Learning Dirac Spectral Transforms for Topological Signals

Leonardo Di Nino, Tiziana Cattai,
Sergio Barbarossa, Ginestra Bianconi, Paolo Di Lorenzo

Dept. of Information Engineering, Electronics, and Telecommunications, Sapienza, Rome
School of Mathematical Sciences, Queen Mary University of London
To appear in *Proc. of EUSIPCO 2026*

June 08, 2026

Outline

Introduction

Interlacing Domains Over Networks

Introduction

- Graph Signal Processing methods mainly revolve around the definition of **shift operators** as **the node Laplacian**;
- Topological Signal Processing as well is built around **higher-order operators emboding discrete Hodge theory**;
- **Laplacian-based signal processing works in one domain at a time**, while we may benefit by considering network signals as a whole undergoing coupling rules: the Dirac operator of networks[1] encodes **first order differential relations between domains**.

---

[1] Bianconi, G., "The topological Dirac equation of networks and simplicial complexes." Journal of Physics: Complexity 2.3 (2021): 035022.

Dirac Operator of Networks

Introduction

- The Dirac operators allows the implementation of holistic processing algorithms:
  - Network dynamics [2],[3];
  - Topological Data Analysis[4];
  - Topological Filtering[5];
  - Topological Neural Networks[6],[7]

---

[2] Calmon, L., et al., "Local Dirac synchronization on networks." Chaos: An Interdisciplinary Journal of Nonlinear Science 33.3 (2023).

[3] Carletti, T., et al., "Global topological Dirac synchronization." Journal of Physics: Complexity 6.2 (2025): 025009.

[4] Wee, J., et al., "Persistent Dirac for molecular representation." Scientific Reports 13.1 (2023): 11183.

[5] Calmon, L., et al., "Dirac signal processing of higher-order topological signals." New Journal of Physics 25.9 (2023): 093013.

[6] Nauck, C., et al. "Dirac–Bianconi Graph Neural Networks-Enabling long-range graph predictions." ICML 2024 Workshop on geometry-grounded representation learning and generative modeling. 2024.

[7] Battiloro, Claudio, et al., "Generalized simplicial attention neural networks." IEEE Transactions on Signal and Information Processing over Networks 10 (2024): 833-850.

Research Question and Contributions

Introduction

- The limits and use cases of Dirac-based learning remain largely underinvestigated;
- We place Laplacian-based and Dirac-based methods within a **common unifying framework** to systematically study their relations.

We study the **spectral localization** *properties of the Dirac and Laplacian operators and propose a* **Mixture-of-Dirac transform** *encompassing both operators as* **limiting cases***: it enables a* **data-driven transform learning method** *for topological signals, interpolating between* **domain coupling** *and* **domain uncoupling**.

Dirac and Laplacian Representations

Topological Signals and Discrete Calculus

Dirac and Laplacian Representations

- We consider a graph $\mathcal{G}(\mathcal{V},\mathcal{E})$ with $|\mathcal{V}| = V$ nodes and $|\mathcal{E}| = E$ edges, and node/edge signals $\mathbf{x}_0 \in \mathcal{C}^0 \cong \mathbb{R}^V$, $\mathbf{x}_1 \in \mathcal{C}^1 \cong \mathbb{R}^E$. A *topological spinor* $\mathbf{s}$ is the composite vector:

$$\mathbf{s} = \begin{pmatrix} \mathbf{x}_0 \\ \mathbf{x}_1 \end{pmatrix} \in \mathcal{C}^0 \oplus \mathcal{C}^1 \cong \mathbb{R}^{V+E}.$$

- The two domains are related by the *discrete gradient* $\delta : \mathcal{C}^0 \to \mathcal{C}^1$ and *divergence* $\delta^* : \mathcal{C}^1 \to \mathcal{C}^0$, acting element-wise as

$$(\delta \mathbf{x}_0)_e = \mathbf{x}_0[v] - \mathbf{x}_0[u], \quad \forall e : u \trianglelefteq e \trianglerighteq v$$

$$(\delta^* \mathbf{x}_1)_v = -\sum_{e : v \trianglelefteq e} \mathbf{x}_1[e] + \sum_{e : v \trianglerighteq e} \mathbf{x}_1[e], \quad \forall v \in \mathcal{V},$$

where $\trianglelefteq, \trianglerighteq$ denote oriented incidence relations.

From Discrete Calculus to Laplacians

Dirac and Laplacian Representations

- The linear operators δ, δ^* can be algebraically represented by the incidence matrix ($\mathbf{B}[n,e] = 1$ for $n \trianglerighteq e$, -1 for $n \trianglelefteq e$, 0 otherwise), so that $\delta \mathbf{x}_0 = \mathbf{B}^\top \mathbf{x}_0$ and $\delta^* \mathbf{x}_1 = \mathbf{B}\mathbf{x}_1$.
- Their composition yields the graph Laplacian (divergence of the gradient) and the 1-Hodge Laplacian (gradient of the divergence):

  $$\mathbf{L}_0 \mathbf{x}_0 = \mathbf{B}\mathbf{B}^\top \mathbf{x}_0$$

  $$\mathbf{L}_1 \mathbf{x}_1 = \mathbf{B}^\top \mathbf{B}\mathbf{x}_1$$

  Their spectra reveal topological structure and enable domain-aware processing.
- However, each Laplacian acts on a single domain at a time, whereas the **underlying constitutive relation suggests joint representation across contiguous domains**.

Dirac Operator: definition and properties

Dirac and Laplacian Representations

- The Dirac operator of networks couples the introduced differential relations. The operator $\mathbf{D} : \mathcal{C}^0 \oplus \mathcal{C}^1 \rightarrow \mathcal{C}^0 \oplus \mathcal{C}^1$ maps one domain into the other, and its square recovers $\mathbf{L}_{\mathcal{G}} = \mathrm{blkdiag}(\mathbf{L}_0, \mathbf{L}_1)$:

$$\mathbf{D}\mathbf{s} = \begin{pmatrix} \mathbf{0} & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{0} \end{pmatrix} \mathbf{s} = \begin{pmatrix} \mathbf{B}\mathbf{x}_1 \\ \mathbf{B}^\top \mathbf{x}_0 \end{pmatrix}, \quad \mathbf{D}^2 \mathbf{s} = \mathbf{L}_{\mathcal{G}} \mathbf{s} = \begin{pmatrix} \mathbf{L}_0 \mathbf{x}_0 \\ \mathbf{L}_1 \mathbf{x}_1 \end{pmatrix}.$$

- The Dirac operator is an anticommutant of $\mathbf{\Gamma} = \mathrm{blkdiag}(\mathbf{I}_V, -\mathbf{I}_E)$:

$$\mathbf{D}\mathbf{\Gamma} + \mathbf{\Gamma}\mathbf{D} = 0$$

This implies the following **chiral** eigenpairs:

$$\mathbf{D}\phi = \sigma\phi$$

$$\mathbf{D}\mathbf{\Gamma}\phi = -\sigma\mathbf{\Gamma}\phi$$

Spectral Structure of Dirac and Laplacian operators

Dirac and Laplacian Representations

- Both $\mathbf{D}$ and $\mathbf{L}_{\mathcal{G}}$ are self-adjoint, so their eigenbases form complete orthonormal bases for spinors, inducing Fourier-like spectral representations.
- Letting $\mathbf{U}$ ($\mathbf{U}_{\mathrm{H}}$) and $\mathbf{V}$ ($\mathbf{V}_{\mathrm{H}}$) collect the left and right singular vectors of $\mathbf{B}$ associated with its nonzero (zero) singular values. The Dirac operator and the super-Laplacian then admit the eigendecompositions $\mathbf{D} = \boldsymbol{\Phi}\boldsymbol{\Gamma}\boldsymbol{\Phi}^{\top}$ and $\mathbf{L}_{\mathcal{G}} = \boldsymbol{\Theta}\boldsymbol{\Lambda}\boldsymbol{\Theta}^{\top}$, with

$$\boldsymbol{\Phi} = \begin{pmatrix} \mathbf{U} & \mathbf{0} & \mathbf{U}_{\mathrm{H}} & \mathbf{U} \\ -\mathbf{V} & \mathbf{V}_{\mathrm{H}} & \mathbf{0} & \mathbf{V} \end{pmatrix}, \quad \boldsymbol{\Theta} = \begin{pmatrix} \mathbf{U}_{\mathrm{H}} & \mathbf{0} & \mathbf{U} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_{\mathrm{H}} & \mathbf{0} & \mathbf{V} \end{pmatrix} \tag{1}$$

$$\boldsymbol{\Gamma} = \mathrm{blkdiag}\big(-\boldsymbol{\Sigma}, \mathbf{0}_{\xi_1}, \mathbf{0}_{\xi_0}, \boldsymbol{\Sigma}\big) \tag{2}$$

$$\boldsymbol{\Lambda} = \mathrm{blkdiag}\big(\mathbf{0}_{\xi_0}, \mathbf{0}_{\xi_1}, \boldsymbol{\Sigma}^2, \boldsymbol{\Sigma}^2\big) \tag{3}$$

where $\boldsymbol{\Sigma}$ holds the nonnull singular values of $\mathbf{B}$, $\xi_0 = \dim(\ker(\mathbf{B}))$, and $\xi_1 = \dim(\ker(\mathbf{B}^{\top}))$.

Localization of Topological Signals

Dirac and Laplacian Representations

PERFECTLY COUPLED

Sparse in Dirac modes

- ✓ 2× More efficient under Dirac
- × 2× Costlier under Laplacian

PERFECTLY UNCOUPLED

Sparse in Laplacian modes

- × 2× Costlier under Dirac
- ✓ 2× More efficient under Laplacian

The Parseval frame $\mathbf{F} = (\mathbf{\Phi} \mid \mathbf{\Theta}) \in \mathbb{R}^{(V+E)\times 2(V+E)}$ optimally represents either perfectly coupled, or perfectly uncoupled signals, and superposition of these two classes, **at the cost of redundancy**.

Topological signals might show **degrees of coupling happening at diverse scales**, admitting no clean separation between coupled and uncoupled components: **Dirac equation of networks** allows to model them.

Dirac-Driven Transform Learning

Dirac Equation of Networks

Dirac-Driven Transform Learning

- The Dirac equation is the topological wave equation:

$$\mathrm{i}\partial_t \mathbf{s} - \mathcal{H}\mathbf{s} = 0, \tag{4}$$

where i denotes the imaginary unit, $\mathcal{H} = \mathbf{D} + m \begin{pmatrix} \mathbf{I}_V & \mathbf{0} \\ \mathbf{0} & -\mathbf{I}_E \end{pmatrix}$ and m parametrizes the *mass* [8].

- The eigenstates ψ of $\mathcal{H}$, satisfying the equation $\mathcal{H}\psi = \varrho\psi$, exhibit the following structure for each nonzero graph frequency λ_i:

$$\psi_i^- = \zeta^- \begin{pmatrix} \frac{\lambda_i}{|\varrho_i|+m}\mathbf{u}_i \\ -\mathbf{v}_i \end{pmatrix}, \quad \psi_i^+ = \zeta^+ \begin{pmatrix} \mathbf{u}_i \\ \frac{\lambda_i}{|\varrho_i|+m}\mathbf{v}_i \end{pmatrix}, \tag{5}$$

where $\zeta^\pm$ are constants used to have unit norm vectors.

---

[8] Bianconi, G., "The mass of simple and higher-order networks." Journal of Physics A: Mathematical and Theoretical 57.1 (2024): 015001.

Interpolating between coupling and uncoupling

Dirac-Driven Transform Learning

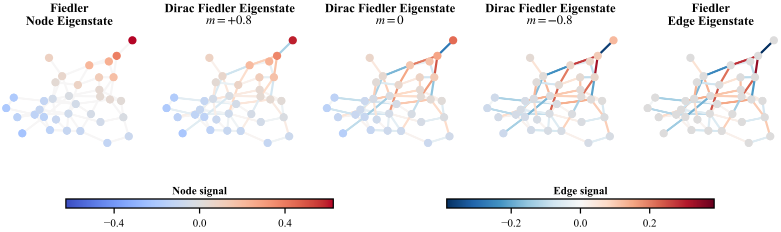

Figure: Effect of the mass parameter on Dirac eigenstates. The outer panels show the node (left) and edge (right) eigenstates at the smallest non-zero graph frequency. The central panel shows the Dirac eigenvector for the smallest positive eigenvalue. Intermediate panels trace the interpolation between coupled and uncoupled regimes as the mass varies.

Spectral Transform Based on Dirac Equation

Dirac-Driven Transform Learning

- We propose a **mass-parameterized Fourier-like transform** $\mathcal{T}_{\mathbf{D},\mathbf{m}}$ encoding mode-dependent coupling between node and edge components.
- Let $\mathbf{m} = [m_1, \ldots, m_r]^\top \in \mathbb{R}^r$, with $r = \frac{1}{2}[V + E - (\xi_0 + \xi_1)]$, assign a scalar mass parameter to each chiral mode pair.
- The operator $\overline{\boldsymbol{\Psi}}(\mathbf{m}) = (\boldsymbol{\Psi}^-(\mathbf{m}) \mid \boldsymbol{\Psi}_{\mathrm{H}} \mid \boldsymbol{\Psi}^+(\mathbf{m}))$ collects Dirac equation eigenstates, replacing each mode's mass with the corresponding m_i.
- The forward and inverse transforms read

$$\mathcal{T}_{\mathbf{D},\mathbf{m}}(\mathbf{s}) = \overline{\boldsymbol{\Psi}}(\mathbf{m})^\top \mathbf{s}, \qquad \mathcal{T}^{-1}_{\mathbf{D},\mathbf{m}}(\hat{\mathbf{s}}) = \overline{\boldsymbol{\Psi}}(\mathbf{m})\, \hat{\mathbf{s}}.$$

- Symmetric mass assignment within each chiral pair preserves orthogonality, yielding a complete basis for spinors — with the **Dirac** and **Laplacian eigenbases** as the two limiting regimes ($m_i = 0$ and $m_i \to \infty$ respectively).

Dirac Driven Transform Learning

Dirac-Driven Transform Learning

- We postulate the following signal model:

$$\mathbf{s} = \overline{\boldsymbol{\Psi}}(\mathbf{m})\boldsymbol{\omega} + \mathbf{w}, \tag{6}$$

  where $\{\omega_i\}_{i\in\mathcal{I}}$ are the sparse expansion coefficients and $\mathbf{w}$ is a zero-mean white Gaussian noise vector with covariance matrix $\sigma^2\mathbf{I}_{V+E}$.
- Let us assume to collect T topological spinors arranged in the data matrix $\mathbf{S} = [\mathbf{s}_1, \ldots, \mathbf{s}_T]$, and denote by $\boldsymbol{\Omega} = [\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_T]$ their corresponding coefficient vectors in the parametrized basis to be learned.
- Under the considered model, the maximum likelihood estimator yields the following least-squares problem:

$$\min_{\mathbf{m},\boldsymbol{\Omega}} \quad \|\mathbf{S} - \overline{\boldsymbol{\Psi}}(\mathbf{m})\boldsymbol{\Omega}\|_{\mathrm{F}}^2. \tag{7}$$

Parametrization

Dirac-Driven Transform Learning

- **Linear reparametrization:** learn the correcting effect $k_i = \frac{\lambda_i}{|\varrho_i|+m_i}$ instead of $\mathbf{m}$;
- **Group sparsity:** restrict codes to $\mathcal{B}_{\eta_0} = \{\boldsymbol{\Omega} : \|\boldsymbol{\Omega}\|_{2,0} = \eta_0\}$;
- **No scale ambiguity:** impose $\overline{\boldsymbol{\Psi}}(\mathbf{m}) \in \mathrm{Ob}(V+E, V+E)$;
- **Decoupling:** transfer structure to splitting variables $\mathbf{P}, \mathbf{X}$.
- By additionally defining a feasible set for $\mathbf{k}$, we finally obtain the following optimization problem:

$$\begin{aligned} \min_{\mathbf{k},\boldsymbol{\Omega},\mathbf{P},\mathbf{X}} \quad & \|\mathbf{S} - \overline{\boldsymbol{\Psi}}(\mathbf{k})\boldsymbol{\Omega}\|_{\mathrm{F}}^2 \\ \text{s.t.} \quad & -c_2\mathbf{1} \leq \mathbf{k} \leq c_1\mathbf{1} \\ & \mathbf{X} \in \mathcal{B}_{\eta_0} \\ & \mathbf{P} \in \mathrm{Ob}(V+E, V+E) \\ & \mathbf{P} - \overline{\boldsymbol{\Psi}}(\mathbf{k}) = \mathbf{0} \\ & \mathbf{X} - \boldsymbol{\Omega} = \mathbf{0} \end{aligned} \tag{P1}$$

Alternating Solution Scheme

Dirac-Driven Transform Learning

To solve it, we resort to an ADMM approach, based on the constrained minimization of the augmented Lagrangian:

$$\mathcal{L} = \|\mathbf{S} - \overline{\mathbf{\Psi}}(\mathbf{k})\mathbf{\Omega}\|_{\mathrm{F}}^2 + \frac{\rho_1}{2}\|\overline{\mathbf{\Psi}}(\mathbf{k}) - \mathbf{P} + \mathbf{H}\|_{\mathrm{F}}^2 + \frac{\rho_2}{2}\|\mathbf{\Omega} - \mathbf{X} + \mathbf{M}\|_{\mathrm{F}}^2.$$

Mass scalings (QP)

$$\mathbf{k}^{t+1} = \underset{-c_2\mathbf{1}\le\mathbf{k}\le c_1\mathbf{1}}{\arg\min} \|\mathbf{S} - \overline{\mathbf{\Psi}}(\mathbf{k})\mathbf{\Omega}\|_{\mathrm{F}}^2 + \tfrac{\rho_1}{2}\|\overline{\mathbf{\Psi}}(\mathbf{k}) - \mathbf{P} + \mathbf{H}\|_{\mathrm{F}}^2$$

Sparse Coding

$$\mathbf{\Omega}^{t+1} = \left[2\,\overline{\mathbf{\Psi}}(\mathbf{k})^{\top}\overline{\mathbf{\Psi}}(\mathbf{k}) + \rho_2\mathbf{I}\right]^{-1}\left[\rho_2(\mathbf{X} - \mathbf{M}) + \overline{\mathbf{\Psi}}^{\top}(\mathbf{k})\mathbf{S}\right]$$

Oblique Retraction

$$\mathbf{P}^{t+1} = \mathcal{R}_{\mathrm{Ob}}\{\mathbf{H} + \overline{\mathbf{\Psi}}(\mathbf{k})\}$$

Hard Row-Sparsity

$$\mathbf{X}^{t+1} = \mathcal{R}_{\mathcal{B}_{\eta_0}}\{\mathbf{\Omega} + \mathbf{M}\}$$

Dual Updates

$$\mathbf{H}^{t+1} = \mathbf{H} + \overline{\mathbf{\Psi}}(\mathbf{k}) - \mathbf{P}, \quad \mathbf{M}^{t+1} = \mathbf{M} + \mathbf{\Omega} - \mathbf{X}$$

Sparse Representation

Dirac-Driven Transform Learning

- Sparse-representation experiment on an Erdős–Rényi random graph with $V = 40$ nodes and $E = 80$ edges.
- We consider four coupling regimes, parametrized by the mode-wise mass $\{m_i\}_{i \in [r]}$:
  - *Fully coupled* (Dirac regime): $m_i = 0, \ \forall i \in [r]$;
  - *Fully decoupled* (Laplacian regime): $m_i \gg 0, \ \forall i \in [r]$;
  - *Partially coupled* (Dirac–Laplacian regime): $m_i = 0$ for $i \in \mathcal{I}_{\mathcal{C}}$ and $m_j \gg 0$ for $j \in \mathcal{I}_{\mathcal{U}}$, with $\mathcal{I}_{\mathcal{C}} \cup \mathcal{I}_{\mathcal{U}} = [r]$;
  - *Mixture-of-Dirac*: a non-trivial mass profile across modes given by a Cauchy-like radial basis kernel centered at $i = 0$,

    $$m_i = \frac{m_0}{1 + (i/\gamma)^2}, \qquad i \in [r],$$

    where m_0 sets the peak mass and γ the spread.
- We average 10 realizations, each learning over 600 signals with OMP projection.

Sparse Representation: results

Dirac-Driven Transform Learning

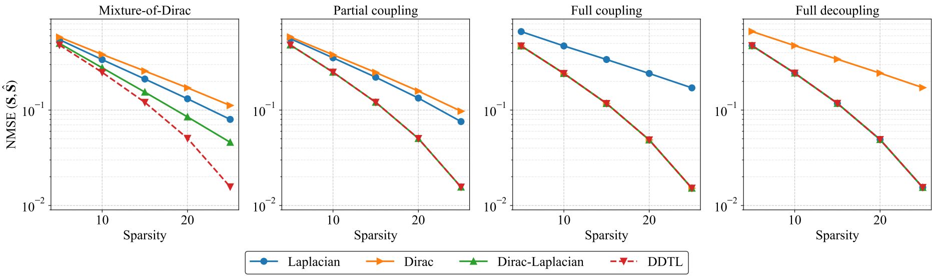

Figure: Synthetic results for sparsity-reconstruction trade-off with different classes of signals on different bases

Denoising

Dirac-Driven Transform Learning

- **Setting:** Water distribution networks (WDNs)
  - Node signals = pressure, edge signals = flow;
- **Data:** Anytown network, $V = 22$ nodes, $E = 41$ edges
  - First $T = 240$ pressure & flow-rate measurements[9];
- **Corruption:** ground truth $\mathbf{S}$ + AWGN at various SNR levels
  - Solve DDTL across bandwidths; optimal filtered signal $\overline{\boldsymbol{\Psi}}(\mathbf{k}^*)\boldsymbol{\Omega}^* \approx \mathbf{S}$;
- **Baseline:** Compare against iterative Dirac filtering (IDESP)[10]

---

[9] Tello, Andrés, et al. "Large-scale multipurpose benchmark datasets for assessing data-driven deep learning approaches for water distribution networks." *Engineering Proceedings* 69.1 (2024): 50.

[10] Wang, Runyue, et al. "Dirac-equation signal processing: Physics boosts topological machine learning." *PNAS nexus* 4.5 (2025): pgaf139.

Denoising: results

Dirac-Driven Transform Learning

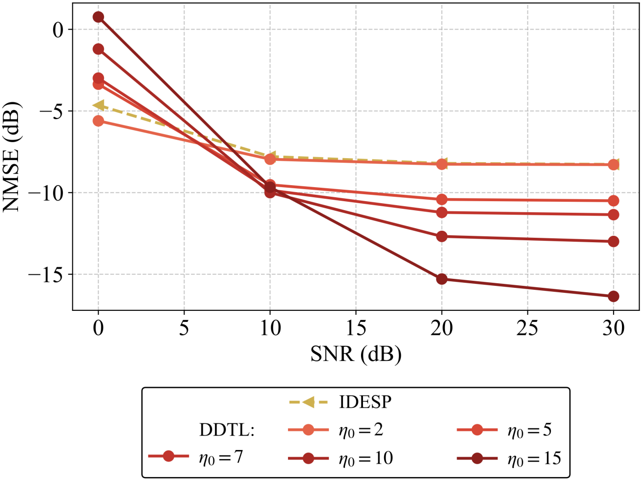

Figure: NMSE vs. SNR, for different algorithms.

Conclusion

- In this work, we studied the **localization properties** of Dirac and Laplacian operators under a **unifying framework**;
- We derived a notion of Dirac Driven Transform building on top of **sparse superpositions of differently graded Dirac equation of networks**;
- We posed the Dirac Driven Transform Learning problem as a **data-driven optimization problem** with an alternated solution scheme;
- Interesting directions include extension to **higher-order networks** and full learning pipelines.

Thank you for your attention!
Questions are welcome :)