Computational Graph Theory in the Twenty-First Century

Sawyer Jack Robertson
sawyerrobertson.xyz

1 Introduction

The term *computational graph theory* was commonly used from the 1980s to the 1990s as an umbrella term for such topics as the design and analysis of algorithms on graphs and computer-assisted advances in combinatorial graph theory (see, for example, the book collection [39]). Since then, the term has waned in popularity. Over the past few years, my body of work has aimed to deliver a principled, fresh perspective on this topic by combining methods and problems from areas such as graph-based algorithms in data science, statistical random graph theory, and geometric graph theory. This research program is driven by my fascination with graphs as mathematical objects that are simultaneously very simple (a graph is a collection of abstract points and some distinguished pairwise relationships between them) and yet deeply expressive in their ability to capture the structure of complex systems.

Generally speaking, my work has centered on three main research directions, each of which I describe in greater detail in the sections that follow. During my Ph.D., I am grateful to have been supported and mentored by my advisors Fan Chung and Alex Cloninger, who have provided complementary perspectives stemming, respectively, from combinatorics and spectral graph theory to signal processing and machine learning. As I look to the next phase of my career, I envision this program continuing to evolve and expand in exciting new directions.

Random graph theory (Section 2)

I am interested in random graphs primarily for two reasons: first, for their role in analyzing and benchmarking the performance of graph-based algorithms emerging from data science applications; and second, for their connections to stochastic processes, specifically Markov chains and algebraic combinatorics. In [27], I established operator norm concentration inequalities for adjacency and Laplacian matrices of inhomogeneous Erdős–Rényi-type *signed* graphs, providing insights into community detection in signed stochastic block models. In [22] (with D. Kohli, A. Cloninger, and G. Mishne), we also advanced this theme by analyzing the concentration of noisy graph kernel Laplacians to their clean counterparts, providing theoretical foundations for robust manifold learning.

Stochastically evolving graphs, the second component of my body of work in random graph theory, appear frequently in computer science as a means to develop sampling algorithms, dynamic network models, and randomized optimization methods. In [9] (with F. Chung), we introduced a model for dynamically evolving graphs based on random edits, deriving closed-form formulas for transition probabilities and mixing times using semigroup spectral theory. In doing so, we applied analysis techniques borrowed from algebraic combinatorics which had yet to be applied to random graph theory and opened the door to a variety of new analysis techniques for Markov chains on combinatorial state spaces.

Graph-based methods in data science (Section 3)

Graphs are ubiquitous in data science, serving as models for social networks, biological systems, transportation networks, and many other complex phenomena. Generically speaking, I am interested in optimization problems defined on the probability simplex of a graph: how to meaningfully construct such programs, how to solve them efficiently at scale, and how to leverage their solutions for downstream data science tasks. In [32] (with A. Cloninger and Z. Wan), we studied p-norm minimum cost flow problems on the probability simplex of a graph and connected these to optimal transport theory, in the process developing scalable algorithms and applying them to downstream clustering tasks. This led to a sequel [29] (with C. Holtz, Z. Wan, G. Mishne, and A. Cloninger), in which we used these problems to develop and analyze a robust graph-based semi-supervised learning method that balances edge sparsification with class separation, achieving state-of-the-art performance in low-label and corrupted-label regimes. These two efforts have highlighted rich and unexpected connections between optimal transport, effective resistance, and graph cuts, and the implications of these findings are still being explored.

Last but certainly not least, in [22] (with D. Kohli, A. Cloninger, and G. Mishne), we developed a spectral method for robust tangent space estimation on noisy data manifolds by orthogonalizing gradients of low-frequency graph Laplacian eigenvectors and provided theoretical guarantees and empirical evidence for improvements over local PCA.

Geometric graph theory (Section 4)

I am interested in geometric graph theory from at least two different angles. First, over the past three years or so I have become interested in various notions of curvature on graphs; particularly from the perspective of understanding the connections between the combinatorial properties of graphs and their curvature notions. Second, I am interested in foundational theory for the graph connection Laplacian, which is a matrix-valued operator that extends the classical graph Laplacian by incorporating geometric data in the form of parallel transport operators on edges.

In [11] (with A. Cloninger, G. Mishne, A. Oslandsbotn, Z. Wan, and Y. Wang), we formulated matrix-valued Dirichlet problems on connection graphs, defining connection voltage functions and conductance matrices that extend classical notions to the connection setting. In [30] (with D. Kohli, G. Mishne, and A. Cloninger), we generalized discrete optimal transport to vector fields on connection graphs and established duality results for these programs.

In [12] (with J. A. De Loera, J. Eddy, and J. A. Samper), we studied discrete curvature notions on convex polytopes, proving finiteness and abundance theorems for two different notions of discrete curvature. In [28], I derived explicit formulas for Wasserstein distances and curvature measures on trees, providing insights into their geometric properties. Finally, in [31] (with F. Southerland and E. Surya), we investigated a special class of graphs arising from a notion of curvature of Steinerberger [38] and generalized a number of known theorems in this field.

2 Statistical random graph theory and stochastically evolving graphs

Random graphs serve as fundamental models for networks that appear in a variety of domains. These arise both from the need for principled simulations of real-world phenomena and data and for benchmarking the performance and behavior of graph-based algorithms. My work in this area primarily focuses on concentration inequalities for random graph matrices (see Section 2.1) and spectral/mixing time analyses for stochastically evolving graphs (see Section 2.2).

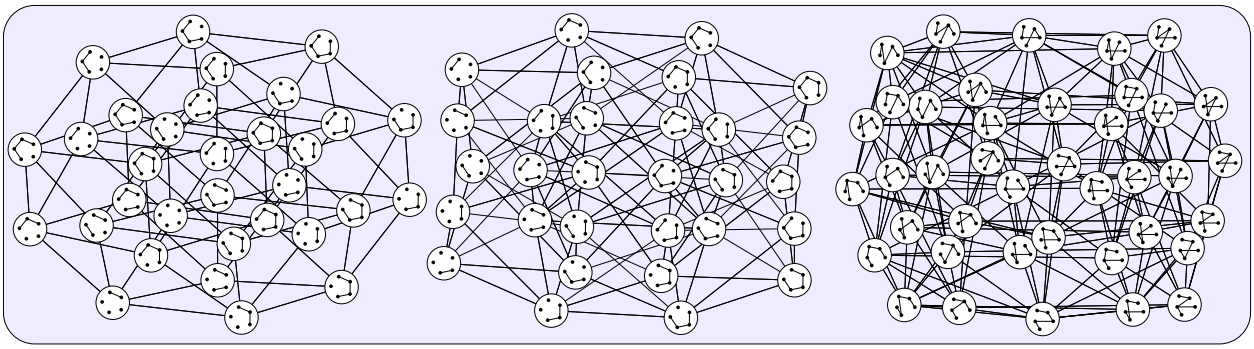

Figure 1: This figure contains three examples of *state graphs*, which are used to model the state spaces of stochastically evolving graph processes. The nodes correspond to labeled subgraphs of a host graph $\mathcal{H}$, and the links appear whenever it is possible to transition between the subgraphs according to the process under consideration. From left to right: host graph $\mathcal{H} = C_5$ according to a simple edit process of adding/deleting edges at random (a hypercube); host graph $\mathcal{H} = C_5$ corresponding to the Moran forest process of adding/deleting groups of edges in tandem; host graph given by the HouseX graph, with respect to the Glauber dynamics process on its spanning trees.

2.1 Concentration inequalities for matrix-valued statistics of random graphs

One of the oldest methods in graph-based data analysis is *spectral clustering*, which can be described at a high level as follows [34]. Given a weighted graph $G = (V, E, w)$ with n vertices, we define the *graph cut* between two disjoint subsets $S, T \subset V$ as

$$\texttt{cut}(S, T) = \sum_{\substack{i \in S \\ j \in T}} w_{ij}.$$

The *normalized cut* (Ncut) between S and its complement S^c is defined as

$$\texttt{Ncut}(S, S^c) = \texttt{cut}(S, S^c) \left(\frac{1}{\mathrm{vol}(S)} + \frac{1}{\mathrm{vol}(S^c)} \right),$$

where $\mathrm{vol}(S) = \sum_{i \in S} \sum_{j \in V} w_{ij}$ is the volume of the set S. The goal of spectral clustering is to find a partition of the graph that minimizes the `Ncut` value. Via a semidefinite relaxation, this problem can be viewed in terms of the graph Laplacian matrix $L = D - W$, where W is the weighted adjacency matrix of G and D is the diagonal degree matrix with entries $D_{ii} = \sum_j W_{ij}$. In turn, the leading eigenvector of L provides an approximate solution to the `Ncut` problem. We note that the variational characterizations of eigenvalues and eigenvectors of L are closely tied to connectivity properties and isoperimetric inequalities on graphs, such as Cheeger's inequality (see [10, 16]). There are also variants of spectral clustering for *signed graphs*, which instead seek to identify so-called *balanced communities* by targeting objectives that combine sparse edge removal and well-behaved signatures on the resulting components.

Certain random graph models are designed to simulate the environment of clustering or other downstream data analysis tasks, and it is of interest to certify that spectral clustering techniques can recover the ground truth information with high probability. These include stochastic block models and so-called *information-plus-noise models*, which consider randomly weighted geometric graphs. There is a rich history on this topic in the statistics literature (see, for starters, [8, 13, 23]).

As a variant, one may consider *inhomogeneous Erdős–Rényi-type signed graphs*, which are random graphs where each pair of vertices is assigned an edge and a positive or negative sign according to some probability distribution. These models are useful for studying networks with both positive and negative relationships, such as social networks with friendships and rivalries. In [27], I

established operator norm concentration inequalities for adjacency and Laplacian matrices of inhomogeneous Erdős–Rényi-type signed graphs, providing insights into community detection in signed stochastic block models.

In [22], detailed in greater length in Section 3.2, we also advanced this theme by analyzing the concentration of noisy graph kernel Laplacians to their clean counterparts, providing theoretical foundations for robust manifold learning. In particular, we adopted an information-plus-noise model consisting of a dataset $\{X_i\}_{i=1}^n$ given by $X_i = y_i + Z_i$, where $\{y_i\}_{i=1}^n$ are "clean" points on the manifold and $\{Z_i\}_{i=1}^n$ are independent sub-Gaussian noise vectors. We proved that under certain conditions on the noise distribution and the kernel used to construct the corresponding Laplacian matrix, the noisy graph Laplacian L_{noisy} concentrates to the clean graph Laplacian L_{clean} with high probability as $n \to \infty$, and hence by Weyl's inequality and the Davis-Kahan theorem, so do the corresponding eigenvalues and eigenvectors. This result was instrumental in establishing theoretical guarantees for the robustness of our tangent space estimation method.

Theorem 2.1 (Informal statement of eigenvector stability [22])**.** *Let $\{y_i\}_{i=1}^n$ be fixed with $y_i \in \mathbb{R}^d$ and write*

$$X_i = y_i + Z_i,$$

where $\{Z_i\}_{i=1}^n$ are i.i.d. mean-zero sub-Gaussian noise vectors with variance proxy $\xi \geq 0$. Assume that $\xi = o(1)$ and d is fixed as $n \to \infty$. Let L_{clean} and L_{noisy} be the normalized kernel graph Laplacians constructed from $\{y_i\}_{i=1}^n$ and $\{X_i\}_{i=1}^n$, respectively. Then, under certain assumptions on the eigenvalues of L_{clean}, the leading $O(1)$ eigenvectors of L_{noisy} converge to those of L_{clean} with high probability as $n \to \infty$.

2.2 New directions in stochastically evolving graphs

Stochastically evolving graphs are created when the structure of a graph changes over time according to some stochastic process. In theoretical computer science, these models appear frequently in the form of sampling algorithms for structures such as matchings [15], spanning trees [1, 5], and colorings [7]. In applied settings, these models are useful for studying dynamic networks, such as social networks [20], communication networks [19], and biological networks [26].

Often a key objective in the analysis of these models is the *mixing time* of the underlying Markov chain, which measures how quickly the graph converges to its stationary distribution. Understanding mixing times is crucial for designing efficient algorithms and for characterizing the long-term behavior of dynamic networks. Classical estimates relate the mixing time of an ergodic, reversible Markov chain P to its largest nontrivial eigenvalue, thereby motivating the investigation of the spectral decomposition of Markov kernels as a means to certify their mixing times. In many cases, however, the state space of the Markov chain is combinatorial in nature, making direct spectral analysis challenging.

In [9], we introduced a model for dynamically evolving graphs based on random edits and derived closed-form formulas for transition probabilities and mixing times using semigroup spectral theory. Specifically, we considered a random walk on the lattice of subgraphs of a fixed graph G where at each step a randomly chosen "edit" (adding/removing one edge, or compound edits affecting many edges) is applied. By showing that the edits form a left regular band semigroup, we derived closed-form formulas for the eigenvalues of the associated transition probability matrix and sharp mixing time estimates. The framework extends to compound edits such as the Moran forest model (see [4]) and random intersection graph processes, with mixing governed by the largest nontrivial eigenvalue. Below I provide an informal statement of one of our main results.

Theorem 2.2 (Informal statement, eigenvalues of compound edit processes [9])**.** *Let $G = (V, E)$ denote a fixed host graph and let* A *be a given set of* edits *(defined as idempotent maps on 2^E). Let $w \in \mathbb{R}^{\mathsf{A}}$ be a given probability distribution on* A*. Consider the random walk obtained as follows: start with a labeled subgraph G_0 of G, and at each time $t \geq 1$, sample an edit $x \sim w$ and write $G_t = xG_{t-1}$. Denote the transition probability matrix of this process by $\mathcal{P}$. Then $\mathcal{P}$ is diagonalizable and its eigenvalues are related to a certain join semilattice associated to G and* A*, and can be computed in closed form.*

2.3 Future directions

- The framework developed in [9] suggests that a variety of related stochastic processes on graphs rapidly mix. Can we leverage these insights to design new sampling algorithms for combinatorial structures or diffusion-based generative graph models?
- If G is a fixed graph and we retain each edge with probability p (or $k = O(1)$ neighbors of each vertex), how low can p (or k) be while still ensuring that the low-frequency eigenvectors of the percolated graph approximate those of G with high probability as the number of vertices grows? These are related to *randomized near neighbor graphs* (see [21]).

3 Graph-based methods in data science

Graphs have been used to develop algorithms and models for a variety of data science tasks, including clustering, classification, and dimensionality reduction. My work in this area primarily focuses on optimal transport and graph-based semi-supervised learning (see Section 3.1), and geometric data analysis (see Section 3.2).

3.1 Optimal transport and graph-based SSL

Optimal transport (OT) has emerged as a powerful tool for comparing and analyzing probability distributions, with applications in machine learning, computer vision, and economics. In the graph setting, one can define OT problems on the probability simplex $\mathcal{P}(V)$ of the graph, which consists of all probability measures supported on the vertices V. Specifically, if $\mu, \nu \in \mathcal{P}(V)$, then the *1-Wasserstein distance* $W_1(\mu, \nu)$ is given by

$$W_1(\mu, \nu) = \inf_{\pi \in \Pi(\mu,\nu)} \sum_{i,j \in V} d_G(i, j)\pi(i, j), \tag{1}$$

where $\Pi(\mu, \nu)$ is the set of couplings between μ and ν, and $d_G(i, j)$ is the shortest path distance between vertices i and j. In the graph setting, the program Eq. (1) is equivalent to the following minimum-cost flow problem (also known as the *Beckmann problem*):

$$W_1(\mu, \nu) = \min_{J:E\to\mathbb{R}} \left\{ \sum_{e \in E} w_e |J(e)| : \mathrm{div}(J) = \mu - \nu \right\}. \tag{2}$$

Here, J is a flow on the edges of the graph, w_e denotes the edge weights, and $\mathrm{div}(J)$ is the divergence of the flow at each vertex defined by the action of the vertex-edge oriented incidence matrix. In [32], we explored connections between $W_1(\mu, \nu)$ and extensions of Eq. (2) to include p-norm costs on flows, which we termed p-Beckmann problems. Specifically, if $p \in [1, \infty)$, we defined

$$B_p(\mu, \nu) = \min_{J:E\to\mathbb{R}} \left\{ \sum_{e \in E} w_e |J(e)|^p : \mathrm{div}(J) = \mu - \nu \right\}.$$

For $p = \infty$, we set

$$B_\infty(\mu, \nu) = \min_{J:E\to\mathbb{R}} \left\{ \max_{e\in E} |J(e)| : \operatorname{div}(J) = \mu - \nu \right\}.$$

We developed duality theory for these problems and demonstrated their utility in various clustering tasks. Moreover, we showed that the 2-Beckmann distance generalizes effective resistance on graphs to the probability simplex and linked it to random walks and homogeneous Sobolev seminorms.

The p-Beckmann problems also served as the basis for a graph-based *semi-supervised learning* (SSL) method in [29]. SSL, generally speaking, aims to leverage both labeled and unlabeled data to improve model performance. In graph-based SSL, one typically has a small set of labeled vertices $L \subset V$ with known labels $\{y_i\}_{i\in L}$ and a larger set of unlabeled vertices $U = V \backslash L$. The goal is to infer labels for the unlabeled vertices based on the graph structure and the labeled data. There is a rich history of methods developed along these lines (see, e.g., [6, 40] for a small selection of starting points). In [29], we proposed a p-conductance learning method that balances edge sparsification with class separation (see Fig. 2 for an illustration). Specifically, we considered the following p-conductance program between measures $\mu, \nu \in \mathcal{P}(V)$:

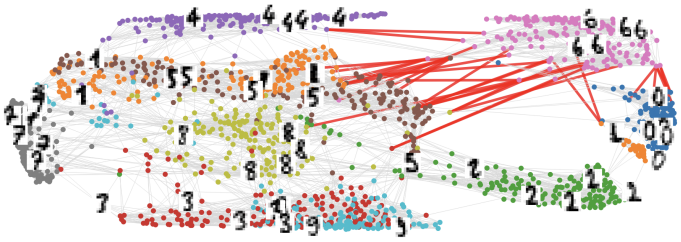

Figure 2: This figure contains a graph with vertices corresponding to the entries of a small digits dataset [2]. Edges are included whenever two images are similar. We seek to classify images of the digit 6 from images of all other digits. The reference measure μ is uniform on the sixes with their images overlaid in the figure, and the reference measure ν is uniform on those vertices with images overlaid as well. We highlight in red edges that are cut according to the program $C_1(\mu, \nu)$.

$$C_p(\mu, \nu) = \min_{\substack{\phi\in\mathbb{R}^V \\ \phi^\top(\mu-\nu)\geq 1}} \sum_{e=\{x,y\}\in E} w_e |\phi(x) - \phi(y)|^p.$$

By leveraging gauge duality (see [18]), we established a relationship between $C_p(\mu, \nu)$ and $B_p(\mu, \nu)$, connecting potentials used in our SSL method to the p-Beckmann problems (see Theorem 3.1).

Theorem 3.1 (Equivalences between Beckmann and Conductance programs [29])**.** *Let $\mu, \nu \in \mathcal{P}(V)$ be distinct. Then the optimal values of C_p and B_p are related as:*

$$C_p(\mu, \nu) = \begin{cases} 1/B_{\infty, w^{-1}}(\mu, \nu) & \text{if } p = 1, \\ 1/B_{q, w^{1-q}}(\mu, \nu) & \text{if } p \in (1, \infty) \text{ and } 1/p + 1/q = 1, \\ 1/B_{1, w^{-1}}(\mu, \nu) & \text{if } p = \infty. \end{cases}$$

In particular, if $G = (V, E)$ has unit weights, then $C_\infty(\mu, \nu) = 1/W_1(\mu, \nu)$ and

$$C_2(\mu, \nu)^2 = 1/B_2(\mu, \nu)^2 = 1/(\mu - \nu)^\top L^\dagger (\mu - \nu),$$

where $L^\dagger$ is the Moore-Penrose inverse of L.

3.2 Geometric data analysis

In many data science applications, one encounters data that lies on or near a low-dimensional manifold embedded in a high-dimensional ambient space. Capturing the geometric structure of such data is crucial for effective analysis and learning. One of the specific subtasks in this area is *tangent space estimation*, which involves estimating the tangent spaces of the manifold at various points based on the graph structure.

Consider a method such as *local PCA* [36], which is designed to estimate tangent spaces as follows. Given a collection of data points $\{x_i\}_{i=1}^n$ in $\mathbb{R}^d$ that lie on or near a k-dimensional manifold $\mathcal{M}$, one constructs a graph $G = (V, E, w)$ as before. For each vertex $i \in V$, one identifies its local neighborhood $N(i)$ (e.g., the k-nearest neighbors) and performs PCA on the points $\{x_j : j \in N(i)\}$ to estimate the tangent space $T_{x_i}\mathcal{M}$ after first estimating the intrinsic dimension k of the manifold beforehand. While local PCA is straightforward and widely used, it can be sensitive to noise and perturbations in the data, leading to inaccurate tangent space estimates.

Our work on this topic begins by utilizing *kernel Laplacians*, which are discrete approximations of the Laplace-Beltrami operator on the manifold constructed from the graph. Specifically, one defines the graph Laplacian L as

$$L = D - W, \tag{3}$$

where W is the weighted adjacency matrix of the graph and D is the diagonal degree matrix with entries $D_{ii} = \sum_j W_{ij}$. The normalized graph Laplacian is given by

$$L_{\text{norm}} = I - D^{-1/2} W D^{-1/2}.$$

When the kernel used to construct the weights W is chosen appropriately (e.g., the random walk kernel), the graph Laplacian converges to the Laplace-Beltrami operator on the manifold as the number of data points increases and the neighborhood size decreases (see, e.g., [3]).

Eigenvectors of the kernel Laplacian capture important geometric information about the manifold, and their gradients can be used to estimate tangent spaces. Moreover, we observe empirically that these low-frequency eigenvectors are more robust to noise than the eigenvectors appearing deeper in the spectrum. This observation motivates our approach in [22], where we developed a spectral method for robust tangent space estimation on noisy manifolds. Our method, termed `LEGO`, proceeds by orthogonalizing the gradients of the low-frequency eigenvectors of the graph Laplacian to estimate the tangent spaces at each vertex. Specifically, for each vertex $i \in V$, we compute the gradients $\nabla\phi_j(x_i)$ of the first m eigenvectors $\{\phi_j\}_{j=1}^m$ of the graph Laplacian and then apply a localized orthogonalization procedure to obtain an estimate of the tangent space $T_{x_i}\mathcal{M}$. We provided two theoretical arguments to support our empirical observations: *(i)* a differential geometric analysis on tubular neighborhoods showing that eigenfunctions with large gradients in noise directions lie deeper in the spectrum; *(ii)* a random matrix analysis proving the noisy graph Laplacian concentrates to the clean operator w.h.p. in certain noise regimes, so low-frequency eigenvectors (and their gradients) are noise stable (see Theorem 2.1).

3.3 Future directions

- In ongoing work (with C. Holtz and M. Black), I am investigating the relationship between existing notions of edge centrality metrics such as the *betweenness centrality* and *biharmonic distance*. We aim to establish equivalences such as the ones outlined in Theorem 3.1.
- We are also currently working on formalizing guarantees for the observed robustness of the graph cut-based SSL method proposed in [29].

4 Geometric graph theory: connections and curvatures

Graphs can be equipped, both intrinsically and extrinsically, with geometric features that capture various properties of the underlying graph or data. My work in this area primarily focuses on connection Laplacians (see Section 4.1) and discrete curvature notions (see Section 4.2).

4.1 Connection Laplacians

One way to incorporate extrinsic geometric information into graph-based methods is through the usage of *connection Laplacians*, which are Laplacians that encode parallel transport operators between vector spaces at each vertex. Specifically, a *connection graph* is a tuple (G, σ) where $G = (V, E, w)$ is a weighted graph and $\sigma : E \to O(d)$ is a *connection* assigning an orthogonal transformation to each edge. The connection Laplacian L_σ is defined as

$$(L_\sigma f)(i) = \sum_{j:\{i,j\}\in E} w_{ij}(f(i) - \sigma_{ij} f(j)),$$

for $f : V \to \mathbb{R}^d$. Connection Laplacians have been used in various applications, including computer vision [35], manifold learning [36], and shape analysis [33].

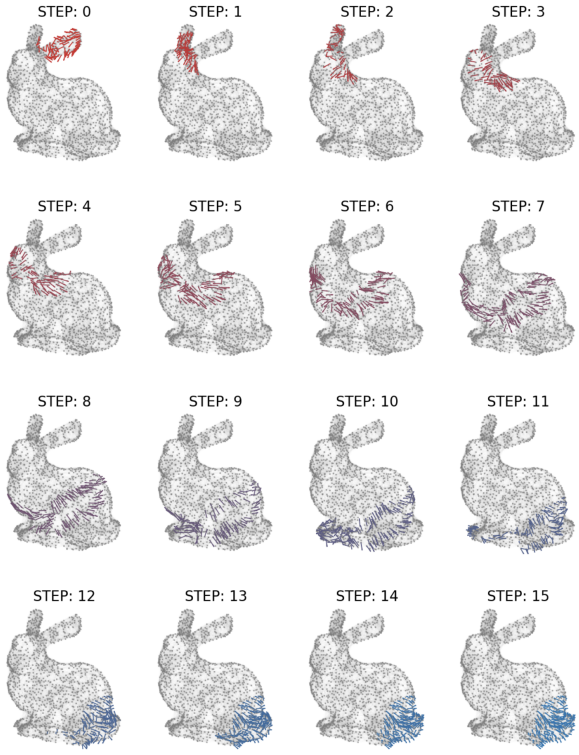

Figure 3: By defining a notion of *minimum cost parallel transport* between vector fields on a geometric graph [30], we introduced an algorithm for vector field interpolation on graphs. Here, we illustrate such an interpolation between vector fields defined on the Stanford bunny [37].

In [11], we formulated various matrix-valued Dirichlet problems on connection graphs and used their solutions to define connection voltage functions and conductance matrices that extend classical notions to the connection setting. We showed that these matrices encode the effective conductance between nodes while also reflecting the connection structure, and we established invariance properties under switching equivalence and direct sums of connections. We also extended the notion of effective resistance to connection graphs by defining the *connection resistance matrix* R_σ based on the solution of a similar Poisson-type equation. We proved that R_σ is, up to a structured congruence, the pseudoinverse of the conductance matrix.

In [30], we generalized discrete optimal transport to vector fields on connection graphs by defining the *connection Beckmann problem* as follows:

$$W_{1,\sigma}(\alpha, \beta) = \min_{J:E\to\mathbb{R}^d} \left\{ \sum_{e\in E} w_e \|J(e)\| : \mathrm{div}_\sigma(J) = \alpha - \beta \right\},$$

where div_σ is the connection divergence operator that incorporates the connection σ. This problem models the optimal parallel transport cost between vector fields $\alpha, \beta : V \to \mathbb{R}^d$ on a data manifold. We established feasibility criteria and duality results for the connection Beckmann problem, as well as a relaxed and regularized variant thereof. We demonstrated that our work can be applied to various data science tasks with a geometric component, such as color image processing and directional data analysis (see Fig. 3).

4.2 Discrete curvatures

Curvature is a fundamental concept in differential geometry that quantifies how a geometric object deviates from being flat. In recent years, there has been rapidly growing interest in developing discrete notions of curvature for graphs and networks, which can provide insights into their structural properties. Various discrete curvature notions have been proposed, including Ollivier-Ricci curvature [25], Lin–Lu–Yau curvature [24], Forman-Ricci curvature [17], and more recently, Steinerberger curvature [38] and effective resistance curvature [14].

In [12], we studied discrete curvature notions (specifically, resistance curvature and Forman–Ricci curvature) on convex polytopes, proving finiteness and abundance theorems in various dimensions. This marked the first systematic study of discrete curvature notions on polytopes, revealing new connections between combinatorial geometry and discrete curvature. In [28], I derived explicit formulas for Wasserstein distances and curvatures on trees, and obtained comparison results between different curvature notions. This marked one of only a handful of papers which obtain explicit quantitative comparisons between multiple notions of discrete curvature in one setting.

Finally, in [31] (in preparation, with F. Southerland and E. Surya), we investigated a special class of graphs arising from a notion of curvature of Steinerberger [38]. Specifically, we studied *distance exceptional graphs*, which are graphs where the equation $Dx = \mathbf{1}$ does not have a solution. Here, D is the distance matrix of the graph and $\mathbf{1}$ is the all-ones vector. These graphs appear naturally within the curvature framework of Steinerberger, but are mysterious in many respects. Prior to our work, constructions of these graphs were extremely constrained and poorly understood. We introduced a graph invariant called the *curvature index*, which measures the extent to which a graph deviates from being distance exceptional (and, in particular, vanishes when G is distance exceptional). In doing so, we generalized a number of known theorems in this field and proved an embedding theorem, stated below.

Theorem 4.1 (Informal statement of distance exceptional embeddings [31])**.** *Let G be any graph, connected or otherwise. Then there exists a distance exceptional graph G' such that G is isomorphic to an induced subgraph of G'. Moreover, if G satisfies $\iota(G) < \infty$, where $\iota(G)$ is the curvature index of G (definition omitted), then the embedding can be chosen to be an isometry.*

4.3 Future directions

- State graphs (see Fig. 1) provide a rich class of large graphs related to stochastically evolving graph processes (see Section 2.2). Can we characterize their discrete curvature properties and relate them to mixing times of the underlying processes?
- In the Erdős–Rényi random graph $G(n, p)$, can we understand the behavior of the curvature index as $n \to \infty$? What fraction of graphs are distance exceptional?

References

[1] D. J. Aldous, *The random walk construction of uniform spanning trees and uniform labelled trees*, SIAM J. Discrete Math., 3 (1990), pp. 450–465.

[2] E. Alpaydin and C. Kaynak, *Optical recognition of handwritten digits*. UCI Machine Learning Repository, 1998, https://archive.ics.uci.edu/dataset/80/optical+recognition+of+handwritten+digits. [Dataset].

[3] M. Belkin and P. Niyogi, *Towards a theoretical foundation for Laplacian-based manifold methods*, J. Comput. System Sci., 74 (2008), pp. 1289–1308.

[4] F. c. Bienvenu, J.-J. Duchamps, and F. Foutel-Rodier, *The Moran forest*, Random Structures Algorithms, 59 (2021), pp. 155–188.

[5] A. Z. Broder, *Generating random spanning trees*, in Foundations of Computer Science, vol. 89, 1989, pp. 442–447.

[6] J. Calder, B. Cook, M. Thorpe, and D. Slepcev, *Poisson learning: Graph based semi-supervised learning at very low label rates*, in International conference on machine learning, PMLR, 2020, pp. 1306–1316.

[7] C. Carlson, D. Frishberg, and E. Vigoda, *Improved distributed algorithms for random colorings*, in 27th International Conference on Principles of Distributed Systems, vol. 286 of LIPIcs. Leibniz Int. Proc. Inform., Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2024, pp. Art. No. 13, 18.

[8] F. Chung and M. Radcliffe, *On the spectra of general random graphs*, Electron. J. Combin., 18 (2011).

[9] F. Chung and S. J. Robertson, *Stochastically evolving graphs via edit semigroups*, Proceedings of the National Academy of Sciences of the USA, (2025), https://arxiv.org/abs/2509.19678. *To appear.*

[10] F. R. K. Chung, *Spectral graph theory*, vol. 92 of CBMS Regional Conference Series in Mathematics, Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1997.

[11] A. Cloninger, G. Mishne, A. Oslandsbotn, S. J. Robertson, Z. Wan, and Y. Wang, *Random walks, conductance, and resistance for the connection graph laplacian*, SIAM Journal on Matrix Analysis, 45(3) (2024), https://arxiv.org/abs/2308.09690.

[12] J. A. De Loera, J. Eddy, S. J. Robertson, and J. A. Samper, *Discrete curvatures and convex polytopes*, arXiv preprint, (2025), https://arxiv.org/abs/2510.11894.

[13] S. Deng, S. Ling, and T. Strohmer, *Strong consistency, graph Laplacians, and the stochastic block model*, J. Mach. Learn. Res., 22 (2021), pp. Paper No. 117, 44.

[14] K. Devriendt, A. Ottolini, and S. Steinerberger, *Graph curvature via resistance distance*, Discrete Appl. Math., 348 (2024), pp. 68–78.

[15] M. Dyer, L. A. Goldberg, and M. Jerrum, *Counting and sampling H-colourings*, Inform. and Comput., 189 (2004), pp. 1–16.

[16] M. Fiedler, *Algebraic connectivity of graphs*, Czechoslovak Math. J., 23(98) (1973), pp. 298–305.

[17] R. Forman, *Bochner's method for cell complexes and combinatorial Ricci curvature*, Discrete Comput. Geom., 29 (2003), pp. 323–374.

[18] M. P. Friedlander, I. Macêdo, and T. K. Pong, *Gauge optimization and duality*, SIAM J. Optim., 24 (2014), pp. 1999–2022.

[19] M. Grossglauser and D. N. Tse, *Mobility increases the capacity of ad hoc wireless networks*, IEEE/ACM transactions on networking, 10 (2002), pp. 477–486.

[20] S. Hanneke, W. Fu, and E. P. Xing, *Discrete temporal models of social networks*, Electron. J. Stat., 4 (2010), pp. 585–605.

[21] A. Jaffe, Y. Kluger, G. C. Linderman, G. Mishne, and S. Steinerberger, *Randomized near-neighbor graphs, giant components and applications in data science*, J. Appl. Probab., 57 (2020).

[22] D. Kohli, S. J. Robertson, G. Mishne, and A. Cloninger, *Robust tangent space estimation via laplacian eigenvector gradient orthogonalization*, arXiv preprint, (2025), https://arxiv.org/abs/2412.20620.

[23] J. Lei and A. Rinaldo, *Consistency of spectral clustering in stochastic block models*, Ann. Statist., 43 (2015), pp. 215–237.

[24] Y. Lin, L. Lu, and S.-T. Yau, *Ricci curvature of graphs*, Tohoku Math. J. (2), 63 (2011), pp. 605–627.

[25] Y. Ollivier, *Ricci curvature of Markov chains on metric spaces*, J. Funct. Anal., 256 (2009), pp. 810–864.

[26] B.-E. Perrin, L. Ralaivola, A. Mazurie, S. Bottani, J. Mallet, and F. d'Alché Buc, *Gene networks inference using dynamic bayesian networks*, Bioinformatics-Oxford, 19 (2003), pp. 138–148.

[27] S. J. Robertson, *Matrix concentration for random signed graphs and community recovery in the signed stochastic block model*, arXiv preprint, (2024), https://arxiv.org/abs/2412.20620.

[28] S. J. Robertson, *A comparative study of curvature on trees*, Bulletin of the Institute of Combinatorics and its Applications, (2025), https://arxiv.org/abs/2412.20661. *To appear.*

[29] S. J. Robertson, C. Holtz, Z. Wan, G. Mishne, and A. Cloninger, *Robust graph-based semi-supervised learning via p-conductances*, arXiv preprint, (2025), https://arxiv.org/abs/2502.08873.

[30] S. J. Robertson, D. Kohli, G. Mishne, and A. Cloninger, *On a generalization of Wasserstein distance and the beckmann problem to connection graphs*, SIAM Journal on Scientific Computing, 47(5) (2025), https://arxiv.org/abs/2312.10295.

[31] S. J. Robertson, F. Southerland, and E. Surya, *Distance exceptional graphs and the curvature index*, arXiv preprint, (2025), https://arxiv.org/abs/2511.03719.

[32] S. J. Robertson, Z. Wan, and A. Cloninger, *All you need is resistance: On the equivalence of effective resistance and certain optimal transport problems on graphs*, arXiv preprint, (2024), https://arxiv.org/abs/2404.15261.

[33] N. Sharp, Y. Soliman, and K. Crane, *The vector heat method*, ACM Transactions on Graphics (TOG), 38 (2019), pp. 1–19.

[34] J. Shi and J. Malik, *Normalized cuts and image segmentation*, IEEE Transactions on pattern analysis and machine intelligence, 22 (2000), pp. 888–905.

[35] A. Singer, *Angular synchronization by eigenvectors and semidefinite programming*, Appl. Comput. Harmon. Anal., 30 (2011).

[36] A. Singer and H.-T. Wu, *Vector diffusion maps and the connection Laplacian*, Comm. Pure Appl. Math., 65 (2012), pp. 1067–1144.

[37] Stanford Computer Graphics Laboratory, *The stanford 3d scanning repository*, 1994, https://graphics.stanford.edu/data/3Dscanrep/ (accessed 2025-11-07).

[38] S. Steinerberger, *Curvature on graphs via equilibrium measures*, J. Graph Theory, 103 (2023), pp. 415–436.

[39] G. Tinhofer, E. Mayr, and H. Noltemeier, eds., *Computational graph theory*, vol. 7 of Computing Supplementum, Springer-Verlag, Vienna, 1990.

[40] X. Zhu, Z. Ghahramani, and J. D. Lafferty, *Semi-supervised learning using gaussian fields and harmonic functions*, in Proceedings of the 20th International conference on Machine learning (ICML-03), 2003, pp. 912–919.