Stationarity and Spectral Characterization of Random Signals on Simplicial Complexes

Andrei Buciulea

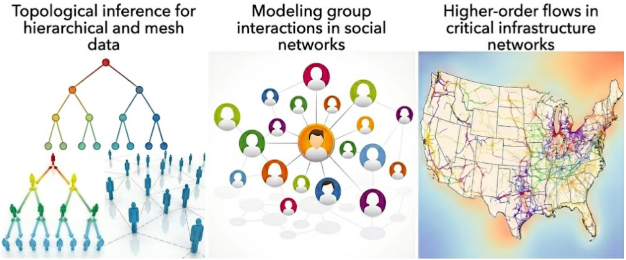
King Juan Carlos University - Madrid, Spain

Joint work with
M. Navarro, S. Segarra, and A. G. Marques

GSPW 2026, Madrid, Spain

Why Go Beyond Graphs?

- We are surrounded by networks generating data at multiple scales
- These data encode rich structure about the underlying system
- Graphs are the standard tool to model these systems



- Graphs capture **pairwise interactions only**
- Systems exhibit higher-order interactions (groups, multi-agent relations,..)
- **Simplicial complexes (SCs)** tractable tool to encode **multi-way relations**
- Higher-order data are also random: noisy, incomplete, and sample-limited

Our Contribution: Topological Signal Processing

- Most existing works rely on **deterministic models** for topological signals
- **Challenge:** Real data involves noise, missing observations, and uncertainty

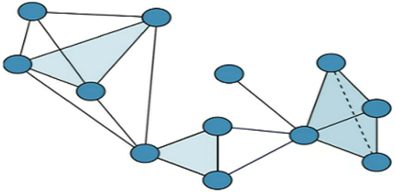
- **Key Idea**
  - $\Rightarrow$ Probabilistic framework for signals on simplicial complexes
  - $\Rightarrow$ Define stationary topological processes as outputs of **topological filters driven by white noise**

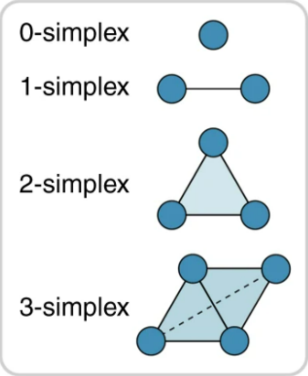
- **Key contributions:**
  - $\Rightarrow$ Extension of stationarity to SC via Hodge/Dirac spectra
  - $\Rightarrow$ Spectral characterization of second-order statistics
  - $\Rightarrow$ Generalization of classical SP tools to topological domain

What is a Simplicial Complex (SC)?

- SC: Hypergraph with a strong structure (simplicial closure constraint)
- Let $\mathcal{P}^K$ be a SC of order K

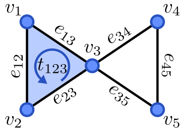
  $\Rightarrow$ Vertices, edges, triangles, tetrahedra, etc
- Graphs as (K=1)-SCs
- If a simplex belongs to $\mathcal{P}^K$, all its faces also belong to $\mathcal{P}^K$ $\Rightarrow$ Well-defined Laplacians
- E.g.: Social structure





Operators: Incidence and Hodge Laplacians

- **Incidence matrix:** $\mathbf{B}_k \in \{-1, 0, 1\}^{N_{k-1} \times N_k}$
  - $\Rightarrow$ links $(k-1)$- and k-simplices
  - $\Rightarrow$ canonical ordering of simplex vertices
  - $\Rightarrow$ signs encode orientation agreement

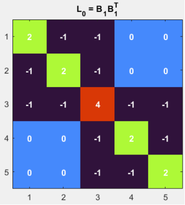


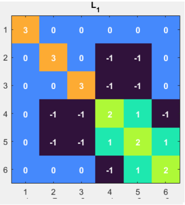
- **Hodge Laplacian**: topological connectivity and signal operator across simplicial orders

$$\mathbf{L}_0 = \mathbf{B}_1 \mathbf{B}_1^\top, \qquad \mathbf{L}_K = \mathbf{B}_K^\top \mathbf{B}_K$$

$$\mathbf{L}_k = \mathbf{L}_{k,\ell} + \mathbf{L}_{k,u}, \quad \mathbf{L}_{k,\ell} = \mathbf{B}_k^\top \mathbf{B}_k, \quad \mathbf{L}_{k,u} = \mathbf{B}_{k+1} \mathbf{B}_{k+1}^\top$$

  - $\Rightarrow \mathbf{L}_{k,\ell}$: connects k-simplices through $(k-1)$-faces
  - $\Rightarrow \mathbf{L}_{k,u}$: connects k-simplices through $(k+1)$-cofaces





- **A k-simplicial signal** is

$$\mathbf{s}^k = [s_1^k, \ldots, s_{N_k}^k]^\top \in \mathbb{R}^{N_k},$$

assigning a value (and orientation) to each k-simplex

Multiorder Topological Signals and Dirac Operator

- A **multiorder simplicial signal** stacks all orders in a single vector

$$\mathbf{s} = \left[(\mathbf{s}^0)^\top, \dots, (\mathbf{s}^K)^\top\right]^\top \in \mathbb{R}^N, \quad N = \textstyle\sum_{k=0}^{K} N_k.$$

- A **multiorder operator** accounting for all orders of the SC

Dirac operator

$$\mathbf{D} = \begin{bmatrix} \mathbf{0} & \mathbf{B}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{B}_1^\top & \mathbf{0} & \mathbf{B}_2 & \ddots & \vdots \\ \mathbf{0} & \mathbf{B}_2^\top & \mathbf{0} & \ddots & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \mathbf{B}_K \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{B}_K^\top & \mathbf{0} \end{bmatrix}$$

Squared Dirac

$$\mathbf{D}^2 = \begin{bmatrix} \mathbf{L}_0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_1 & \mathbf{0} & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{L}_2 & \ddots & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{L}_K \end{bmatrix}$$

- The Dirac operator couples simplices of adjacent orders
  - $\Rightarrow$ $\mathbf{D}$: transfers information between orders $k-1 \leftrightarrow k \leftrightarrow k+1$
- The squared Dirac operator recovers the Hodge Laplacians
  - $\Rightarrow$ $\mathbf{D}^2$: decouples into independent Hodge Laplacians per order

Weak Topological Stationarity

Definition of weak topological stationarity

A zero-mean topological signal $\mathbf{s}$ is weakly stationary if $\mathbf{s} = \mathbf{Hw}$, where $\mathbf{w}$ is white simplicial noise with $\mathbb{E}[\mathbf{w}] = \mathbf{0}$, $\mathbb{E}[\mathbf{w}\mathbf{w}^\top] = \mathbf{I}$, and $\mathbf{H}$ is a **topological filter** that is a polynomial in $\mathbf{T} \in \{\mathbf{L}_k, \mathbf{D}\}$

Topological filters: $\mathbf{H}$ aggregate signals over faces/cofaces via $\mathbf{T}$

$$\mathbf{H} = \sum_{r=0}^{R-1} h_r \mathbf{T}^r, \qquad \mathbf{h} = [h_0, \ldots, h_{R-1}]^\top$$

- **Spectral representation:** If $\mathbf{T} = \mathbf{U}\mathrm{diag}(\boldsymbol{\lambda})\mathbf{U}^\top$, then

  $\Rightarrow \mathbf{H} = \mathbf{U}\,\mathrm{diag}\big(h(\boldsymbol{\lambda})\big)\,\mathbf{U}^\top, \qquad h(\lambda_i) = \sum_{r=0}^{R-1} h_r \lambda_i^r$

- **Covariance structure:** $\mathbf{C} = \mathbb{E}[\mathbf{s}\mathbf{s}^\top] = \mathbf{H}\mathbf{H}^\top = \mathbf{H}^2 = \mathbf{U}\,\mathrm{diag}\big(h(\boldsymbol{\lambda})^2\big)\,\mathbf{U}^\top$

  $\Rightarrow$ Covariance fully determined by $(\mathbf{T}, \mathbf{h})$

  $\Rightarrow$ Hodge case: local (single order) filtering

  $\Rightarrow$ Dirac case: multiorder joint filtering

Spectral Stationarity and Topological PSD

Definition of spectral stationarity

A zero-mean random topological signal $\mathbf{s}$ *is weakly stationary with respect to* $\mathbf{T} \in \{\mathbf{L}_k, \mathbf{D}\}$ *if* $\mathbf{C}$ *and* $\mathbf{T}$ *are simultaneously diagonalizable* $\mathbf{C} = \mathbf{U}\mathrm{diag}(\mathbf{p})\mathbf{U}^\top$ *where* $\mathbf{U}$ *denotes the shared eigenvectors of* $\mathbf{C}$ *and* $\mathbf{T}$

Definition of power spectral density

For a random topological signal $\mathbf{s}$ *that is stationary with respect to* $\mathbf{T} = \mathbf{U}\mathrm{diag}(\boldsymbol{\lambda})\mathbf{U}^\top \in \{\mathbf{L}_k, \mathbf{D}\}$*, its power spectral density (PSD) is defined as*

$$\mathbf{p} := \mathrm{diag}(\mathbf{U}^\top \mathbf{C} \mathbf{U})$$

Spectral domain interpretation. Let $\tilde{\mathbf{s}} = \mathbf{U}^\top \mathbf{s}$. Then $\mathbb{E}[\tilde{\mathbf{s}}\,\tilde{\mathbf{s}}^\top] = \mathrm{diag}(\mathbf{p})$

- $\Rightarrow$ Spectral coefficients are **uncorrelated**
- $\Rightarrow$ $\mathbf{p}$ distributes signal energy over topological frequencies

Filter interpretation. If $\mathbf{s} = \mathbf{H}\mathbf{w}$ with $\mathbf{H} = h(\mathbf{T})$, then $p_i = h(\lambda_i)^2$

Covariance estimation I

Nonparametric covariance and PSD estimation

Given M realizations $\mathbf{S} = [\mathbf{s}_1, \dots, \mathbf{s}_M]$ of a stationary topological signal, estimate the covariance matrix $\mathbf{C} = \mathbb{E}[\mathbf{s}\mathbf{s}^\top]$ and its PSD exploiting the spectral structure induced by $\mathbf{T} \in \{\mathbf{L}_k, \mathbf{D}\}$

- **Sample covariance.** $\hat{\mathbf{C}} = \frac{1}{M}\mathbf{S}\mathbf{S}^\top$. Complexity: $\mathcal{O}(MN^2)$
  - $\Rightarrow$ Does not exploit stationarity
- **Correlogram-based estimator.** Complexity: $\mathcal{O}(MN^2 + N^3)$

$$\hat{\mathbf{p}}_{cg} = \mathrm{diag}\left(\mathbf{U}^\top \hat{\mathbf{C}} \mathbf{U}\right), \qquad \hat{\mathbf{C}}_{cg} = \mathbf{U}\mathrm{diag}(\hat{\mathbf{p}}_{cg})\mathbf{U}^\top$$

- **Periodogram-based estimator.** Complexity: $\mathcal{O}(MN^2)$

$$\tilde{\mathbf{S}} = \mathbf{U}^\top \mathbf{S}, \qquad [\hat{\mathbf{p}}_{pg}]_i = \frac{1}{M}\sum_{m=1}^{M} [\tilde{\mathbf{s}}_m]_i^2, \qquad \hat{\mathbf{C}}_{pg} = \mathbf{U}\mathrm{diag}(\hat{\mathbf{p}}_{pg})\mathbf{U}^\top$$

- Correlogram and periodogram yield the same PSD estimate
  - $\Rightarrow$ Periodogram avoids explicit cov. estimation (preferable for $M < N$)

Covariance estimation II

Parametric covariance and PSD estimation

Given M realizations $\mathbf{S} = [\mathbf{s}_1, \ldots, \mathbf{s}_M]$ of a stationary topological signal, estimate the covariance matrix $\mathbf{C} = \mathbb{E}[\mathbf{s}\mathbf{s}^\top]$ and its PSD assuming the stationary topological signal follows a low-order filter model

- **MA topological process**

  $\Rightarrow$ $\mathbf{s} = \mathbf{H}(\boldsymbol{\beta})\mathbf{w}, \qquad \mathbf{H}(\boldsymbol{\beta}) = \sum_{r=0}^{R-1} \beta_r \mathbf{T}^r$

Spatial estimation

$$\hat{\boldsymbol{\beta}}_{\mathbf{C}} = \operatorname{argmin}_{\boldsymbol{\beta}} \mathcal{D}_{\mathbf{C}}\left(\hat{\mathbf{C}}, \mathbf{C}^{\beta}(\boldsymbol{\beta})\right)$$

$\Rightarrow$ with $\mathbf{C}^{\beta}(\boldsymbol{\beta}) = \mathbf{H}(\boldsymbol{\beta})\mathbf{H}(\boldsymbol{\beta})^\top = \sum_{r=0}^{R-1} \sum_{r'=0}^{R-1} \beta_r \beta_{r'} \mathbf{T}^{r+r'}$

Spectral estimation

$$\hat{\gamma}_{\mathbf{p}} = \operatorname{argmin}_{\gamma} \mathcal{D}_{\mathbf{p}}\left(\hat{\mathbf{p}}_{\mathrm{pg}}, \boldsymbol{\Psi}^{\gamma}\gamma\right),$$

$\Rightarrow$ with $\gamma_r := \sum_{r'+r''=r} \beta_{r'} \beta_{r''}$,

$\Rightarrow$ $\boldsymbol{\Psi}^{\gamma}$: Vandermonde matrix associated with the spectrum of $\mathbf{T}$

Covariance estimation III

Parametric covariance and PSD estimation

Given M realizations $\mathbf{S} = [\mathbf{s}_1, \ldots, \mathbf{s}_M]$ of a stationary topological signal, estimate the covariance matrix $\mathbf{C} = \mathbb{E}[\mathbf{s}\mathbf{s}^\top]$ and its PSD assuming the stationary topological signal follows a low-order filter model

- **AR topological process:** $\mathbf{s} = \sum_{r=1}^{R} \alpha_r \mathbf{T}^r \mathbf{s} + \mathbf{w}, \qquad \mathbf{s} = \mathbf{H}^{-1}(\boldsymbol{\alpha})\mathbf{w}$

  Spatial estimation $\hat{\boldsymbol{\alpha}}_{\mathbf{C}} = \operatorname{argmin}_{\boldsymbol{\alpha}} \mathcal{D}_{\mathbf{C}}\Big(\hat{\mathbf{C}}\big(\mathbf{C}^{\alpha}(\boldsymbol{\alpha})\big)^{-1}, \mathbf{I}\Big)$

  $\Rightarrow$ With $\mathbf{C}^{\alpha}(\boldsymbol{\alpha}) = \mathbf{H}^{-2}(\boldsymbol{\alpha})$

  Spectral estimation $\hat{\boldsymbol{\alpha}}_{\mathbf{p}} = \operatorname{argmin}_{\boldsymbol{\alpha}} \mathcal{D}_{\mathbf{p}}\Big(\operatorname{diag}^{-1}\big(\mathbf{p}(\boldsymbol{\alpha})\big)\hat{\mathbf{p}}_{\mathrm{pg}}, \mathbf{1}\Big)$, with

  $\Rightarrow$ $[\mathbf{p}(\boldsymbol{\alpha})]_i := \frac{1}{\left(h^{\alpha}(\lambda_i)\right)^2} = \frac{1}{[\mathbf{1} - \boldsymbol{\Psi}^{\alpha}\boldsymbol{\alpha}]_i^2}$ for all $i \in [N]$,

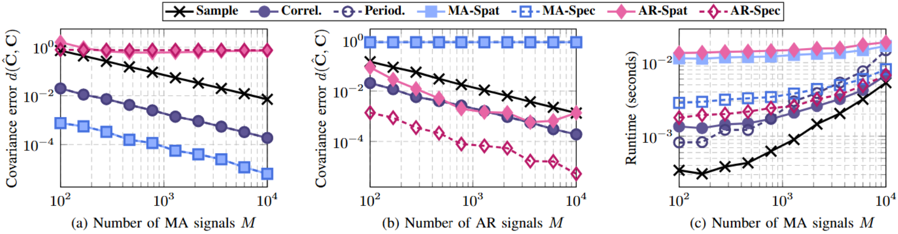
- **Other parametric estimators**

  $\Rightarrow$ Non-polynomial filtering (exponential or rational responses)

  $\Rightarrow$ Kernel parametrizations (Gaussian or Laplacian kernels)

Covariance estimation: MA vs AR models

- Covariance estimation error for MA and AR signal models (a,b)
- Runtime comparison as number of samples M increases (c)



- Model-matched estimators perform best; mismatch increases error
- Correl $\equiv$ Period and MA-Spat $\equiv$ MA-Spec (equivalent formulations)
- AR-Spec outperforms AR-Spat in sample efficiency
- Runtime:
  - $\Rightarrow$ Periodogram faster for small M, correlogram better for large M
  - $\Rightarrow$ Spectral parametric methods are most efficient overall

Topological Denoising via Wiener Filtering

Topological Denoising via Wiener Filtering

Given a noisy topological signal $\mathbf{y} = \mathbf{s} + \mathbf{v}$*, the goal is to recover the unknown clean signal* $\mathbf{s}$ *corrupted by additive noise* $\mathbf{v}$*, with* $\mathbf{C}_v = \sigma_v^2 \mathbf{I}$ *and* $\sigma_v^2 > 0$*, assuming* $\mathbf{s}$ *stationary in* $\mathbf{T} \in \{\mathbf{L}_k, \mathbf{D}\}$ *and* $\mathbf{C} = \mathbb{E}[\mathbf{s}\mathbf{s}^\top]$

- We seek the **Minimum MSE (MMSE)** estimate: of $\mathbf{s}$ from $\mathbf{y}$

$$\hat{\mathbf{s}} = \arg\min_{\mathbf{z}} \mathbb{E}[\|\mathbf{s} - \mathbf{z}\|_2^2 \mid \mathbf{y}] = \mathbf{C}(\mathbf{C} + \sigma_v^2 \mathbf{I})^{-1}\mathbf{y} = \mathbf{G}\mathbf{y}$$

  ⇒ with $\mathbf{G}$ the Wiener filter matrix

  ⇒ Infeasible to compute $\mathbf{G}$ for large SC

- Denoising in the spectral domain (considering stationarity)

$$\tilde{\hat{\mathbf{s}}} = \mathbf{U}^\top \mathbf{G} \mathbf{U} \tilde{\mathbf{y}} = \mathrm{diag}(\mathbf{p})\mathrm{diag}^{-1}\left(\mathbf{p} + \sigma_v^2 \mathbf{1}\right)\tilde{\mathbf{y}} = \tilde{\mathbf{g}} \circ \tilde{\mathbf{y}}$$

  ⇒ with $\tilde{\mathbf{y}} = \mathbf{U}^\top \mathbf{y}$ and $\tilde{g}_i = \frac{p_i}{p_i + \sigma_v^2}$ for all $i \in [N]$

- **Scalable:** Reduces complexity from $\mathcal{O}(N^3)$ to $\mathcal{O}(N^2)$ (TFT)

Topological Signal Reconstruction & Interpolation

Topological Signal Reconstruction & Interpolation

Given partially observed topological signal $\bar{\mathbf{s}} = \mathbf{\Theta}\mathbf{s} + \mathbf{v}$*, where* $\mathbf{\Theta} \in \{0,1\}^{P\times N}$ *selects the observed entries, and* $\mathbf{v}$ *is additive white noise with* $\mathbf{C}_v = \sigma_v^2\mathbf{I}$*. The goal is to recover the full unknown topological signal* $\mathbf{s}$*, that is stationary with respect to* $\mathbf{T} \in \{\mathbf{L}_k, \mathbf{D}\}$ *and with* $\mathbf{C} = \mathbb{E}[\mathbf{s}\mathbf{s}^\top]$

- We seek the **Minimum MSE (MMSE)** estimate of $\mathbf{s}$ from $\bar{\mathbf{s}}$

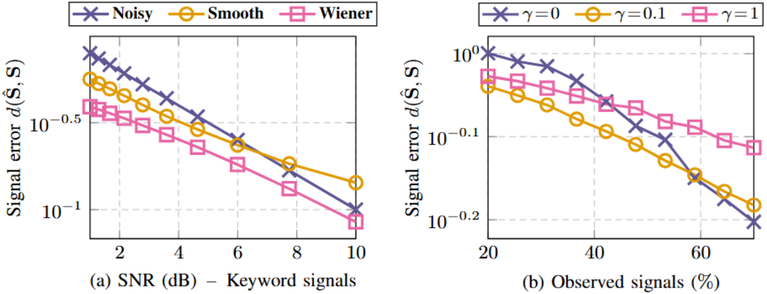
$$\hat{\mathbf{s}} = \arg\min_{\mathbf{z}} \frac{1}{\sigma_v^2}\|\bar{\mathbf{s}} - \mathbf{\Theta}\mathbf{z}\|_2^2 + \mathbf{z}^\top\mathbf{C}^{-1}\mathbf{z}$$

  $\Rightarrow$ $\mathbf{z}^\top\mathbf{C}^{-1}\mathbf{z}$ promotes stationarity behaviour for $\hat{\mathbf{s}}$

- **Key effect:** balances data fidelity ($\mathbf{\Theta}$) with topological prior structure
- **Interpretation:** $\mathbf{C}^{-1}$ encodes prior structure
  - $\Rightarrow$ **Smoothness** $\Rightarrow \hat{\mathbf{s}} = \text{argmin}_{\mathbf{z}} \frac{1}{\sigma_v^2}\|\bar{\mathbf{s}} - \mathbf{\Theta}\mathbf{z}\|_2^2 + \mathbf{z}^\top\mathbf{D}^2\mathbf{z}$
  - $\Rightarrow$ **AR-1** $\Rightarrow \hat{\mathbf{s}} = \text{argmin}_{\mathbf{z}} \frac{1}{\sigma_v^2}\|\bar{\mathbf{s}} - \mathbf{\Theta}\mathbf{z}\|_2^2 + \mathbf{z}^\top\big(\mathbf{I} + \alpha^2\mathbf{T}^2 - 2\alpha\mathbf{T}\big)\mathbf{z}$

Denoising and reconstruction (Coauthorship dataset)

- Real-world simplicial complex of academic coauthorships
- Node, edge, triangle signals: $M = 1903$ keyword frequencies per author
  - $\Rightarrow$ (a) Denoising under additive Gaussian noise $\mathbf{Y} = \mathbf{S} + \mathbf{V}$
  - $\Rightarrow$ (b) Interpolation from partial observations of simplices



- Wiener filtering (stationary model) achieves lowest error
- Smoothness-based method helps only in high-noise / low-data regimes
- A trade-off between promoting smoothness and stationarity is preferable

Additional Directions

- **Time-varying topological signals**
  - ⇒ Extension of stationarity to dynamic settings
  - ⇒ State-space models and temporal filtering on simplicial complexes
- **Robust covariance estimation**
  - ⇒ Joint covariance estimation from incomplete or noisy observations
  - ⇒ Unified frameworks combining denoising + covariance learning
- **Beyond simplicial complexes**
  - ⇒ Extension to hypergraphs and cell complexes
  - ⇒ Generalization of filtering and stationarity notions
- **Simplicial complex estimation**
  - ⇒ Recover structure $(\{\mathbf{B}_k\}, \mathbf{D}, \{\mathbf{L}_k\})$ from data
  - ⇒ Stationarity-driven learning of higher-order topology
- **Random simplicial complexes**
  - ⇒ Joint modeling of uncertainty in signals and topology
  - ⇒ Probabilistic frameworks for higher-order connectivity

Conclusions

- **Characterization of stationarity for random topological signals**
  1. We define stationarity for random topological signals via Dirac filters
  2. Stationarity yields a PSD that characterizes second-order statistics
  3. Enables spectral-domain formulations with reduced complexity
- Addressing classical SP tasks using stationary topological signals
  - $\Rightarrow$ Covariance estimation
  - $\Rightarrow$ Signal denoising
  - $\Rightarrow$ Signal reconstruction and interpolation
- Additional directions for topological stationary signals

- **THANKS!**
  - $\Rightarrow$ Additional details can be found in the paper
  - $\Rightarrow$ Feel free to contact me for questions and code: **andrei.buciulea@urjc.es**