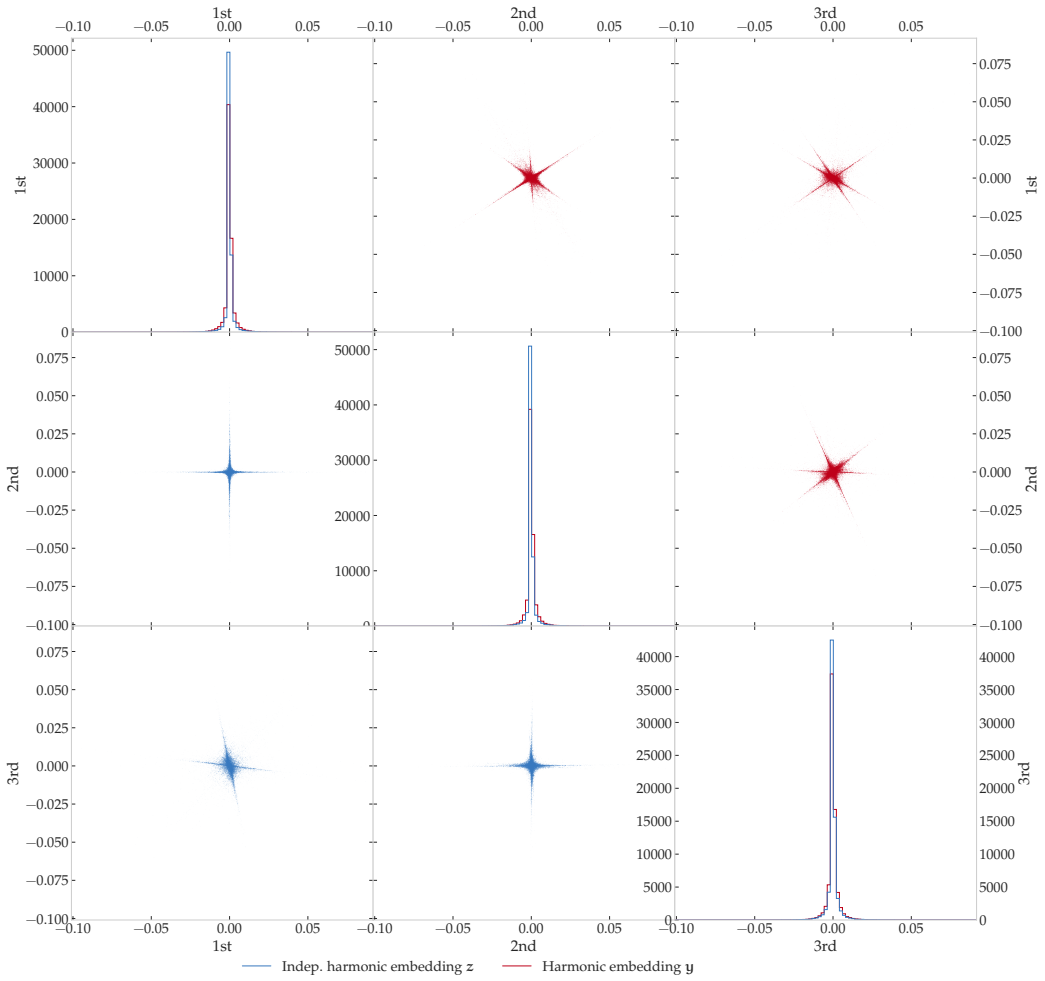

Figure S4: The independent ($\mathbf{z}$, in blue) and the coupled ($\mathbf{y}$, in red) homology embeddings of `ISLAND`.

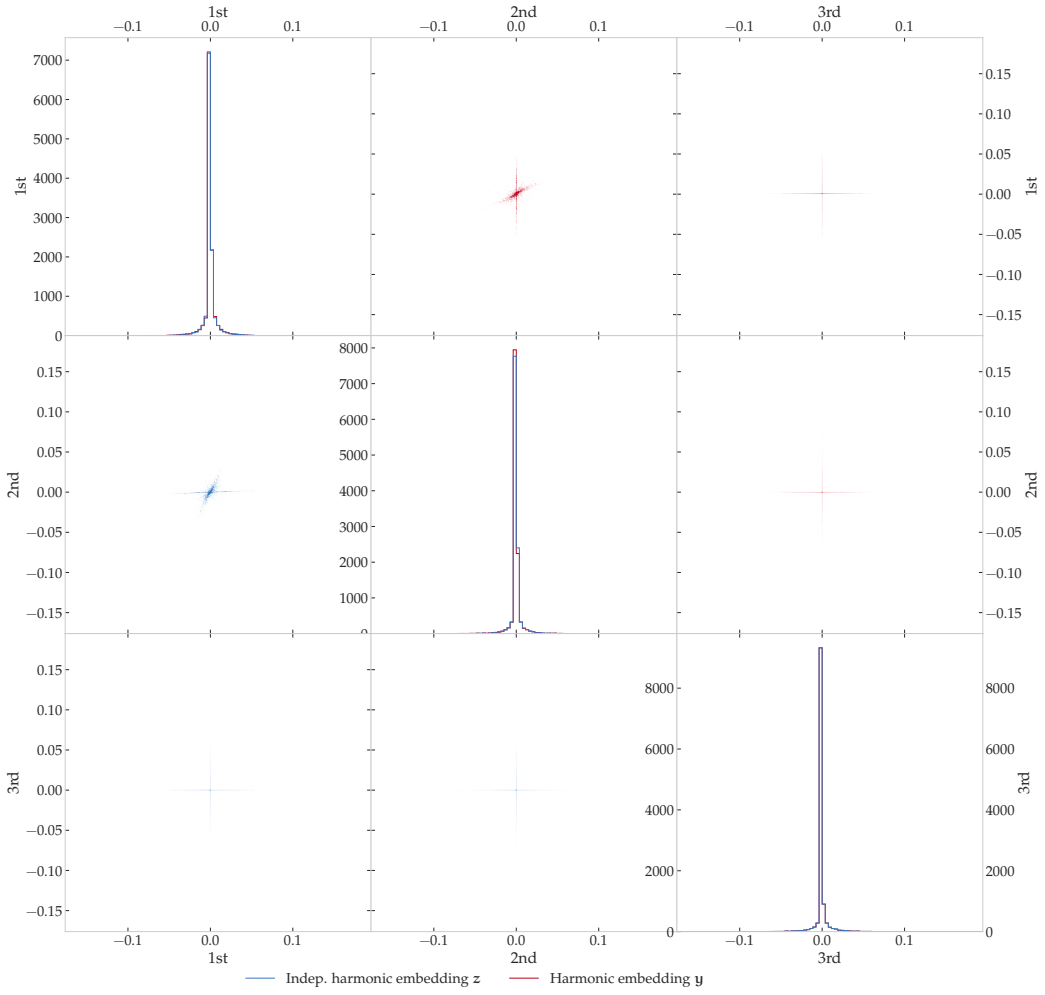

Figure S5: The independent ($\mathbf{z}$, in blue) and the coupled ($\mathbf{y}$, in red) homology embeddings of PANCREAS.

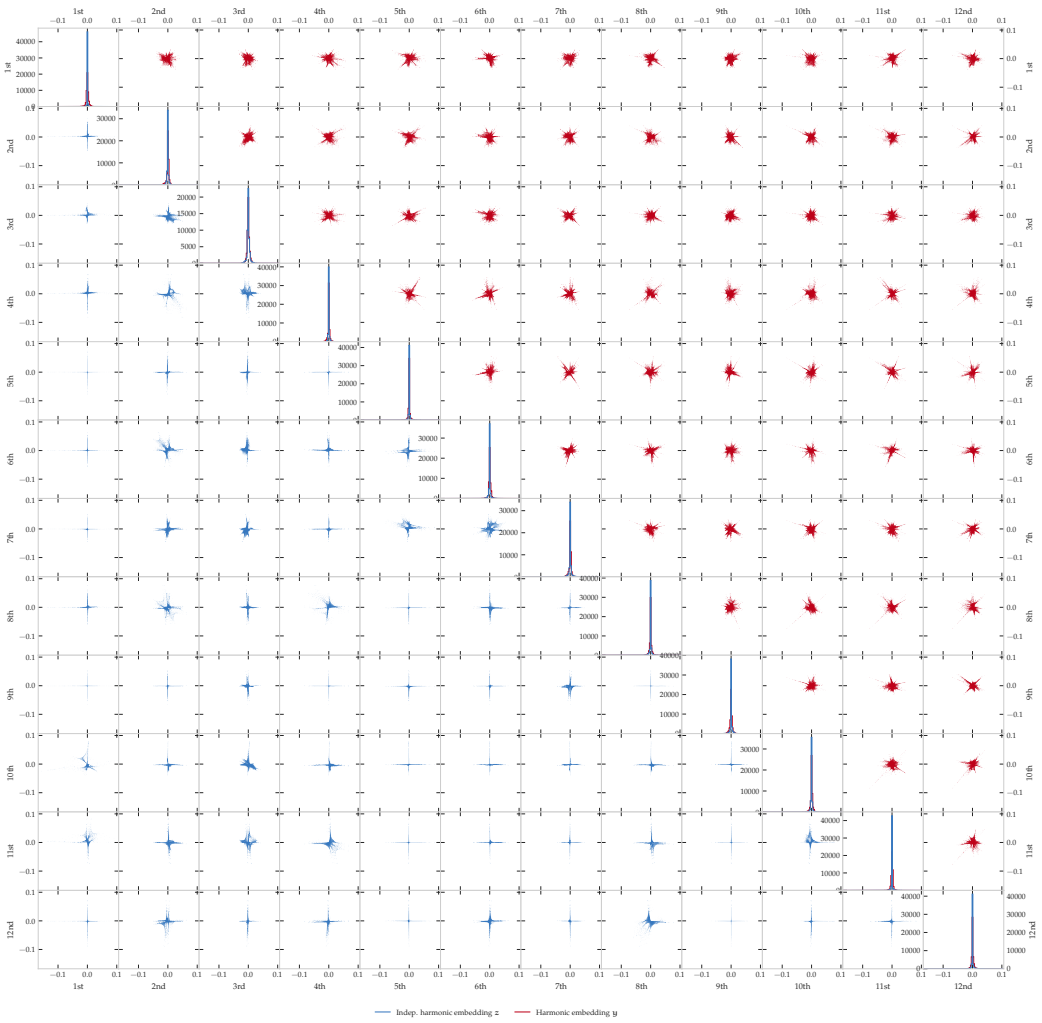

Figure S6: The independent (**z**, in blue) and the coupled (**y**, in red) homology embeddings of RETINA.

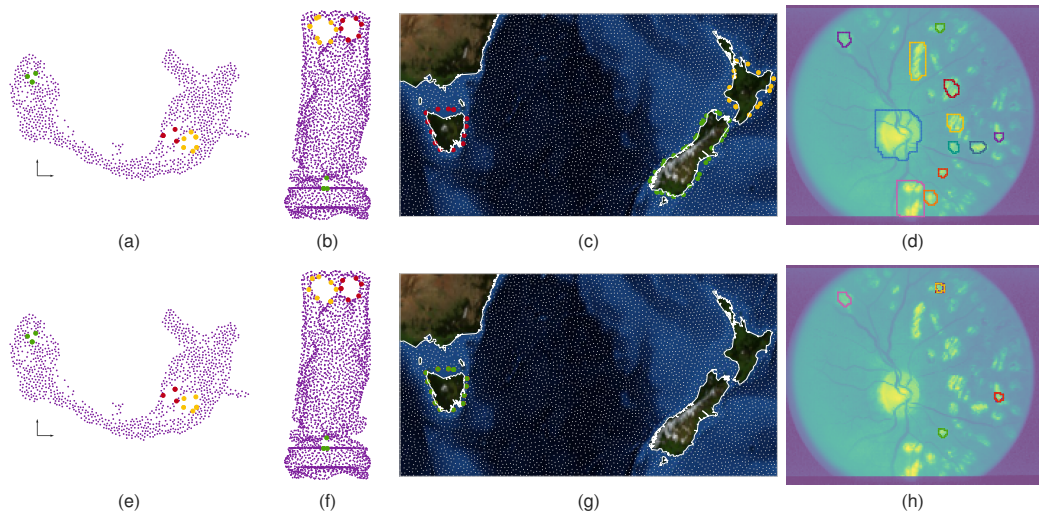

Figure S7: Comparison of the homologous loop detections on $\mathbf{Z}$ (the first row) and $\mathbf{Y}$ (the second row). The first, the second, the third, and the fourth columns present the results on PANCREAS, 3D-GRAPH, ISLAND, and RETINA, respectively. Note that (a)–(d) are identical to Figures 3g–3j.

D.4 Experiments and discussions on the shortest homologous loops detection algorithm

D.4.1 Shortest homologous loops obtained from the coupled embedding $\mathbf{Y}$

Figure S7 shows the results of the shortest homologous loop detection algorithm applied on the coupled homology embeddings $\mathbf{Y}$ on the real datasets. Note that Figures S7a–S7d are identical to Figures 3g–3j; they are presented here as comparisons to the loops detected from $\mathbf{Y}$ (the second row). As shown in Figures S7g and S7h, duplicated loops might be extracted if using the coupled embedding $\mathbf{Y}$; these loops are clearly sub-optimal.

D.4.2 Comparison to existing homologous loop detection algorithms

In this section, we discuss some existing methods in solving the problem of homologous loop detection. Some of these works offer efficient and elegant alternatives in cases when extra information is available, e.g., a triangular mesh, having an initial non-trivial loop, or in the special case of surfaces in 3D. Therefore, we see the proposed Algorithm 2, which does apply to point clouds/VR complexes with any ambient or intrinsic dimension, as complementary to these methods.

Dey et al. [17] proposed an integer programming formulation of finding the optimal homology loops given a non-trivial cycle; this problem is intrinsically different from our spectral loops extraction algorithm since no loops are given up-front. Nevertheless, the integer programming problem can be relaxed to linear programming when the boundary matrix $\mathbf{B}_2$ is total unimodular, which can be solved in $\mathcal{O}^*(\sqrt{n_1}n_2 + n_1^{5/2})$ by Lee and Sidford [34]. An example of the total unimodularity assumption to hold is when SC is constructed from a triangularization of the finite manifold. Even in this scenario, the proposed spectral loop extraction method (Algorithm 2) scales better asymptotically.

The method by Dey et al. [16] has runtime $\mathcal{O}(N^4)$, where $N = n + n_1 + n_2$ is the size of SC_2 (the detailed runtime is in the paper). By contrast, the persistent cycles extraction methods by Dey et al. [19], Wu et al. [60] use the annotation algorithm [10]. The annotation stage has a time complexity of $\mathcal{O}(N^{2.37\cdots})$. As discussed in the main paper, they are tractable for the simplicial complexes built from triangularization; however, this is not the case for VR complex since n_2 will grow fast as n increases.

Lastly, Chambers et al. [11], Dey et al. [18], Feng and Tong [22] have relatively fast runtime (in the order of $\mathcal{O}(n \log n)$). However, they either assume that the points are sampled from surfaces in 3D

(i.e., a genus-g surface due to the classification theorem) or the graph can be triangularized. It is oftentimes not the case for real datasets such as `ETH` or `MDA` in Figure 3.

E Pseudocodes

Algorithm S1: Spectral homologous loop detection—an alternative to Algorithm 2

Input : $\mathbf{Z} = [\mathbf{z}_1, \cdots, \mathbf{z}_{\beta_1}]$, V, E, edge distance $\mathbf{d}$

for $i = 1, \cdots, \beta_1$ **do**
- $E_i^+ \leftarrow \{(s,t) : e = (s,t) \in E \text{ and } [\mathbf{z}_i]_e > 0\}$
- $E_i^- \leftarrow \{(t,s) : e = (s,t) \in E \text{ and } [\mathbf{z}_i]_e < 0\}$
- $E_i \leftarrow E_i^+ \cup E_i^-$
- $G_i \leftarrow (V, E_i)$, with weight of $e \in E_i$ being $[\mathbf{d}]_e$
- $e_* = (t, s_0) \leftarrow \operatorname{argmax}_{e \in E_i} [\mathbf{z}_i]_e$
- $[s_0, s_1, \cdots, t] \leftarrow$ DIJKSTRA$(G_i, \texttt{from=}s_0, \texttt{to=}t)$
- $\mathcal{C}_i \leftarrow [t, s_0, s_1, \cdots, t]$

Return : $\mathcal{C}_1, \cdots, \mathcal{C}_{\beta_1}$

Algorithm S2: BOUNDARYMAPS

Input : $\mathrm{SC}_\ell = (\Sigma_0, \cdots, \Sigma_\ell)$, k

▷ `Requires` $\ell \geq k+1$

$\mathbf{B}_k \leftarrow$ BOUNDARYMAP(Σ_{k-1}, Σ_k) ▷ `Algorithm S3`

$\mathbf{B}_k \leftarrow$ BOUNDARYMAP(Σ_k, Σ_{k+1})

Return : Boundary maps $\mathbf{B}_k$, $\mathbf{B}_{k+1}$

Algorithm S3: BOUNDARYMAP

Input : Set of $(k-1)$ and k-simplices Σ_{k-1}, Σ_k (or cubes K_{k-1}, K_k)

$\mathbf{B}_k \leftarrow \mathbf{0}_{n_{k-1}} \mathbf{0}_{n_k}^\top \in \mathbb{R}^{n_{k-1} \times n_k}$

for *every* $\sigma_{k-1} \in \Sigma_{k-1}$ **do**
- **for** *every* $\sigma_k \in \Sigma_k$ **do**
  - **if** σ_{k-1} *is a face of* σ_k **then**
    - $[\mathbf{B}_k]_{\sigma_{k-1}, \sigma_k} \leftarrow$ ORIENTATION(σ_{k-1}, σ_k)
  - **else**
    - $[\mathbf{B}_k]_{\sigma_{k-1}, \sigma_k} \leftarrow 0$

Return : k-th boundary map $\mathbf{B}_k$

Algorithm S4: Furthest points sampling

Input : Initial point cloud $\tilde{\mathbf{X}} \in \mathbb{R}^{n' \times D}$, number of furthest points n

$\mathbf{X} \leftarrow \emptyset$

Pick a point $\widehat{\mathbf{x}} \in \mathbb{R}^D$ randomly from $\tilde{\mathbf{X}}$

for $i = 1, \cdots, n-1$ **do**
- $\mathbf{X} \leftarrow \mathbf{X} \cup \{\widehat{\mathbf{x}}\}$ ▷ `Add` $\widehat{\mathbf{x}}$ `to` $\mathbf{X}$
- $\tilde{\mathbf{X}} \leftarrow \tilde{\mathbf{X}} \backslash \{\widehat{\mathbf{x}}\}$ ▷ `Remove` $\widehat{\mathbf{x}}$ `from` $\tilde{\mathbf{X}}$
- $\widehat{\mathbf{x}} \leftarrow \operatorname{argmax}_{\mathbf{x} \in \mathbf{X}} \min_{\tilde{\mathbf{x}} \in \tilde{\mathbf{X}}} \|\mathbf{x} - \tilde{\mathbf{x}}\|$
- ▷ `Find the point` $\widehat{\mathbf{x}}$ `in` $\mathbf{X}$ `that is furthest from` $\tilde{\mathbf{X}}$

Return : Point cloud $\mathbf{X} \in \mathbb{R}^{n \times D}$