Density of States of Hodge Laplacians: for Simplicial Sparsification

Tony Savostianov[1] • Francesco Tudisco[2] • Michael Schaub[3] • Nicola Guglielmi[4]

[1] *Computational Network Science, RWTH Aachen , email:* savostianov@cs.rwth-aachen.de
[2] *Maxwell Institute & School of Mathematics, University of Edinburgh, UK , email:* f.tudisco@ed.ac.uk
[3] *Computational Network Science, RWTH Aachen , email:* schaub@cs.rwth-aachen.de
[4] *Gran Sasso Science Institute, L'Aquila, Italy, , email:* nicola.guglielimi@gssi.it

Abstract:

Keywords: density of states, simplicial complexes, sparsification, generalized effective resistance

I Introduction

Well, it would be nice to have one.
You know, somewhere here, maybe...

Related works.

- benson
- osting, speilman
- kyng and fast solvers
- cohen and geoemetrical solvers
- hecs precon

Main Contributions. Our main contributions are as follows:

- we show that sparsification measure (which is necessary to sample sparse spectral approximation of an up-Laplacian operator of simplicial complex) is directly related to local densities of states of a higher-order down-Laplacian;
- we develop a novel method of kernel ignoring decomposition (KID) for efficient approximation of LDoS;
- finally, we exploit existing error estimates to derive KID parameters sufficient for an arbitrary approximation quality and demonstrate the corresponding competitive algorithmic complexity for the method.

Outline. The rest of the paper is organized as follows: we provide the brief introduction of simplicial complexes and its densities of states as a fast computable proxies of spectral information, CREF. CREF states the sparsification result and demonstrates its connection to the local spectral densities with CREF describing a novel approach which efficiently computes the sparsifying probability measure. We provide numerical experiments in CREF, CREF concludes.

II Preliminaries

I. Simplicial Complexes, [Lim20]

Simplicial complexes Let $V = \{v_1, \dots v_{m_0}\}$ be the set of nodes. Each subset of nodes $\sigma = \{v_{i_1}, \dots v_{i_k}\}$ is known as **simplex** of order $k-1$ with all its maximal proper subsets knowns as **faces** of σ. Then a collection of simplices $\mathcal{K}$ is a **simplicial complex** if and only if each simplex σ enters $\mathcal{K}$ with all its faces (inclusion rule). As a result, each simplicial complex can be written as a collections of simplices of fixed orders, $\mathcal{K} = \{\mathcal{V}_0(\mathcal{K}), \mathcal{V}_1(\mathcal{K}), \mathcal{V}_2(\mathcal{K}), \dots\}$ where $\mathcal{V}_0(\mathcal{K})$ is a set of 0-simplices (nodes), $\mathcal{V}_1(\mathcal{K})$ is a set of edges, $\mathcal{V}_2(\mathcal{K})$ of triangles, and so on, with $m_k = |\mathcal{V}_k(\mathcal{K})|$ denoting the number of simplices in the corresponding set.

Boundary operators and Laplacians Due to the inclusion principle, simplices inside $\mathcal{K}$ form a nested structure through boundary relation; specifically, one can map a k-order simplex to a collection of $(k-1)$-order simplices on its border. Formally, let $\mathcal{C}_k$ be a linear space of all formal sums of simplices in $\mathcal{V}_k(\mathcal{K})$; then the boundary map B_k is defined as an alternating sum:

$$B_k : \mathcal{C}_k \mapsto \mathcal{C}_{k-1}, \qquad B_k[v_1, v_2, \dots v_k] = \sum_{i=1}^{k} (-1)^{i-1} [v_1, \dots v_{i-1}, v_{i+1}, \dots v_k] \tag{Eqn. 1}$$

such that boundary of a boundary is 0, $B_k B_{k+1} = 0$, (fundamental lemma of homology). Provided one fixes an order of simplices and their orientation (lexicographically, so the simplicity's sake), one can define the following entities describing the higher-order structure of $\mathcal{K}$:

Def. 1 **(Homology group and higher-order Laplacian)** Since $\operatorname{im} B_{k+1} \subset \ker B_k$, the quotient space $\mathcal{H}_k = {}^{\ker B_k}/_{\operatorname{im} B_{k+1}}$, known as k-th homology group, is correctly defined and the following isomorphisms hold

$$\mathcal{H}_k \cong \ker B_k \cap \ker B_{k+1}^\top \cong \ker \left(B_k^\top B_k + B_{k+1} B_{k+1}^\top \right).$$

The matrix $L_k = B_k^\top B_k + B_{k+1} B_{k+1}^\top$ is called the k-th order **Laplacian operator**; the two terms $L_k^\downarrow = B_k^\top B_k$ and $L_k^\uparrow = B_{k+1} B_{k+1}^\top$ are referred to as the **down-Laplacian** and the **up-Laplacian**, respectively.

The homology group $\mathcal{H}_k$ describes the k-th topology of the simplicial complex $\mathcal{K}$: $\beta_k = \dim \mathcal{H}_k = \dim \ker L_k$ coincides exactly with the number of k-dimensional holes in the complex, known as the **k-th Betti number**. In the case $k = 0$, the operator $L_0 = L_0^\uparrow$ is exactly the classical graph Laplacian whose kernel corresponds to the *connected components* of the graph, while $L_0^\downarrow = 0$. For $k = 1$ and $k = 2$, the elements of $\ker L_1$ and $\ker L_2$ describe the simplex 1-dimensional holes and voids respectively, and are frequently used in the analysis of trajectory data, [SBH+20, BGL16]. redo

Weighted case and spectral structures Considering weighted generalization of simplicial complexes, one needs to preserve the fundamental lemma of homology, $B_k B_{k+1} = 0$. As a result, let us assume the collection of weight matrices $W_k \in \mathbb{R}_{m_k \times m_k}$ such that W_k is diagonal with each $(W_k)_{ii} > 0$ containing the weight of the i-the simplex in $\mathcal{V}_k(\mathcal{K})$. Then a generic weighting scheme for boundary operators can be written as:

$$B_k \mapsto W_{k-1}^{-1} B_k W_k \tag{Eqn. 2}$$

Note that with such weighting scheme the dimensionality of the homology group is preserved, $\dim\ker L_k = \dim\ker \widehat{L}_k$, [GST23].

At each order k, weighted boundary operators B_k define the space decomposition (Hodge decomposition):

$$\mathbb{R}^{m_k} = \overbrace{\operatorname{im} B_k^\top \oplus \underbrace{\ker\left(B_k^\top B_k + B_{k+1}B_{k+1}^\top\right)}_{\ker B_k} \oplus \operatorname{im} B_{k+1}}^{\ker B_{k+1}^\top} \tag{Eqn. 3}$$

implying that that kernels of up- and down-Laplacians are necessarily non-trivial and high-dimensional, $\ker L_k^\downarrow \supseteq \operatorname{im} B_{k+1}$ and $\ker L_k^\uparrow \subseteq \operatorname{im} B_k^\top$. More explicit relation between these subspaces is known as a principle spectral inheritance, [GST23].

II. Networks' Density of States, [DBB19]

Spectral densities As we established above, the spectral structure of the higher-order Laplacian operators carry meaningful information of the complex's topological structure, YADA YADA ——— the usual song

At the same time, the overall cost of extracting full spectral information from corresponding Laplacian operators L_k is computationally expensive with a complexity $\mathcal{O}(m_k^3)$. Instead, one can consider the following functional description:

Def. 2 **(Density of States)** Fro a given symmetric matrix $A = Q\Lambda Q^\top$ with $Q^\top Q = I$ and diagonal $\Lambda = \operatorname{diag}(\lambda_1, \dots \lambda_n)$, the spectral density or density of states (DoS)

$$\mu(\lambda \mid A) = \frac{1}{n}\sum_{i=1}^{n} \delta(\lambda - \lambda_i) \tag{Eqn. 4}$$

Additionally, let $\mathbf{q}_i$ be a corresponding unit eigenvector of A (such that $A\mathbf{q}_i = \lambda_i \mathbf{q}_i$ and $Q = (\mathbf{q}_1 \mid \mathbf{q}_2 \mid \cdots \mid \mathbf{q}_n)$); then one can define a set of local (entry-wise) densities of states (LDoS):

$$\mu_k(\lambda \mid A) = \sum_{i=1}^{n} \left|\mathbf{e}_k^\top \mathbf{q}_i\right|^2 \delta(\lambda - \lambda_i) \tag{Eqn. 5}$$

with $\mathbf{e}_k$ being the corresponding versor.

Here DoS function $\mu(\lambda \mid A)$ contains the overall spectrum of the operator A while the family of LDoS $\mu_k(\lambda \mid A)$ describe the contribution of the simplex $\sigma_k \in \mathcal{V}_k(\mathcal{K})$ to the spectral information.

Finally, one should note that by its definition DoS and LDoS are generalized functions, hence the quality of their computation is difficult to assess directly; instead, one considers its histogram representations:

$$h_i = \int_{x_i}^{x_i+\Delta_x} \mu(\lambda \mid A)d\lambda, \qquad h_i^{(k)} = \int_{x_i}^{x_i+\Delta_x} \mu_k(\lambda \mid A)d\lambda \tag{Eqn. 6}$$

which correspond to discretized output of the convolution of spectral densities with a smooth approximation of identity function $K_{\Delta h}$, $[\mu(\lambda \mid A) * K_{\Delta h}]$.

Kernel Polynomial Method (KPM) By its very definition, the introduction of the spectral densities $\mu(\lambda \mid L_k)$ and $\{\mu_k(\lambda \mid A)\}$ requires the complete spectral information of the original operator L_k and, hence, is not immediately

computationally more efficient. At the same time, one can use the functional nature of DoS and LDoS to obtain approximation through expansion.

Namely, let $T_m(x)$ be Chebyhev polynomials of the first kind. Specifically,

$$T_0(x) = 1,\ T_1(x) = x, \qquad T_{m+1}(x) = 2xT_m(x) - Tm - 1(x) \tag{Eqn. 7}$$

which as an orthonormal (with respect to the weight function $w(x) = \frac{2}{(1+\delta_{0m})\pi\sqrt{1-x^2}}$) on $[-1, 1]$. Let us transform the operator $A \to H$ such that $\sigma(H) \in [-1, 1]$ (e.g. $H = \frac{2}{\lambda_{\max}} A - I$); then one can decompose

$$\mu(\lambda \mid A) = \sum_{m=0}^{\infty} d_m w(\lambda) T_m(\lambda), \qquad \mu_k(\lambda \mid A) = \sum_{m=0}^{\infty} d_{mk} w(\lambda) T_m(\lambda) \tag{Eqn. 8}$$

Due to the orthonormality,

$$d_m = \langle \mu(\lambda \mid A), T_m(x) \rangle = \sum_{\lambda_i \in \sigma(A)} T_m(\lambda_i) = \operatorname{tr} T_m(A), \quad d_{m\bullet} = \operatorname{diag}(T_m(A)) \tag{Eqn. 9}$$

Note that instead of computing $T_m(A)$ one can use Monte-Carlo estimations for the trace and diagonal, specifically:

$$\operatorname{tr} X = \mathbb{E}\left[\mathbf{z}^\top X \mathbf{z}\right], \qquad \operatorname{diag} X = \mathbb{E}\left[\mathbf{z} \odot X\mathbf{z}\right], \tag{Eqn. 10}$$

where $\mathbf{z}$ is a vector of i.i.d. random variables with zero mean and unit variance. The MC-sampling reduces $\operatorname{tr} T_m(A)$ and $\operatorname{diag}(T_m * (A))$ calculations to simple matvec operations which can be efficiently updates for the next values of m due to the recurrent definition of $T_m(x)$.

KPM approximation of DoS/LDoS relies on two introduced errors: reduction of the expansion to the finite sum of M terms and sampling of the moments d_m. The computational cost of the approximation is then fixed to $\mathcal{O}(N_z M \operatorname{nnz}(A))$ where $\mathcal{O}(\operatorname{nnz}(A))$ is the cost of a matvec operation for the operator A.

III Sparsification of Simplicial Complexes

Th III.II.1 **(Simplicial Sparsification, [OPW22])** Let $\mathcal{K}$ be a simplicial complex restricted to its p-skeleton, $\mathcal{K} = \bigcup_{i=0}^{p} \mathcal{V}_i(\mathcal{K})$. Let $L_k^{\uparrow}(\mathcal{K})$ be its k-th up-Laplacian and let $m_k = |\mathcal{V}_k(\mathcal{K})|$. For any $\varepsilon > 0$, a sparse simplicial complex $\mathcal{L}$ can be sampled as follows:

(1) compute the probability measure $\mathbf{p}$ on $\mathcal{V}_{k+1}(\mathcal{K})$ proportional to the generalized resistance vector $\mathbf{r} = \operatorname{diag}\left(B_{k+1}^\top (L_k^{\uparrow})^\dagger B_{k+1}\right)$, where $\operatorname{diag}(A)$ denotes the vector of the diagonal entries of A;

(2) sample q simplices τ_i from $\mathcal{V}_{k+1}(\mathcal{K})$ according to the probability measure $\mathbf{p}$, where q is chosen so that $q(m_k) \geq 9C^2 m_k \log(m_k/\varepsilon)$, for some absolute constant $C > 0$;

(3) form a sparse simplicial complex $\mathcal{L}$ with all the sampled simplexes of order k and all its faces with the weight $\frac{w_{k+1}(\tau_i)}{q(m_k)\mathbf{p}(\tau_i)}$; weights of repeated simplices are accumulated.

Then, with probability at least $1/2$, the up-Laplacian of the sparsifier $\mathcal{L}$ is

ε-close to the original one, i.e. it holds $L_k^{\uparrow}(\mathcal{L}) \underset{\varepsilon}{\approx} L_k^{\uparrow}(\mathcal{K})$.

The bottleneck of the subsampling above is the construction of the appropriate measure **p** or, more precisely, generalized effective resistance **r**. Indeed, one needs a fast pseudo-inverse operator $\left(L_k^{\uparrow}\right)^{\dagger}$ in order to compute **r**.

Rem II.1 **(Sensitivity of the Sparsification vis-a-vis sampling measure p)** Since **p** is a probability measure on m_{k+1} simplices, it is natural consider size of its perturbation in terms of $\frac{1}{m_{k+1}}$.

we need to assume somewhere $W_k = I$ for simplicity and do not forget about it

Th III.II.2 **(GER through LDoS)** For a given simplicial complex mcK with the k-th order up-Laplacian $L_k^{\uparrow} = B_{k+1} W_{k+1}^2 B_{k+1}^{\top}$, a generalized effective resistance r can be computed through family of local densities of states $\{\mu_i(\lambda \mid L_{k+1}^{\downarrow})\}$:

$$\mathbf{r}_i = \int_{\mathbb{R}} (1 - \mathbb{1}_0(\lambda)) \mu_i(\lambda \mid L_{k+1}^{\downarrow}) d\lambda \qquad \text{(Eqn. 11)}$$

IV KID for LDoS

Note that up- and down-Laplacians $L_k^{\uparrow}$ and $L_k^{\downarrow}$ by their definition have non-trivial kernels of high dimensionality. Indeed, recalling the Hodge decomposition :

ref up

$$\mathbb{R}^{m_k} = \overbrace{\operatorname{im} B_k^{\top} \oplus \underbrace{\ker\left(B_k^{\top} B_k + B_{k+1} B_{k+1}^{\top}\right)}_{\ker B_k} \oplus \operatorname{im} B_{k+1}}^{\ker B_{k+1}^{\top}} \qquad \text{(Eqn. 12)}$$

the subspaces $\ker B_k = \ker L_k^{\downarrow}$ and $\ker B_{k+1}^{\top} = \ker L_k^{\uparrow}$ include at least $\operatorname{im} B_{k+1}$ and $\operatorname{im} B_{k+1}^{\top}$. As a result, the zero eigenvalue in the corresponding DoS and LDoS for both operators exhibits a dominating pick severely affecting the quality of KPM-approximation.

do we need an illustration

Rem II.2 **(Filtration of the kernels)** As given by the Hodge decomposition above, the most part of the peak in $L_k^{\downarrow}$ is explained by the elements of $\operatorname{im} B_{k+1}$. Then one can avoid the dominating kernel pick in the corresponding LDoS family $\{\mu_k(\lambda \mid L_k^{\downarrow})$ by filtering $\operatorname{im} B_{k+1}$ out provided one can obtain an orthonormal basis of this subspace. As a result, the filtration of the kernel becomes the question of the efficient range finder of a sparse operator $\operatorname{im} B_{k+1}$. Although we do not claim that computationally efficient range finder for such case does not exist, the majority of the methods at the current moment run into the bottleneck of the QR-decomposition which is far more computationally complex then one can allow in comparison with the straightforward computation of GER for $L_k^{\uparrow}$.

Theorem III.II.2 implies that one indeed does not need the full profile $\mu_k(\lambda \mid L_{k+1}^{\downarrow})$, but LDoS ignoring the kernel, $(1 - \mathbb{1}_0)\mu_k(\lambda \mid L_{k+1}^{\downarrow})$. As a result, one can attempt to construct an approximation different from Chebyshev KPM

that avoids the kernel of the operator in its entirety.

V Conclusion

Motive filtration is still a thing

V References

[BGL16] Austin R Benson, David F Gleich, and Jure Leskovec. Higher-order organization of complex networks. *Science*, 353(6295):163–166, 2016.

[BGVKS23] Jess Banks, Jorge Garza-Vargas, Archit Kulkarni, and Nikhil Srivastava. Pseudospectral shattering, the sign function, and diagonalization in nearly matrix multiplication time. *Foundations of computational mathematics*, 23(6):1959–2047, 2023.

[DBB19] Kun Dong, Austin R Benson, and David Bindel. Network density of states. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 1152–1161, 2019.

[GST23] Nicola Guglielmi, Anton Savostianov, and Francesco Tudisco. Quantifying the Structural Stability of Simplicial Homology. *Journal of Scientific Computing*, 97(1):2, August 2023.

[Lim20] Lek-Heng Lim. Hodge laplacians on graphs. *Siam Review*, 62(3):685–715, 2020.

[OPW22] Braxton Osting, Sourabh Palande, and Bei Wang. Spectral sparsification of simplicial complexes for clustering and label propagation. *Journal of computational geometry*, 11(1), 2022.

[SBH+20] Michael T Schaub, Austin R Benson, Paul Horn, Gabor Lippner, and Ali Jadbabaie. Random walks on simplicial complexes and the normalized Hodge 1-Laplacian. *SIAM Review*, 62(2):353–391, 2020.

III. Appendix

III.I Simplicial complexes

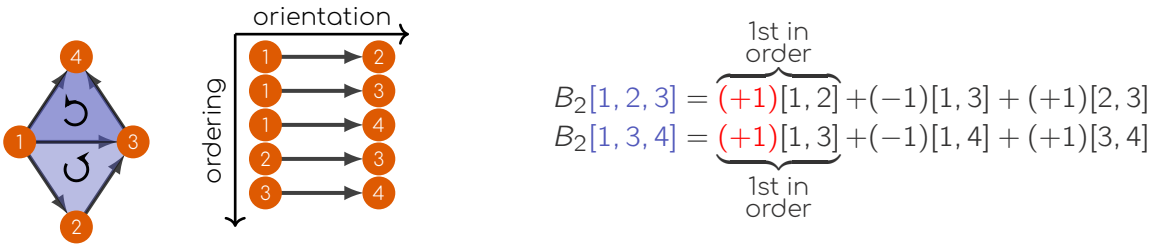

Figure III.1: Example of the simplicial complex with ordering and orientation: nodes from $\mathcal{V}_0(\mathcal{K})$ in orange, triangles from $\mathcal{V}_2(\mathcal{K})$ in blue. Orientation of edges and triangles is shown by arrows; the action of B_2 operator is given for both triangles.

III.II Proof of Theorem 1

Proof Let $B_{k+1}W_{k+1} = USV^\top$ where S is diagonal and invertible and both U and V are orthogonal (so it is a truncated SVD decomposition of $B_{k+1}W_{k+1}$ matrix with eliminated obsolete kernel). Then:

$$(L_k^{\uparrow})^\dagger = \left(B_{k+1}W_{k+1}^2B_{k+1}^\top\right)^\dagger = \left(US^2U^\top\right)^\dagger = US^{-2}U^\top \qquad \text{(Eqn. 13)}$$

$$\begin{aligned}\mathbf{r} &= \operatorname{diag}\left(W_{k+1}B_{k+1}^{\top}(L_k^{\uparrow})^{\dagger}B_{k+1}W_{k+1}\right) = \\ &= \operatorname{diag}\left(VSU^{\top}US^{-2}U^{\top}USV^{\top}\right) = \operatorname{diag}\left(VV^{\top}\right)\end{aligned} \tag{Eqn. 14}$$

As a result, $\mathbf{r}_i = \|V_{i\cdot}\|^2 = \sum_j |v_{ij}|^2$, so the i-th entry of the resistance is defined by the sum of square of i-th components of eigenvectors $\mathbf{v}_j$ of $L_{k+1}^{\downarrow} = W_{k+1}B_{k+1}^{\top}B_{k+1}W_{k+1}$ operator where $\mathbf{v}_j \perp \ker L_{k+1}^{\downarrow}$.
Note that

$$\mu_i(\lambda \mid L_{k+1}^{\downarrow}) = \sum_{j=1}^{m_{k+1}} \left|\mathbf{e}_i^{\top}\mathbf{q}_j\right|^2 \delta\left(\lambda - \lambda_j\right) = \sum_{j=1}^{m_{k+1}} \left|q_{ij}\right|^2 \delta\left(\lambda - \lambda_j\right) \tag{Eqn. 15}$$

so

$$\begin{aligned}\mathbf{r}_i = \|V_{i\cdot}\|^2 = \sum_j |v_{ij}|^2 = \int_{\mathbb{R}\setminus\{0\}} \sum_{j=1}^{m_{k+1}} \left|q_{ij}\right|^2 \delta\left(\lambda - \lambda_j\right) d\lambda = \\ = \int_{\mathbb{R}\setminus\{0\}} \mu_i(\lambda \mid L_{k+1}^{\downarrow}) d\lambda = \int_{\mathbb{R}} (1 - \mathbb{1}_0(\lambda))\mu_i(\lambda \mid L_{k+1}^{\downarrow}) d\lambda\end{aligned} \tag{Eqn. 16}$$

■