

Hodge-Compositional Edge Gaussian Processes

Edge flow; difference vs. node data; graph vs. simplicial complex
Smoothness of edge flows: div and curl; Hodge decomposition
GP modeling of edge functions: divfree, curl-free kernels ...



AISTATS 2024

Maosheng Yang (TU Delft),

Viacheslav Borovitskiy (ETH Zürich) Elvin Isufi (TU Delft)

Graphs vs Simplicial 2-Complexes 0-, 1-, 2-, 3-simplices e_{10} *e*₃ e_3 *e*₉ *t*₃(e_8 6

Graph Simplicial 1-complex G = (V, E)

Simplicial 2-complex $SC_2 = (V, E, T)$

- Oriented simplices (equivalence class of permutations)



Where are SCs used?

- Network analysis
- Topological data analysis
- Topological signal processing
- Topological deep learning
- Numerical methods
- Computer graphics

- . . .

Functions on simplices Signals on nodes, edges, triangles, ...

0.5 e_3



Node function $f_0: V \to \mathbb{R}$ $\mathbf{f}_0 = (f_0(1), \dots, f_0(N_0))^\top$ $\mathbf{f}_1 = (f_1(e_1), \dots, f_1(e_{N_1}))^\top$

Edge function $f_1: E \to \mathbb{R}$

- Alternating property - Magnitude and sign

- Flow-type data (natural)

- . . .

- Physical world: traffic flow, water flow, information flow...
- Forex: exchange rates
- Game theory (Candogan et al. 2011)
- Ranking data (Jiang et al. 2011)
- Edge-based vector field discretisation (computer graphics)

Triangle function $f_2: T \to \mathbb{R}$

0-, 1-, 2-cochains in topology



Algebraic reps. of simplicial 2-complex Incidences & Laplacians



Graph Laplacian: $\mathbf{L}_0 = \mathbf{B}_1 \mathbf{B}_1^{\mathsf{T}}$ 1-Hodge Laplacian: $\mathbf{L}_1 = \mathbf{B}_1^{\mathsf{T}} \mathbf{B}_1 + \mathbf{B}_2 \mathbf{B}_2^{\mathsf{T}} := \mathbf{I}$

$$= \mathbf{L}_{1,d} + \mathbf{L}_{1,u}$$

Down

GPs on graphs Modeling node functions

- $\mathbf{f}_0 \sim GP(\mathbf{0}, \mathbf{K}_0)$ (Borovitskiy et al. 2021)
- Matérn graph kernel

$$\Phi(\mathbf{L}_0)\mathbf{f}_0 = \mathbf{W}_0$$
, with

$$\Phi(\mathbf{L}_0) = \left(\frac{2\nu}{\kappa^2}\mathbf{I} + \mathbf{L}_0\right)^{\frac{\nu}{2}} \text{ and } \mathbf{w}_0 \sim N(\mathbf{0}, \sigma^2)$$

• The solution has kernel

$$\mathbf{K}_{0} = \sigma^{2} \sum_{n=0}^{N_{0}-1} \psi(\lambda_{n}) \mathbf{u}_{n} \mathbf{u}_{n}^{\mathsf{T}} = \sigma^{2} \left(\frac{2\nu}{\kappa^{2}} \mathbf{I} + \mathbf{L}_{0} \right)$$
$$\psi(\lambda) = \begin{cases} \left(\frac{2\nu}{\kappa^{2}} + \lambda \right)^{-\nu} & \nu < \infty, \text{Matern} \\ e^{-\frac{\kappa^{2}}{2}\lambda} & \nu = \infty, \text{Diffusion} \end{cases}$$

GPs from Euclidean to non-Euclidean

GP in Euclidean settings Function on a set $f: X \to \mathbb{R}$ $f \sim GP(\mu, k)$ -Predictive distribution $f_{|y|}$ -Matérn GP family, e.g., diffusion $k(x, x') = \sigma^2 \exp\left(-\frac{d(x, x')^2}{2\kappa^2}\right)$

- Distance-based: geometry-aware, but not welldefined for manifolds, graphs ...
- Instead, as solutions of SDEs (Whittle (1963); Lindgren et al. (2011))

$$\left(\frac{2\nu}{\kappa^2} - \Delta\right)^{\frac{\nu}{2} + \frac{d}{4}} f = w$$

- Δ : Laplacian, w: white noise
- implicit, generalizable, domain-aware
- explicit for some domains



Matérn Edge GPs Derived from SDEs on the edge set

- $\mathbf{f}_1 \sim \operatorname{GP}(\mathbf{0}, \mathbf{K}_1)$
- Matérn graph kernel

$$\Phi(\mathbf{L}_1)\mathbf{f}_1 = \mathbf{w}_1$$
, with

EVD: $\mathbf{L}_1 = \mathbf{U}_1 \mathbf{\Lambda}_1$

$$\Phi(\mathbf{L}_1) = \left(\frac{2\nu}{\kappa^2}\mathbf{I} + \mathbf{L}_1\right)^{\frac{\nu}{2}} \text{ and } \mathbf{w}_1 \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$$

• The solution gives edge GPs

Matérn:
$$\mathbf{f}_1 \sim \mathrm{GP}\left(0, \left(\frac{2\nu}{\kappa^2}\mathbf{I} + \mathbf{L}_1\right)^{-\nu}\right)$$

Diffusion: $\mathbf{f}_1 \sim \mathrm{GP}\left(0, e^{-\frac{\kappa^2}{2}\mathbf{L}_1}\right)$



Smoothness Node function — 0-form (scalar field) Edge function — 1-form (vector field)

> Divergence Curl



 $[\mathbf{B}_{1}^{\mathsf{T}}\mathbf{v}]_{[1,2]} = -1.34 - 0.96 = -2.30$

-Node signal **v** -Edge flows **f**

Gradient of node signal: $[\mathbf{f}_G]_{[i,j]} = [\mathbf{B}_1^{\top}\mathbf{v}]_{[i,j]} = [\mathbf{v}_j]_j - [\mathbf{v}_j]_i$ Divergence of edge flows: $[\mathbf{B}_1 \mathbf{f}]_{[i]} = \sum \mathbf{f}_{[i,i]} - \sum \mathbf{f}_{[i,k]}$ Curl of edge flows: $[\mathbf{B}_{2}^{\mathsf{T}}\mathbf{f}]_{t} = \mathbf{f}_{[i,i]} + \mathbf{f}_{[i,k]} - \mathbf{f}_{[i,k]}$, for t = [i, j, k]





 $[\mathbf{B}_1\mathbf{f}]_5 = 0.5 + 2.6 - (0.9 + 2.6) = -0.4$ $[\mathbf{B}_{2}'\mathbf{f}]_{[1,2,3]} = -1.2 + 1.8 - (-2.9) = 3.5$

Gradient of node signal: $\mathbf{B}_{1}^{\mathsf{T}}\mathbf{v}$ $[\mathbf{f}_{G}]_{[i,j]} = [\mathbf{v}]_{j} - [\mathbf{v}]_{i}$ Divergence of edge flows: $[\mathbf{B}_1 \mathbf{f}]_{[i]} = \sum \mathbf{f}_{[j,i]} - \sum \mathbf{f}_{[i,k]}$ *i*<*k* Net-flow = in_flow - out_flow j < iCurl of edge flows: $[\mathbf{B}_2^{\mathsf{T}}\mathbf{f}]_t = \mathbf{f}_{[i,i]} + \mathbf{f}_{[i,k]} - \mathbf{f}_{[i,k]}$, for t = [i, j, k]







 $[\mathbf{B}_1\mathbf{f}]_5 = 0.5 + 2.6 - (0.9 + 2.6) = -0.4$ $[\mathbf{B}_{2}^{\mathsf{T}}\mathbf{f}]_{[1,2,3]} = -1.2 + 1.8 - (-2.9) = 3.5$

Gradient of node signal: $\mathbf{B}_{1}^{\mathsf{T}}\mathbf{v}$ $[\mathbf{f}_{G}]_{[i,j]} = [\mathbf{v}]_{j} - [\mathbf{v}]_{i}$ Divergence of edge flows: $[\mathbf{B}_1 \mathbf{f}]_{[i]} = \sum \mathbf{f}_{[j,i]} - \sum \mathbf{f}_{[i,k]}$ Curl of edge flows: $[\mathbf{B}_{2}^{T}\mathbf{f}]_{t} = \mathbf{f}_{[i,j]} + \mathbf{f}_{[j,k]} - \mathbf{f}_{[i,k]}$, for t = [i, j, k]**Net-circulation in triangles**







 $[\mathbf{B}_1\mathbf{f}]_5 = 0.5 + 2.6 - (0.9 + 2.6) = -0.4$ Hodge Laplacians = Grad Div + Curl* Curl $[\mathbf{B}_2^{\mathsf{T}}\mathbf{f}]_{[1,2,3]} = -1.2 + 1.8 - (-2.9) = 3.5$ Hodge Laplacian: $\mathbf{L}_1 = \mathbf{B}_1^{\mathsf{T}} \mathbf{B}_1 + \mathbf{B}_2 \mathbf{B}_2^{\mathsf{T}}$

Gradient of node signal: $\mathbf{B}_{1}^{\top}\mathbf{v}$ $[\mathbf{f}_{G}]_{[i,j]} = [\mathbf{v}]_{j} - [\mathbf{v}]_{i}$ Divergence of edge flows: $[\mathbf{B}_1 \mathbf{f}]_{[i]} = \sum \mathbf{f}_{[j,i]} - \sum \mathbf{f}_{[i,k]}$ Net-flow = in_flow - out_flow j < i*i*<*k* Curl of edge flows: $[\mathbf{B}_2^{\mathsf{T}}\mathbf{f}]_t = \mathbf{f}_{[i,j]} + \mathbf{f}_{[j,k]} - \mathbf{f}_{[i,k]}$, for t = [i,j,k]**Net-circulation in triangles**







Curl flow Div-free, solenoidal

Applications of Hodge decomposition



Ocean currents



 $f_{[a,b]} + f_{[b,c]} - f_{[a,c]} = 0$ Curl-free

Gradient flow Curl-free, irrotational

Curl flow Div-free, solenoidal

Chen, Yu-Chia et al. (2021) "Helmholtzian Eigenmap."

- Water flows (div-free) - Electrical currents (KCL), voltages (KVL)

- Brain networks (Anand et al. 2022)
- Game theory (Candogan et al. 2011)
- Ranking problems (Jiang et al. 2011)
- Random walks (Strang et al. 2020)

- . . .

Eigenspace of L_1 spans Hodge subspaces

- Nonzero Eigenspace of down Laplacian spans the gradient space
- Nonzero Eigenspace of up Laplacian spans the curl space
- Zero Eigenspace of Laplacian spans the harmonic space



Simplicial Fourier transform

Frequency – eigenvalues Fourier basis — eigenvectors

Curl eigenvector Fourier basis reflecting rotational properties

EVD: $\mathbf{L}_1 = \mathbf{U}_1 \mathbf{\Lambda}_1 \mathbf{U}_1^{\mathsf{T}}$

 $\mathbf{U}_1 = [\mathbf{U}_H \ \mathbf{U}_G \ \mathbf{U}_C]$ $\operatorname{span}(\mathbf{U}_H) = \operatorname{ker}(\mathbf{L}_1)$ $\operatorname{span}(\mathbf{U}_G) = \operatorname{im}(\mathbf{B}_1^{\top})$ $\operatorname{span}(\mathbf{U}_{C}) = \operatorname{im}(\mathbf{B})$

Yang et al. Simplicial Convolutional Filters, 2022







Eigenspace spans Hodge subspaces

- Down Laplacian, its nonzero eigenspace spans the gradient space







 λ_G , more divergent

- Up laplacian, its nonzero eigenspace spans the curl space













 λ_C , more rotational



Hodge-compositional Edge GPs **Curl-free, div-free GPs**

• Gradient kernel $\mathbf{K}_G = \mathbf{U}_G \Psi_G(\Lambda_G) \mathbf{U}_G^{\dagger}$; Curl kernel $\mathbf{K}_C = \mathbf{U}_C \Psi_C(\Lambda_C) \mathbf{U}_C^{\dagger}$

• Matérn family:
$$\Psi_{\Box}(\Lambda_{\Box}) = \sigma_{\Box}^2 \left(\frac{2\nu_{\Box}}{\kappa_{\Box}^2}\mathbf{I} + I\right)$$

Also as solutions of SDEs, e.g.,

 $\Phi_C(\mathbf{L}_{1,\mathbf{u}})$ $\mathbf{f}_1 = \mathbf{w}_C$, with curl noise $\mathbf{w}_C \sim N(0, \sigma_C^2 \mathbf{U}_C \mathbf{U}_C^{\mathsf{T}})$ and $\Phi(\mathbf{L}_{1,u}) = \left(\frac{2\nu_C}{\kappa_C^2}\mathbf{I} + \mathbf{L}_{1,u}\right)^{\frac{\nu_C}{2}} \text{ or } \Phi(\mathbf{L}_{1,u}) = e^{\frac{\kappa_C^2}{4}\mathbf{L}_{1,u}}$

 $\begin{aligned} \mathbf{f}_G \sim \mathbf{GP}(\mathbf{0}, \mathbf{K}_G) \\ \mathbf{f}_H \sim \mathbf{GP}(\mathbf{0}, \mathbf{K}_H) \\ \mathbf{f}_C \sim \mathbf{GP}(\mathbf{0}, \mathbf{K}_C) \end{aligned}$

 Λ_{\Box})^{$-\nu_{\Box}$}, $\Box = H, G, C$



Hodge-compositional Edge GPs

Composition of three GPs on the Hodge subspaces

- Kernel: $K_1 = K_G + K_H + K_C$
- Mutual independence hypothesis
- Separate learning of different components
- Automatic determination of Hodge components, instead of solving Hodge decomp.
- Edge Fourier Feature perspective

Alternative formulation

via node-edge-triangle interactions - Derivatives of GPs are also GPs - Induce edge GPs from node and triangle GPs $K_1 = K_H + B_1^\top K_0 B_1 + B_2 K_2 B_2^\top$ - Induce node GPs from edge GPs





GP based Forex prediction





True





non-Hodge





GP based Ocean current analysis

Original







Predictive mean







Pointwise variance





Learned kernel

State estimation in water supply networks **Based on the node-edge joint GPs**





Head Source 4 3 2 1 Flow -2.0 -1.5 -1.0 -0.5 0.00.5 1.0 Hodge edge GP





Conclusion

- How to generalize GPs to non-Euclidean domains? SDE framework
- How to measure edge functions? Div and curl, like VFs
- What is a good edge GP? Edge dependency + Hodge decomposition
- Node-edge-triangle joint GPs Alain et al. 2023
- Continuous version: Euclidean VF Berlinghieri et al. 2023; Manifold VF Robert-Nicoud et al. 2024

Thank you!



Paper Code



Appendix Diffusion kernels from diffusion processes



Diffusion on nodes



Diffusion on edges

Tabular results

Table 1: Forex rates inference results.

Method	RMSE		NLPD		
	Diffusion	Matérn	Diffusion	Matérn	
Euclidean Line-Graph Non-HC HC	$\begin{array}{c} 2.17 \pm 0.13 \\ 2.43 \pm 0.07 \\ 2.48 \pm 0.07 \\ 0.08 \pm 0.12 \end{array}$	$\begin{array}{c} 2.19 \pm 0.12 \\ 2.46 \pm 0.07 \\ 2.47 \pm 0.08 \\ 0.06 \pm 0.12 \end{array}$	$\begin{array}{c} 2.12 \pm 0.07 \\ 2.28 \pm 0.04 \\ 2.36 \pm 0.07 \\ -3.52 \pm 0.02 \end{array}$	$\begin{array}{c} 2.20 \pm 0.18 \\ 2.32 \pm 0.03 \\ 2.34 \pm 0.04 \\ -3.52 \pm 0.02 \end{array}$	

Table C.1: Ocean current inference results.

Method	RMSE			NLPD		
	Diffusion	Matérn	Hodge Laplacian	Diffusion	Matérn	Hodge Laplacian
Euclidean	1.00 ± 0.01	1.00 ± 0.00		1.42 ± 0.01	1.42 ± 0.10	
Line-Graph	0.99 ± 0.00	0.99 ± 0.00		1.41 ± 0.00	1.41 ± 0.00	
Non-HC	0.35 ± 0.00	0.35 ± 0.00	0.35 ± 0.00	0.33 ± 0.00	0.36 ± 0.03	0.33 ± 0.01
HC	0.34 ± 0.00	0.35 ± 0.00	0.35 ± 0.00	0.33 ± 0.01	0.37 ± 0.04	0.33 ± 0.01

Method	Node	Heads	Edge Flowrates		
memori	RMSE	NLPD	RMSE	NLPD	
Diffusion, non-HC Matérn, non-HC	$\begin{array}{c} 0.16 \pm 0.05 \\ 0.16 \pm 0.04 \end{array}$	$\begin{array}{c} 0.72 \pm 2.06 \\ 0.71 \pm 2.39 \end{array}$	$\begin{array}{c} 0.32 \pm 0.05 \\ 0.26 \pm 0.05 \end{array}$	$\begin{array}{c} 0.97 \pm 1.80 \\ 0.10 \pm 0.13 \end{array}$	
Diffusion, HC Matérn, HC	$\begin{array}{c} 0.15 \pm 0.04 \\ 0.15 \pm 0.04 \end{array}$	$\begin{array}{c} -0.47 \pm 0.14 \\ -0.25 \pm 0.48 \end{array}$	$\begin{array}{c} 0.22 \pm 0.03 \\ 0.23 \pm 0.03 \end{array}$	$\begin{array}{c} -0.20 \pm 0.13 \\ -0.45 \pm 0.49 \end{array}$	

Table 3: WSN inference results.

Sampling gradient and curl edge GPs

Proof. We focus on the case of gradient GPs. First, we can decompose the gradient kernel in terms of $U_1 =$ $[U_H \ U_G \ U_C]$ as

$$\boldsymbol{K}_{G} = \boldsymbol{U}_{1} \begin{pmatrix} \boldsymbol{0} & \\ \boldsymbol{\Psi}_{G}(\boldsymbol{\Lambda}_{G}) & \\ & \boldsymbol{0} \end{pmatrix} \boldsymbol{U}_{1}^{\top}. \tag{B.9}$$

From a vector $\boldsymbol{v} = (v_1, \dots, v_{N_1})^\top$ of variables following independent normal distribution, we can draw a random sample of gradient function as

$$\boldsymbol{f}_{G} = \boldsymbol{U}_{1} \operatorname{diag}([\boldsymbol{0}, \Psi_{G}^{\frac{1}{2}}(\boldsymbol{\Lambda}_{G}), \boldsymbol{0}])\boldsymbol{v} \tag{B.10}$$

where diag([a, b, c]) is the diagonal matrix with $(a, b, c)^{\top}$ on its diagonal. Therefore, their curls are

$$\operatorname{curl} \boldsymbol{f}_G = \boldsymbol{B}_2^{\top} \boldsymbol{U}_1 \operatorname{diag}([\boldsymbol{0}, \Psi_G^{\frac{1}{2}}(\boldsymbol{\Lambda}_G), \boldsymbol{0}]) = \boldsymbol{B}_2^{\top} \boldsymbol{U}_G \Psi_G^{\frac{1}{2}}(\boldsymbol{\Lambda}_G) = \boldsymbol{0}.$$
(B.11)

Likewise, we can show the samples of a curl GP are div-free.

Posterior distribution of Hodge components

$$\begin{bmatrix} f_{H}(\boldsymbol{x}) \\ f_{H}(\boldsymbol{x}^{*}) \\ f_{G}(\boldsymbol{x}) \\ f_{G}(\boldsymbol{x}) \\ f_{C}(\boldsymbol{x}) \\ f_{C}(\boldsymbol{x}^{*}) \\ f_{L}(\boldsymbol{x}^{*}) \\ f_{1}(\boldsymbol{x}^{*}) \end{bmatrix} \sim \mathcal{N} \begin{pmatrix} \mathbf{K}_{H} & \mathbf{K}_{H}^{*} & \mathbf{K}_{H}^{*} & \mathbf{K}_{H}^{*} \\ \mathbf{K}_{H}^{*\top} & \mathbf{K}_{H}^{**} & \mathbf{K}_{G}^{*} & \mathbf{K}_{G}^{*} & \mathbf{K}_{H}^{**} \\ & \mathbf{K}_{G} & \mathbf{K}_{G}^{*} & \mathbf{K}_{G}^{*} & \mathbf{K}_{G}^{*} \\ & \mathbf{K}_{G}^{*\top} & \mathbf{K}_{G}^{**} & \mathbf{K}_{G}^{*} & \mathbf{K}_{G}^{*} \\ & & \mathbf{K}_{G}^{*\top} & \mathbf{K}_{G}^{**} & \mathbf{K}_{G}^{*} & \mathbf{K}_{G}^{*} \\ & & \mathbf{K}_{G}^{*\top} & \mathbf{K}_{G}^{**} & \mathbf{K}_{G}^{*} & \mathbf{K}_{G}^{*} \\ & & \mathbf{K}_{G}^{*\top} & \mathbf{K}_{G}^{**} & \mathbf{K}_{G}^{*} & \mathbf{K}_{G}^{*} & \mathbf{K}_{G}^{*} \\ & & \mathbf{K}_{H}^{*\top} & \mathbf{K}_{H}^{**} & \mathbf{K}_{G}^{*\top} & \mathbf{K}_{G}^{*} & \mathbf{K}_{G}^{*\top} & \mathbf{K}_{G}^{*} & \mathbf{K}_{G}^{*\top} & \mathbf{K}_{G}^{**} \\ \end{pmatrix} \end{pmatrix}$$
(B.26)

where we represent the kernel matrices by $K_1 = k_1(x, x), K_1^* = k_1(x, x^*)$ and $K_1^{**} = k_1(x^*, x^*)$, and likewise for the other kernel matrices. Given this joint distribution, we can obtain the posterior distributions of the three Hodge components as follows

$$f_{H}(\boldsymbol{x}^{*})|f_{1}(\boldsymbol{x}) \sim \mathcal{N}\left(\boldsymbol{K}_{H}^{*\top}\boldsymbol{K}_{1}^{-1}f_{1}(\boldsymbol{x}), \boldsymbol{K}_{H}^{**} - \boldsymbol{K}_{H}^{*\top}\boldsymbol{K}_{1}^{-1}\boldsymbol{K}_{H}^{*}\right)$$
(B.27a)
$$f_{G}(\boldsymbol{x}^{*})|f_{1}(\boldsymbol{x}) \sim \mathcal{N}\left(\boldsymbol{K}_{G}^{*\top}\boldsymbol{K}_{1}^{-1}f_{1}(\boldsymbol{x}), \boldsymbol{K}_{G}^{**} - \boldsymbol{K}_{G}^{*\top}\boldsymbol{K}_{1}^{-1}\boldsymbol{K}_{G}^{*}\right)$$
(B.27b)
$$f_{C}(\boldsymbol{x}^{*})|f_{1}(\boldsymbol{x}) \sim \mathcal{N}\left(\boldsymbol{K}_{C}^{*\top}\boldsymbol{K}_{1}^{-1}f_{1}(\boldsymbol{x}), \boldsymbol{K}_{C}^{**} - \boldsymbol{K}_{C}^{*\top}\boldsymbol{K}_{1}^{-1}\boldsymbol{K}_{C}^{*}\right)$$
(B.27c)

From these posterior distributions, we can directly obtain the means and the uncertainties of the Hodge components of the predicted edge function.

