Computational Graph Theory in the Twenty-First Century

Sawyer Jack Robertson
sawyerrobertson.xyz

1 Introduction

The term *computational graph theory* was commonly used in the 1980s until the 1990s as an umbrella 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 [38]). Since then, the term has largely fallen out of favor. 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, which I describe in greater detail in the sections that follow. As I look to the next phase of my career, I envision this program continuing to evolve and expand in exciting new directions. This statement contains a detailed overview of my past contributions and ongoing work in these areas, as well as a discussion of potential future directions that I am eager to explore.

Graph-based methods in data science 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 [31] (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 [28] (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.

In [21] (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.

Statistical random graph theory and stochastically evolving graphs Random graphs serve as useful models for networks that appear across a variety of domains, arising both from the need for principled simulations of real-world phenomena and data, as well as the need for benchmarking the performance and behavior of graph-based algorithms. In [26], 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 [21], we also contributed to 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, another entry point to random graph theory from the Markovian side of things, 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 in doing so, opened the door to a variety of new analysis techniques for Markov chains on combinatorial state spaces.

Geometric graph theory and curvature The intrinsic geometric properties of graphs, such as curvature, play a crucial role in understanding the structural features of real-world graphs. I have also studied geometric graph models that incorporate the extrinsic contribution of geometric data, namely, connection graphs, which are typically used to model data manifolds along with their parallel transport operators. 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 [29] (with D. Kohli, G. Mishne, and A. Cloninger), we generalized discrete optimal transport to vector fields on connection graphs, establishing duality results and revealing new phenomena related to flux constraints. 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 [27], I derived explicit formulas for Wasserstein distances and curvature measures on trees, providing insights into their geometric properties. Finally, in [30] (with F. Southerland and E. Surya), we investigated a special class of graphs arising from a notion of curvature of Steinerberger [37] and generalized a number of known theorems in this field.

2 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, graph-based semi-supervised learning (SSL), and geometric data analysis. Below, I highlight a few questions that motivate me and other researchers working in this area.

- Many graph-based SSL algorithms are designed by using the pre-labeled vertices to define a boundary value problem or variational equivalent. Often these formulations are interrelated in various ways. Can we develop a unified framework that captures these relationships and provides insights into their performance and limitations?
- Consequently, as the scale of the data increases, how can we design graph-based SSL algorithms that are both computationally efficient and scalable while maintaining high accuracy and robustness to noise?
- When designing algorithms for local tangent space estimation on manifolds, there are fundamental tradeoffs between *locality* (i.e., how small of a neighborhood one uses to estimate the tangent space) and *robustness* (i.e., how sensitive the method is to noise and perturbations in the data). How can we balance these competing demands to achieve accurate and reliable tangent space estimates?

In the subsections to follow, I discuss some of my contributions to these areas in greater detail.

2.1 Optimal transport and graph-based SSL

Often in data science pipelines, one begins with a fixed set of points $\{x_i\}_{i=1}^n$ in some ambient space (e.g., $\mathbb{R}^d$) and constructs a weighted graph $G = (V, E, w)$ with vertices $V = \{1, \dots, n\}$ corresponding to the data points and edges and weights E, w defined according to some similarity measure (e.g., k-nearest neighbors, ϵ-neighborhoods, or fully connected graphs with weights given by a kernel function). Once the graph is constructed, one can leverage its structure to perform various tasks.

Optimal transport 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 of the graph, which consists of all probability measures supported

on the vertices V. Specifically, if $\mu, \nu \in \mathcal{P}(V)$, where $\mathcal{P}(V)$ is the probability simplex on V, then the 1-Wasserstein distance $W_1(\mu, \nu)$ can be defined as

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

where $\Pi(\mu, \nu)$ is the set of couplings between μ and ν, and $d_G(i, j)$ is the *weighted* graph 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):

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

where J is a flow on the edges of the graph, w_e are the edge weights, and $\operatorname{div}(J)$ is the divergence of the flow at each vertex defined by the action of the vertex-edge oriented incidence matrix. In [31], we explored connections between $W_1(\mu, \nu)$ and extensions of Eq. (2) involving p-norm costs on flows, which we termed p-Beckmann problems. Specifically, if $p \in [1, \infty)$, we defined

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

If $p = \infty$, 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 clustering tasks, providing scalable algorithms and empirical results. Moreover, we showed that the 2-Beckmann distance generalizes effective resistance on graphs to the probability simplex, linking it to random walks and negative Sobolev seminorms.

The p-Beckmann problems also served as the basis for a graph-based SSL method in [28]. *Semi-supervised learning*, generally speaking, aims to leverage both labeled and unlabeled data to improve learning 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 $U = V \setminus L$. 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, 39] for a small selection of starting points). In [28], we proposed a p-conductance learning method that balances edge sparsification with class separation. Specifically, we considered the following p-conductance program between measures $\mu, \nu \in \mathcal{P}(V)$:

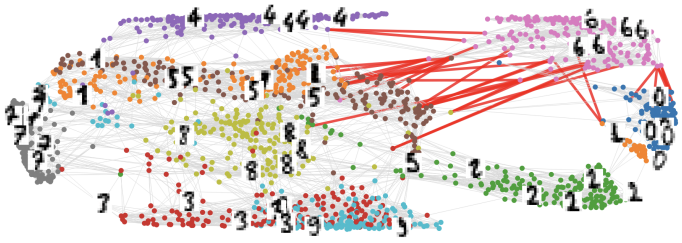

Figure 1: This figure contains a graph with vertices corresponding to the entries of the Sklearn Digits Dataset [2]. Edges are included whenever two images are similar. We seek to classify images of the digit 6 from images of all other digits. The reference measure μ is uniform on those sixes with their 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, we established a reciprocal relationship between $C_p(\mu,\nu)$ and $B_p(\mu,\nu)$, connecting our SSL method to the p-Beckmann problems (see Theorem 2.1). We developed a scalable solver for the optimization problem and demonstrated state-of-the-art performance in low-label, corrupted-label, and partial-label regimes.

Theorem 2.1 (Equivalences between Beckmann and Conductance programs [28])**.** *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 pseudoinverse of the graph Laplacian L.

2.2 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* [35], 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 $T_{x_i}\mathcal{M}$ having estimated 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., [4]).

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 [21], 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 tangent space $T_{x_i}\mathcal{M}$. We verified our initial empirical observations by providing two theoretical pillars: *(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 2.2).

Theorem 2.2. *(Informal statement of eigenvector stability [21]) Let* $\{y_i\}_{i=1}^n$ *be fixed with* $y_i \in mathbbR^d$ *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* $n \to \infty$*.*

3 Statistical random graph theory and stochastically evolving graphs

Random graphs serve as fundamental models for networks appearing across a variety of domains. These arise both from the need for principled simulations of real-world phenomena and data, as well as the need for benchmarking the performance and behavior of graph-based algorithms. My work in this area primarily focuses on concentration inequalities for random graph matrices, community detection in signed stochastic block models, and stochastically evolving graphs. Below, I highlight a few questions that motivate me and other researchers working in this area.

- The graph connection Laplacian has been used to design spectral algorithms for solving synchronization problems. Inspired by the statistics literature on spectral clustering for unsigned graphs, can we develop concentration inequalities for connection graph Laplacians of random graphs to certify the performance of these algorithms?
- There has been an explosion in recent years concerning connections between up-down walks on simplicial complexes and random sampling algorithms for combinatorial structures. The most famous case concerns sampling spanning trees; are there other opportunities to further develop and apply these techniques to other families of graphs (e.g., degree regular, planar graphs)?
- One problem I am particularly interested in is the statistical consistency of spectral estimators derived from percolation-type models on host graphs. Specifically, 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?

In the subsections to follow, I discuss some of my contributions to these areas in greater detail.

3.1 Concentration inequalities for matrix-valued statistics of random graphs

One of the most historic methods in graph-based data analysis is *spectral clustering*, which originates from the following derivation [33]. Given a weighted graph $G = (V, E, w)$ with n vertices, one defines 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

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

where $\text{vol}(S) = \sum_{i \in S} \sum_{j \in V} w_{ij}$ 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 L as in Eq. (3). As such, the leading eigenvector of L provides an approximate solution to the `Ncut` problem. We note that the variational characterizations of eigenvalues and

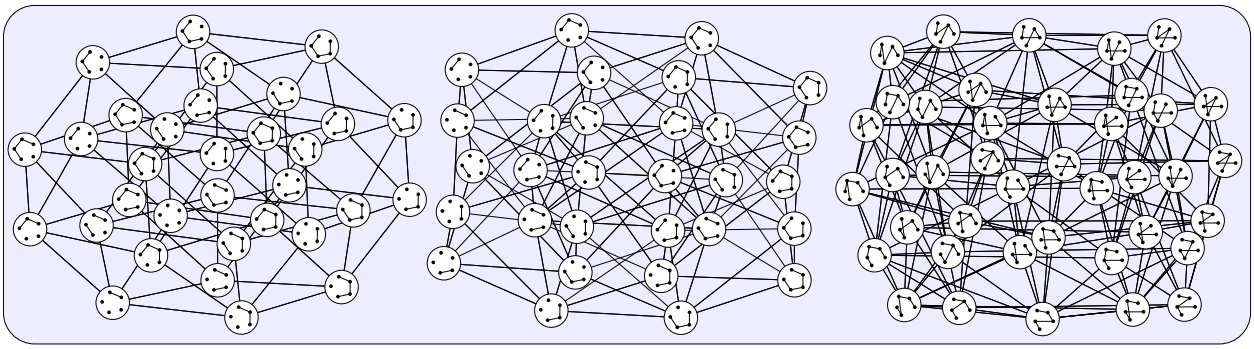

Figure 2: This figure contains three examples of *state graphs*, which are used to model stochastically evolving graph processes driven by, e.g., random edits or structural changes. The nodes of these graphs correspond to permissible labeled subgraphs of a host graph $\mathcal{H}$, and the links appear 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; 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.

eigenvectors of L are closely tied to connectivity properties and isoperimetric inequalities on graphs, such as Cheeger's inequality (see [10, 16]).

It is of interest to determine the extent to which the eigenvalues and eigenvectors of L concentrate around their expected values when G is a random graph as a means to certify the performance of spectral clustering and related methods. There is a rich history on this topic in the statistics literature (see [8, 13, 22]). 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 [26], 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 the separate setting of local tangent space estimation and the `LEGO` algorithm discussed in Section 2.2, we also contributed to this theme by analyzing the concentration of noisy graph kernel Laplacians to their clean counterparts, providing theoretical foundations for robust manifold learning in [21]. In particular, we adopted an information-plus-noise model where the data $\{X_i\}_{i=1}^n$ is 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 graph Laplacian, the noisy graph Laplacian L_{noisy} concentrates to the clean graph Laplacian L_{clean} with high probability as $n \to \infty$. This result provided theoretical guarantees for the robustness of our tangent space estimation method.

3.2 New directions in stochastically evolving graphs

Stochastically evolving graphs are random graph models where the graph structure changes over time according to some stochastic process. In theoretical computer science, these models appear frequently as a means to develop 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 [19], communication networks [18], and biological networks [25].

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. If P is the transition probability matrix of a Markov chain with stationary distribution π, the mixing time

$t_{\text{mix}}(\varepsilon)$ is defined as

$$t_{\text{mix}}(\varepsilon) = \min\left\{t : \max_{x \in V} \|P^t(x,\cdot) - \pi\|_{\text{TV}} \leq \varepsilon\right\},$$

where $\|\cdot\|_{\text{TV}}$ denotes the total variation distance. In the setting where P is reversible and ergodic, the mixing time can be bounded in terms of the spectral gap of P as follows. If $\lambda_1 = 1 > \lambda_2 \geq \cdots \geq \lambda_n \geq -1$ are the eigenvalues of P, then the spectral gap is defined as $\gamma = 1 - \lambda_2$. It is known that

$$t_{\text{mix}}(\varepsilon) \leq \frac{1}{\gamma} \log\left(\frac{1}{\varepsilon \pi_{\min}}\right),$$

where $\pi_{\min} = \min_{x \in V} \pi(x)$. Thus, understanding the spectral properties of P is key to analyzing mixing times.

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 transition probability matrix eigenvalues, eigenvectors (in the simple edit case), and sharp mixing time estimates. Notably, with simple edits the stationary distribution matches an edge-independent random graph model (including Erdős–Rényi, power law, and stochastic block models), and the framework extends to compound edits such as the Moran forest and random intersection graph processes, with mixing governed by the largest nontrivial eigenvalue.

Theorem 3.1 (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 (i.e., 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 $\mathcal{P}$. 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.*

With an eye toward future work, we are currently using these techniques to develop and analyze new generative diffusion-based models for graphs by choosing appropriate edit operations and studying the resulting stochastic processes. Separately, we are also interested in adapting emerging techniques from the study of random walks on simplicial complexes (which have shown great promise in theoretical computer science, see [3, 20]) to analyze existing graph-based sampling algorithms from a new perspective.

4 Geometric graph theory: Connection Laplacians and curvature

Graphs can be endowed, both intrinsically and extrinsically, with geometric structures that capture combinatorial, spectral, and other properties of the underlying graph or data. My work in this area primarily focuses on connection Laplacians and discrete curvature notions on graphs. Below, I highlight a few questions that motivate me and other researchers working in this area.

- How can we extend classical notions of effective resistance and conductance to connection graphs, which incorporate extrinsic geometric information through parallel transport operators? What are the properties of these extended notions, and how can they be applied to downstream tasks such as sparsification and clustering?
- It has long been known that Ollivier-Ricci curvature is connected to the mixing times of Markov chains. Are similar results true for other discrete curvature notions, such as Forman-Ricci curvature or Steinerberger curvature? Can we develop new techniques to analyze the relationship between curvature and mixing times in these settings?
- Can we develop a comprehensive framework for understanding the relationships between different discrete curvature notions on graphs? How do these notions compare in terms of their properties, applications, and computational complexity?

In the subsections to follow, I discuss some of my contributions to these areas in greater detail.

4.1 Connection Laplacians

One way to incorporate extrinsic geometric information into graph-based methods is through *connection Laplacians*, which are Laplacians that encode graphs equipped with additional structure that model parallel transport information. 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 [34], manifold learning [35], and shape analysis [32].

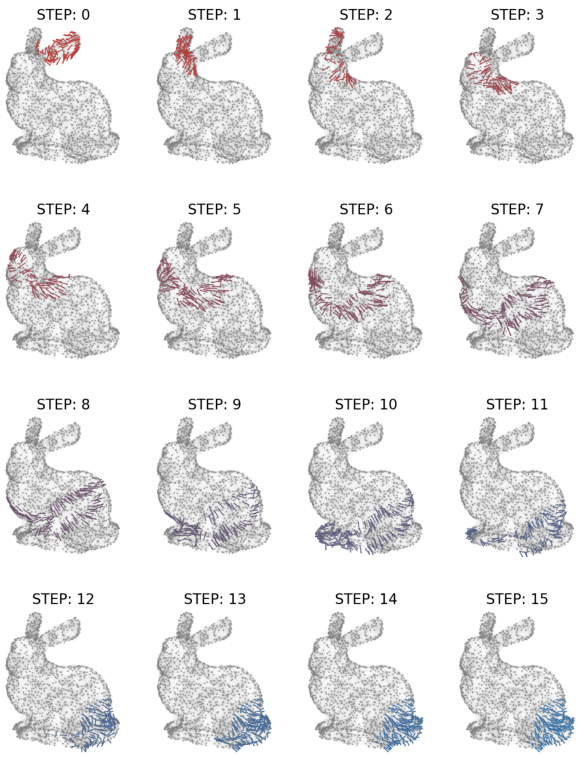

Figure 3: By investigating a notion of *minimum cost parallel transport* between vector fields defined on a geometric graph [29], we introduced an algorithm for vector field interpolation on graphs. Here, we illustrate such an interpolation between vector fields defined on the Stanford bunny [36].

In [11], we formulated matrix-valued Dirichlet problems on connection graphs, defining connection voltage functions and conductance matrices that extend classical notions to the connection setting. Specifically, we defined the *connection conductance matrix* C_σ between node pairs based on the solution to a matrix-valued boundary value problem on G using the connection Laplacian L_σ. We showed that C_σ captures the effective conductance between nodes while respecting 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 [29], 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, and revealed phenomena absent in classical graphs. For example, equal mass marginals can be infeasible under nontrivial one-dimensional (signed) connections, while unequal mass marginals can be feasible, highlighting how the connection reorients flux constraints. 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.

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 [24], Lin–Lu–Yau curvature [23], Forman-Ricci curvature [17], and more recently, Steinerberger curvature [37] 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 [27], I derived explicit formulas for Wasserstein distances and curvature measures on trees, providing insights into their geometric properties. Specifically, I provided closed-form expressions for the Ollivier-Ricci and Lin–Lu–Yau curvatures on trees in terms of the tree structure and edge weights, revealing relationships between these curvature notions and classical graph invariants such as degree sequences and path lengths.

Finally, in [30] (in preparation, with F. Southerland and E. Surya), we investigated a special class of graphs arising from a notion of curvature of Steinerberger [37]. Specifically, we studied *distance exceptional graphs*, which are graphs where the equation $Dx = \mathbf{1}$ does not have a solution, where D is the distance matrix of the graph and $\mathbf{1}$ is the all-ones vector. We introduced a graph invariant called the *curvature index*, which measures the extent to which a graph deviates from being distance exceptional. We developed a detailed calculus for this index under graph operations such as Cartesian products and graph joins, and used these to compute the curvature index for various families of graphs. In doing so, we generalized a number of known theorems in this field and proved an embedding theorem: any graph G can be embedded as an induced subgraph into a distance exceptional graph H, and in many cases, this embedding is an isometry.

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, https://doi.org/10.1137/0403039.

[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] N. Anari, K. Liu, S. Oveis Gharan, and C. Vinzant, *Log-concave polynomials II: High-dimensional walks and an FPRAS for counting bases of a matroid*, Ann. of Math. (2), 199 (2024), pp. 259–299, https://doi.org/10.4007/annals.2024.199.1.4.

[4] M. Belkin and P. Niyogi, *Towards a theoretical foundation for Laplacian-based manifold methods*, J. Comput. System Sci., 74 (2008), pp. 1289–1308, https://doi.org/10.1016/j.jcss.2007.08.006.

[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, https://doi.org/10.4230/lipics.opodis.2023.13.

[8] F. Chung and M. Radcliffe, *On the spectra of general random graphs*, Electron. J. Combin., 18 (2011), https://doi.org/10.37236/702.

[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, https://doi.org/10.1016/j.dam.2024.01.012.

[15] M. Dyer, L. A. Goldberg, and M. Jerrum, *Counting and sampling H-colourings*, Inform. and Comput., 189 (2004), pp. 1–16, https://doi.org/10.1016/j.ic.2003.09.001.

[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, https://doi.org/10.1007/s00454-002-0743-x.

[18] 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.

[19] S. Hanneke, W. Fu, and E. P. Xing, *Discrete temporal models of social networks*, Electron. J. Stat., 4 (2010), pp. 585–605, https://doi.org/10.1214/09-EJS548.

[20] T. Kaufman and I. Oppenheim, *High order random walks: beyond spectral gap*, Combinatorica, 40 (2020), pp. 245–281, https://doi.org/10.1007/s00493-019-3847-0.

[21] 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.

[22] J. Lei and A. Rinaldo, *Consistency of spectral clustering in stochastic block models*, Ann. Statist., 43 (2015), pp. 215–237, https://doi.org/10.1214/14-AOS1274.

[23] Y. Lin, L. Lu, and S.-T. Yau, *Ricci curvature of graphs*, Tohoku Math. J. (2), 63 (2011), pp. 605–627, https://doi.org/10.2748/tmj/1325886283.

[24] Y. Ollivier, *Ricci curvature of Markov chains on metric spaces*, J. Funct. Anal., 256 (2009), pp. 810–864, https://doi.org/10.1016/j.jfa.2008.11.001.

[25] 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.

[26] 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.

[27] 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.*

[28] 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.

[29] 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.

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

[31] 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.

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

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

[34] A. Singer, *Angular synchronization by eigenvectors and semidefinite programming*, Appl. Comput. Harmon. Anal., 30 (2011), https://doi.org/10.1016/j.acha.2010.02.001.

[35] A. Singer and H.-T. Wu, *Vector diffusion maps and the connection Laplacian*, Comm. Pure Appl. Math., 65 (2012), pp. 1067–1144, https://doi.org/10.1002/cpa.21395.

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

[37] S. Steinerberger, *Curvature on graphs via equilibrium measures*, J. Graph Theory, 103 (2023), pp. 415–436, https://doi.org/10.1002/jgt.22925.

[38] G. Tinhofer, E. Mayr, and H. Noltemeier, eds., *Computational graph theory*, vol. 7 of Computing Supplementum, Springer-Verlag, Vienna, 1990, https://doi.org/10.1007/978-3-7091-9076-0.

[39] 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.