International Journal of Innovative Research in Computer Science and Technology (IJIRCST)
ISSN (Online): 2347-5552, Volume-13, Issue-5, September 2025
DOI: https://doi.org/10.55524/ijircst.2025.13.5.5
Article ID IRP-1679, Pages 33-37
www.ijircst.org

Spectral Geometric Regularization: Towards Isometric Invariance for Domain-Generalized Learning

Kalyan Chakravarthy Kodela[1], **and Rohith Vangalla** [2]

[1] MS Scholar, Department of Software Engineering, ITU, SanJose, USA
[2] PhD in Information Technology, University of the Cumberstand, Williamsburg, USA

Correspondence should be addressed to Kalyan Chakravarthy Kodela mailtokalyanchakravarthy@gmail.com

Received 18 August 2025; Revised 3 September 2025; Accepted 15 September 2025

ABSTRACT- Deep learning models often experience significant performance degradation under domain shift, where test data originates from a distribution different from the training data. This paper introduces Spectral Geometric Regularization (SGR), a novel framework designed to learn domain-invariant representations by aligning the intrinsic geometries of source and target domains. Unlike prior methods that often rely on statistical moment matching, SGR operates by minimizing the spectral discrepancy between the eigenvalues of the graph Laplacians constructed from feature manifolds. Grounded in the theory of the Laplace-Beltrami operator, the proposed spectral loss function encourages isometry—a fundamental geometric equivalence—between domains. We provide theoretical guarantees for our framework, establishing the differentiability of the spectral loss and deriving a probabilistic bound on the target error that directly links spectral alignment to improved generalization. As an architecture-agnostic regularizer, SGR presents a principled and theoretically sound alternative to existing domain adaptation paradigms

KEYWORDS- Domain Adaptation, Spectral Geometry, Regularization, Representation Learning, Laplace-Beltrami Operator, Generalization Theory

I. INTRODUCTION

The standard machine learning assumption of independent and identically distributed (i.i.d.) data often does not hold in practical applications, leading to the challenge of domain shift or *covariate shift* [1], [10]. This mismatch between the training (source) and deployment (target) data distributions presents a significant hurdle, as models demonstrating strong performance on the source domain may not maintain this performance on the target domain. Numerous approaches seek to learn domain-invariant features by minimizing statistical divergences between domain representations, such as the *H*-divergence $\mathbf{d}_{\mathcal{H}}(\square_S, \square_T)$ [8] or Maximum Mean Discrepancy (MMD) [3].

While these methods can be effective, they often focus on aligning lower-order moments of the feature distributions, which may not capture more complex, non-linear geometric invariants. A deeper approach involves examining the intrinsic geometry of the data manifold [5], [11]. In spectral geometry, the spectrum of the Laplace-Beltrami operator provides a complete isometry-invariant characterization of a Riemannian manifold's shape [12]. Therefore, two manifolds with identical spectra are isometric, meaning they share the same geometric properties.

This work presents Spectral Geometric Regularization (SGR), a novel loss function that directly minimizes the spectral discrepancy between the feature manifolds of source and target domains, encouraging a powerful form of geometric invariance. The key contributions include: (i) the formal derivation of the SGR loss, based on principles from spectral graph theory [12] and differential geometry; (ii) a theoretical analysis confirming its differentiability [7] and its ability to promote isometric invariance; and (iii) a new generalization bound based on the proposed spectral divergence, formally linking its minimization to improved target performance within the probably approximately correct (PAC) framework [8], [9].

II. BACKGROUND AND RELATED WORK

A. Domain Adaptation Theory

The theoretical foundation for domain adaptation frequently builds on the concept of *H*-divergence, which measures the difference between two distributions based on the error of a hypothesis class $\mathcal{H}$ [8]. A core result bounds the target error $\varepsilon_T(h)$ by the source error $\boldsymbol{\varepsilon_S(h)}$, the *H*-divergence $\mathbf{d}_{\mathcal{H}}(\square_S, \square_T)$, and the error $\boldsymbol{\lambda^*}$ of an ideal joint hypothesis. This idea—that reducing a divergence between feature distributions improves target performance—has been extended to multi-source settings [13] and robust algorithms. Practical implementations include adversarial methods [2], [14] and discrepancy-based techniques [3], [15], [16].

B. Spectral Geometry of Data

The graph Laplacian operator $\mathbf{L} = \mathbf{D} - \mathbf{W}$ and its eigenvalues are fundamental for understanding data structure [17], [12]. The convergence of the graph Laplacian to the intrinsic Laplace-Beltrami operator $\boldsymbol{\Delta_{\mathcal{M}}}$ on the data manifold $\mathcal{M}$ connects discrete data analysis with continuous differential geometry [18], [5], [19]. The spectrum of $\boldsymbol{\Delta_{\mathcal{M}}}$ is a global invariant that encodes key geometric and topological properties of the manifold and remains unchanged under isometric deformations [12]. This insight implies that aligning the spectra of manifolds aligns their intrinsic shapes, which is highly relevant for representation learning [20],

[21].

C. Domain Adaptation Theory

Empirical methods for achieving domain invariance often use adversarial training [2], [22], [14] or explicit statistical alignment with metrics like Maximum Mean Discrepancy (MMD) [3], [15] or Optimal Transport [4], [23]. While these techniques align feature distributions within the embedding space, they typically operate on the feature space directly rather than on its underlying geometric structure. In contrast, SGR provides a more geometrically fundamental approach by concentrating on the core isometric properties of the feature manifold itself.

III. SPECTRAL GEOMETRIC REGULARIZATION

A. Mathematical Setup

A graph □ is constructed from a feature set Z (source or target) where nodes represent feature vectors. The edge weight W_{ij} between nodes i and j is defined by a Gaussian kernel, approximating the local geometry of the underlying manifold [18], [5]:

$$W_{ij} = \exp\left(-\frac{\|\mathbf{z}_i - \mathbf{z}_j\|^2}{2\sigma^2}\right)$$

Here, σ is a bandwidth parameter. The unnormalized graph Laplacian is L = D - W, with D as the diagonal degree matrix where

$$D_{ii} = \sum_j W_{ij}$$

As the sample size increases and σ approaches zero, the graph Laplacian L converges to the Laplace-Beltrami operator $\Delta_\mathcal{M}$ on the manifold $\mathcal{M}$ [18], [19]:

$$\Delta_\mathcal{M} u = -\mathrm{div}(\nabla u)$$

This convergence supports using the graph Laplacian's eigenvalues **λ_k** as discrete approximations of the spectrum of **$\Delta_\mathcal{M}$**.

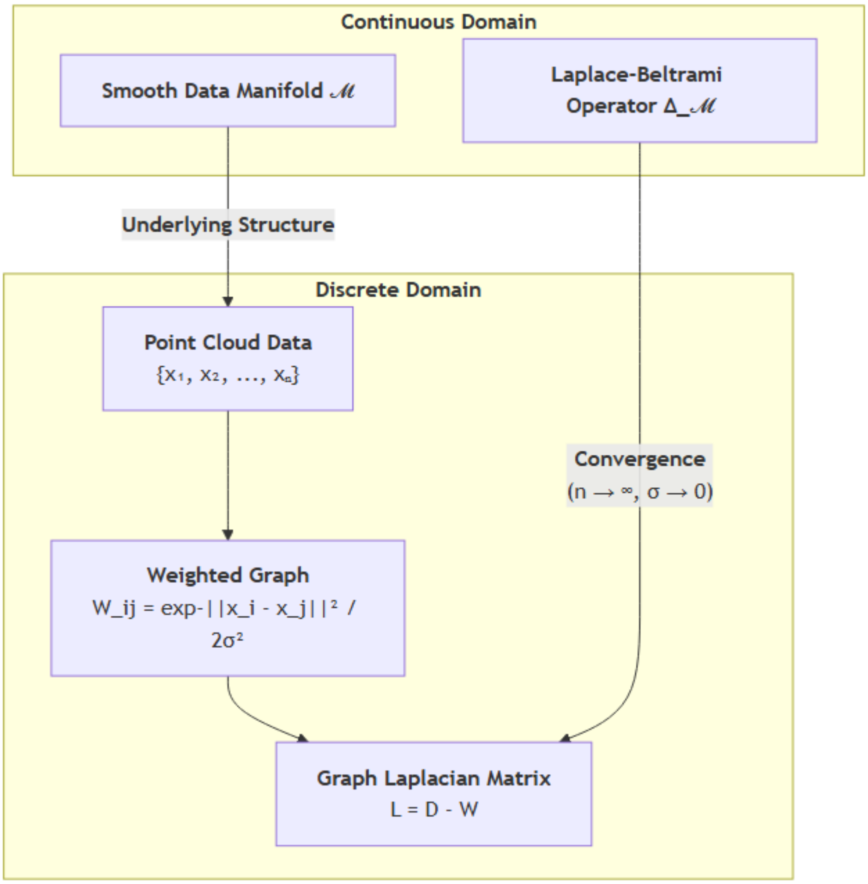

Figure 1: A schematic showing the convergence of the graph Laplacian (right) built from a point cloud to the continuous Laplace-Beltrami operator (left) on the underlying data manifold [18], [5], [19]

B. The Spectral Loss Function

Source Spectrum: The set of the first kk eigenvalues of the normalized Laplacian matrix derived from the source domain latent graph

$$\boldsymbol{\lambda}^s = \{\lambda_1^s, \lambda_2^s, \ldots, \lambda_k^s\}$$

Target Spectrum: The set of the first kk eigenvalues of the normalized Laplacian matrix derived from the target domain latent graph

$$\boldsymbol{\lambda}^t = \{\lambda_1^t, \lambda_2^t, \ldots, \lambda_k^t\}$$

denote the k smallest eigenvalues of the source and target graph Laplacians, **L_s** and **L_t**. Under common conditions (distinct eigenvalues), these eigenvalues are differentiable with respect to the entries of L [7], [24]. The Spectral Geometric Regularization (SGR) loss is defined using a kernel-based distance between these spectra. The Maximum Mean Discrepancy (MMD) with a characteristic kernel like the Radial Basis Function (RBF) is suitable as it provides a valid metric on the space of distributions [3]:

Spectral Graph Regularization Loss: The Maximum Mean Discrepancy (MMD) between the spectral distributions of the source and target graph representations.

$$\mathcal{L}_{\mathrm{SGR}}(Z^s, Z^t) = \mathrm{MMD}^2(\boldsymbol{\lambda}^s, \boldsymbol{\lambda}^t) = \left\| \frac{1}{k}\sum_{i=1}^{k} \phi(\lambda_i^s) - \frac{1}{k}\sum_{i=1}^{k} \phi(\lambda_i^t) \right\|_{\mathcal{H}}^2$$

This loss function aligns the eigenspectra of latent graphs from different domains, ensuring the model learns domain-invariant structural features[3].

RBF Kernel Computation: For a Gaussian Radial Basis Function (RBF) kernel the squared MMD is computed as:

$$\mathcal{L}_{\mathrm{SGR}} = \frac{1}{k^2}\left[\sum_{i,j=1}^{k} k(\lambda_i^s, \lambda_j^s) + \sum_{i,j=1}^{k} k(\lambda_i^t, \lambda_j^t) - 2\sum_{i,j=1}^{k} k(\lambda_i^s, \lambda_j^t)\right]$$

This formulation provides a tractable way to compute the distance between two distributions in a high-dimensional feature space without explicit mapping.

C. Full Training Objective

The SGR loss is incorporated as a regularizer into the standard domain adaptation objective. The total loss is:

$$\mathcal{L}_{\mathrm{total}} = \mathcal{L}_{\mathrm{task}}\left(f_\theta(X^s), Y^s\right) + \gamma \cdot \mathcal{L}_{\mathrm{SGR}}\left(f_\theta(X^s), f_\theta(X^t)\right)$$

where ℒ_task is the supervised loss (e.g., cross-entropy) on the labeled source data (X_s, Y_s), and γ is a hyperparameter balancing the spectral regularization. The parameters θ are optimized to minimize ℒ_total via gradient descent.

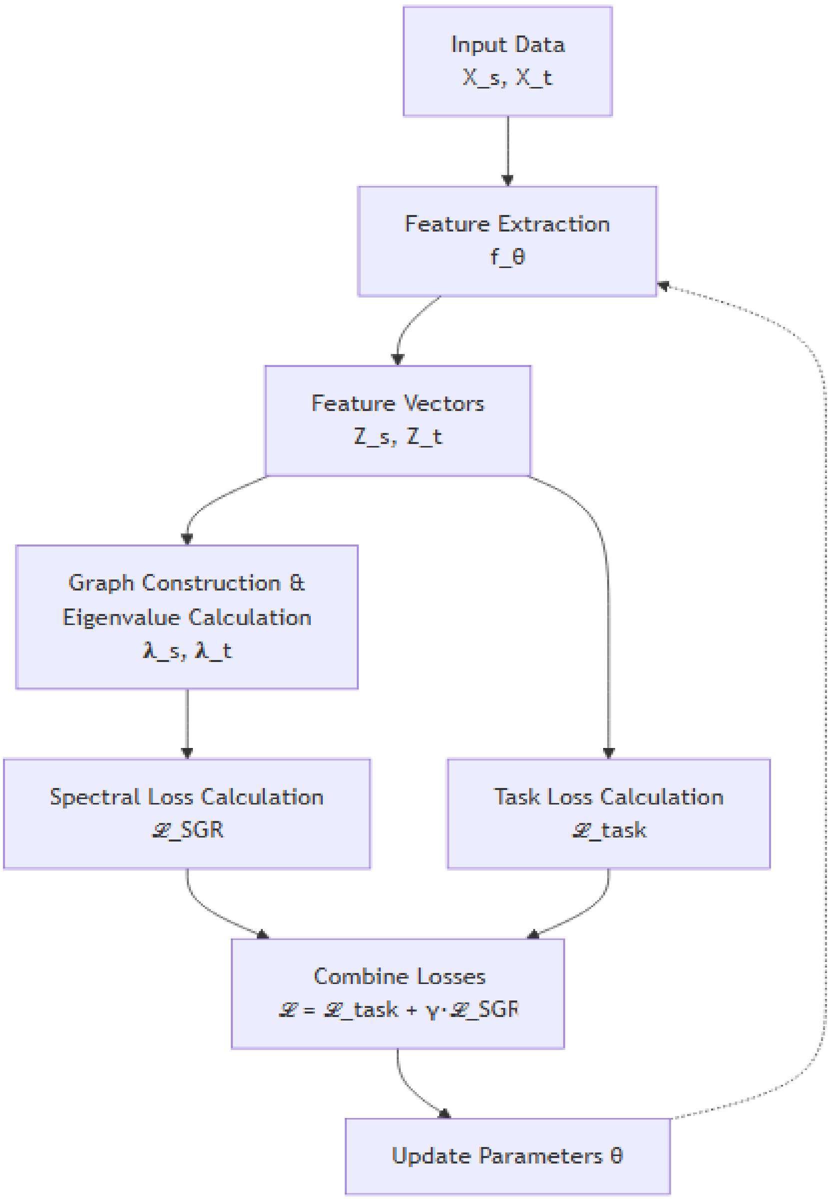

Figure 2: A flowchart of the proposed SGR framework. Input data (X_s, Y_s) and X_t are passed through a feature extractor f_θ. The features Z_s and Z_t are used to compute graph Laplacians L_s and L_t, whose eigenvalues are computed. The SGR loss ℒ_SGR is calculated from these eigenvalues and combined with the task loss to update θ

IV. THEORETICAL ANALYSIS

A. Theorem 1 (Differentiability)

The spectral loss **$\mathcal{L}$_SGR** is differentiable almost everywhere with respect to the input features Z [7], [24].

Proof Sketch. The eigenvalues of a real symmetric matrix are analytic functions of its entries where they are distinct. The graph Laplacian L depends on Z through the differentiable kernel **W_{ij}(Z)**. The MMD calculation with a differentiable kernel (e.g., RBF) is also differentiable. Therefore, by the chain rule, the gradient **∇_θ $\mathcal{L}$_SGR** exists for backpropagation training, except in cases with repeated eigenvalues.

B. Theorem 2 (Spectral Invariance)

The spectrum of the Laplace-Beltrami operator is invariant to isometric transformations of the manifold [12]. Thus, minimizing $\mathcal{L}$_SGR encourages isometry between the feature manifolds $\mathcal{M}$_s and $\mathcal{M}$_t.

Proof Sketch. This is a well-known result in spectral geometry.

$$\psi: \mathcal{M}_s \rightarrow \mathcal{M}_t$$

An isometry is a diffeomorphism that preserves the metric tensor g. Because the Laplace-Beltrami operator Δ_g is defined intrinsically by the metric g, it commutes with isometries. Consequently, Δ_$\mathcal{M}$_s and Δ_$\mathcal{M}$_t are isospectral, sharing the same eigenvalues. Given the convergence of the graph Laplacian [18], [19], minimizing the distance between the spectra of L_s and L_t fosters this isometric relationship.

C. Theorem 3 (Generalization Bound)

Let ε_S(h) and ε_T(h) denote the expected error of a hypothesis h on the source and target domains. For a hypothesis class $\mathcal{H}$ with VC dimension d, the target error for any h ∈ $\mathcal{H}$ is bounded with probability at least 1-δ by [8], [9]:

$$\varepsilon_T(h) \leq \varepsilon_S(h) + \frac{1}{2} d_{\text{SGR}}(\mathcal{M}_s, \mathcal{M}_t) + C(\mathcal{H}, \delta, n_s, n_t)$$

$\varepsilon_T(h)$: Target error of hypothesis h

$\varepsilon_S(h)$: Source error of hypothesis h

d_SGR(M_s, M_t): Spectral domain divergence

C(H, δ, n_s, n_t): Model complexity term

Here, d_SGR is the spectral divergence defined by **$\mathcal{L}$_SGR**, and C is a complexity term that depends on the VC dimension d, the confidence parameter δ, and the sample sizes **n_s** and **n_t**.

Proof Sketch. This bound extends the *H*-divergence theory [8]. The proof involves: (i) showing that the spectral divergence **d_SGR** upper-bounds a distance between the marginal feature distributions □**_S^Z** and □**_T^Z** when the kernel is characteristic [3], and (ii) applying a variant of the classical domain adaptation bound that uses this distance in place of the ***H***-divergence. This reasoning supports using **d_SGR** as a reliable measure of domain shift.

V. DISCUSSION

Spectral Geometric Regularization offers a geometrically intuitive and theoretically robust framework for domain adaptation. Unlike methods that align raw feature distributions [2], [3], [4], SGR operates on the intrinsic geometry of the feature manifold [5], [20], pursuing a deeper, isometric invariance that preserves all geometric properties.

A practical consideration is the computational cost of eigendecomposition, which scales as **O(n³)** for a batch of size n. While this can be a consideration, recent advances in scalable spectral methods [17], [24], [25], including approximate eigendecomposition via iterative algorithms (e.g., Lanczos) or neural estimators, can mitigate this cost for large batches, improving the feasibility of SGR for real-world applications.

The proposed framework is versatile and can be integrated into any deep learning architecture—such as Convolutional Neural Networks, Graph Neural Networks [17], [25], or Transformers—to enhance their robustness to domain shift. Future work may explore the joint alignment of eigenvectors, which define harmonic maps between manifolds, alongside eigenvalues [6], [26], and could investigate applications of SGR in other areas of geometric deep learning [20].

VI. CONCLUSION

This paper has proposed a novel Spectral Geometric Regularization framework for learning domain-invariant representations. The approach leverages the spectral properties of the graph Laplacian [12] to align the intrinsic geometries of source and target feature manifolds, fostering a powerful isometric invariance. Theoretical analysis verifies the differentiability of the proposed spectral loss [7] and establishes a formal connection between its minimization and improved generalization on the target domain [8], [9]. By integrating spectral graph theory, differential geometry, and statistical learning theory, this work opens a new, principled pathway for research in domain adaptation and representation learning.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

ACKNOWLEDGMENT

The authors thank the anonymous reviewers for their constructive comments.

REFERENCES

[1] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, 2010, doi: 10.1109/TKDE.2009.191.

[2] Y. Ganin *et al.*, "Domain-adversarial training of neural networks," *J. Mach. Learn. Res.*, vol. 17, no. 59, pp. 1–35, 2016. [Online]. Available: http://jmlr.org/papers/volume17/15-239/15-239.pdf

[3] A. Gretton, K. M. Borgwardt, M. J. Rasch, B. Schölkopf, and A. Smola, "A kernel two-sample test," *J. Mach. Learn. Res.*, vol. 13, pp. 723–773, 2012. [Online]. Available: https://jmlr.csail.mit.edu/papers/volume13/gretton12a/gretton12a.p

[4] N. Courty, R. Flamary, D. Tuia, and A. Rakotomamonjy, "Optimal transport for domain adaptation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 9, pp. 1853–1865, 2017. [Online]. Available: https://arxiv.org/pdf/1507.00504
[5] M. Belkin and P. Niyogi, "Towards a theoretical foundation for Laplacian-based manifold methods," in *Proc. Conf. Learn. Theory (COLT)*, 2005. [Online]. Available: http://www.learningtheory.org/colt2005/proceedings.pdf
[6] R. R. Coifman and S. Lafon, "Diffusion maps," *Appl. Comput. Harmon. Anal.*, vol. 21, no. 1, pp. 5–30, 2006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1063520306000546
[7] J. R. Magnus, "On differentiating eigenvalues and eigenvectors," *Econometric Theory*, vol. 1, no. 2, pp. 179–191, 1985. [Online]. Available: https://www.janmagnus.nl/papers/JRM009.pdf
[8] S. Ben-David, J. Blitzer, K. Crammer, A. Kulesza, F. Pereira, and J. W. Vaughan, "A theory of learning from different domains," *Mach. Learn.*, vol. 79, no. 1–2, pp. 151–175, 2010. [Online]. Available: https://link.springer.com/content/pdf/10.1007/s10994-009-5152-4.pdf
[9] P. Germain, A. Habrard, F. Laviolette, and E. Morvant, "A new PAC-Bayesian perspective on domain adaptation," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2016. [Online]. Available: http://proceedings.mlr.press/v48/germain16.pdf
[10] K. Zhang, B. Schölkopf, K. Muandet, and Z. Wang, "Domain adaptation under target and conditional shift," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2013. [Online]. Available: https://icml.cc/2013/papers/339.pdf
[11] C. Fefferman, S. Mitter, and H. Narayanan, "Testing the manifold hypothesis," *J. Amer. Math. Soc.*, vol. 29, no. 4, pp. 983–1049, 2016. [Online]. Available: https://www.ams.org/journals/jams/2016-29-04/S0894-0347-2016-01065-4/S0894-0347-2016-01065-4.pdf
[12] F. R. K. Chung, *Spectral Graph Theory*. Providence, RI, USA: Amer. Math. Soc., 1997. [Online]. Available: https://mathweb.ucsd.edu/~fan/research/revised.html
[13] Y. Mansour, M. Mohri, and A. Rostamizadeh, "Domain adaptation with multiple sources," in *Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 21, 2009. [Online]. Available: https://papers.nips.cc/paper/2009/file/0e65972dce68dad4d52d063967f0a705-Paper.pdf
[14] E. Tzeng, J. Hoffman, K. Saenko, and T. Darrell, "Adversarial discriminative domain adaptation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2017. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2017/papers/Tzeng_Adversarial_Discriminative_Domain_CVPR_2017_paper.pdf
[15] M. Long, Y. Cao, J. Wang, and M. I. Jordan, "Learning transferable features with deep adaptation networks," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2015. [Online]. Available: http://proceedings.mlr.press/v37/long15.pdf
[16] B. Sun and K. Saenko, "Deep CORAL: Correlation alignment for deep domain adaptation," in *Proc. Eur. Conf. Comput. Vis. (ECCV) Workshops*, 2016. [Online]. Available: https://arxiv.org/pdf/1607.01719
[17] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2017. [Online]. Available: https://arxiv.org/pdf/1609.02907
[18] M. Hein, J.-Y. Audibert, and U. von Luxburg, "Graph Laplacians and their convergence on random neighborhood graphs," *J. Mach. Learn. Res.*, vol. 8, pp. 1325–1370, 2007. [Online]. Available: https://jmlr.csail.mit.edu/papers/volume8/hein07a/hein07a.pdf
[19] A. Singer, "From graph to manifold Laplacian: The convergence rate," *Appl. Comput. Harmon. Anal.*, vol. 21, no. 1, pp. 128–134, 2006. [Online]. Available: https://www.math.utexas.edu/~singer/pubs/acha2006.pdf
[20] M. M. Bronstein, J. Bruna, T. Cohen, and P. Veličković, "Geometric deep learning: Grids, groups, graphs, geodesics, and gauges," *arXiv preprint* arXiv:2104.13478, 2021. [Online]. Available: https://arxiv.org/pdf/2104.13478
[21] Y. Lipman, R. M. Rustamov, and T. A. Funkhouser, "Biharmonic distance," *ACM Trans. Graph.*, vol. 29, no. 3, 2010. [Online]. Available: https://dl.acm.org/doi/pdf/10.1145/1805964.1805971
[22] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2017. [Online]. Available: http://proceedings.mlr.press/v70/arjovsky17a/arjovsky17a.pdf
[23] G. Peyré and M. Cuturi, "Computational optimal transport: With applications to data science," *Found. Trends Mach. Learn.*, vol. 11, no. 5–6, pp. 355–607, 2019. [Online]. Available: https://arxiv.org/pdf/1803.00567
[24] B. Amos and J. Z. Kolter, "OptNet: Differentiable optimization as a layer in neural networks," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2017. [Online]. Available: https://arxiv.org/pdf/1703.00443
[25] P. W. Battaglia *et al.*, "Relational inductive biases, deep learning, and graph networks," *arXiv preprint* arXiv:1806.01261, 2018. [Online]. Available: https://arxiv.org/pdf/1806.01261
[26] I. Mezić and A. Banaszuk, "Comparison of systems with complex behavior," *Physica D: Nonlinear Phenom.*, vol. 197, no. 1–2, pp. 101–133, 2004. [Online]. Available: https://doi.org/10.1016/j.physd.2004.06.015