Preprint – IJCAI 2025: This is the accepted version made available for conference attendees.
Do not cite. The final version will appear in the IJCAI 2025 proceedings.

`Trace`: Structural Riemannian Bridge Matching for Transferable Source Localization in Information Propagation

Li Sun[1*], **Suyang Zhou**[1], **Bowen Fang**[2], **Hechuan Zhang**[1], **Junda Ye**[3], **Yutong Ye**[4], **Philip S. Yu**[5]

[1]North China Electric Power University, Beijing 102206, China
[2]Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China
[3]Beijing University of Posts and Telecommunications, Beijing 100876, China
[4]East China Normal University, Shanghai 200241, China
[5]Department of Computer Science, University of Illinois at Chicago, IL, USA
ccesunli@ncepu.edu.cn, zhousuyang154@gmail.com, fbw24@mails.tsinghua.edu.cn, psyu@uic.edu

Abstract

Source localization, the inverse problem of information diffusion, shows fundamental importance for understanding social dynamics. While achieving notable progress, existing solutions are typically exposed to the risk of error accumulation, and require a large number of observations for effective inference. However, it is often impractical to obtain quantities of observations in real scenarios, highlighting the need for a transferable model with broad applicability. Recently, Riemannian geometry has demonstrated its effectiveness in information diffusion and offers guidance in knowledge transfer, but has yet to be explored in source localization. In light of the issues above, we propose to study transferable source localization from a fresh geometric perspective, and present a novel approach (**`Trace`**) on the Riemannian manifold. Concretely, we establish a structural Schrödinger bridge to directly model the map between source and final distributions, where a functional curvature, encapsulating the graph structure, is formulated to govern the Schrödinger bridge and facilitate domain adaptation. Furthermore, we design a simple yet effective learning algorithm for Riemannian Schrödinger bridges (geodesic bridge matching) in which we prove the optimal projection holds for Riemannian measures so that the expensive iterative procedure is avoided. Extensive experiments demonstrate the effectiveness and transferability of `Trace` on both synthetic and real datasets.

1 Introduction

Information diffusion is a classic problem, predicting the infected nodes in the future according to the propagation pattern. Source localization, as the inverse problem of information diffusion, seeks to identify the source nodes responsible for the final diffusion observations [Yan *et al.*, 2023]. It not only facilitates a deeper understanding of social dynamics but also enables the vital applications such as information propagation intervention and rumor source detection. In recent years, researchers have visited the advanced models including invertible graph neural network and variational autoencoder [Wang *et al.*, 2022; Ling *et al.*, 2022]. While achieving notable progress, source localization still presents several important issues as follows.

Error Accumulation. Existing methods typically follow a step-by-step backward procedure [Yan *et al.*, 2023; Huang *et al.*, 2023] or model the round trip with forward information diffusion [Ling *et al.*, 2022; Wang *et al.*, 2022]. They are vulnerable to error accumulation especially in the absence of intermediate supervision. Instead, it is preferable to directly model the source and final distributions. Recent generative models achieve notable success in modeling data distributions [Song *et al.*, 2021; Lipman *et al.*, 2023], but they primarily focus on generating distribution from a Gaussian. So far, the transformation between arbitrary distributions, accounting for graph structure, remains under-investigated.

Transferability. Recent efforts leverage deep learning on graphs, which usually requires a large quantity of observations for effective inference [Ling *et al.*, 2024; Dong *et al.*, 2019]. However, acquiring observations is often laborious or expensive, limiting their widespread use in real scenarios. It calls for a transferable model that leverages the knowledge learned from previous observations to enhance the localization in new graphs. We emphasize that, given the underlying social dynamics, graph structures significantly influence information propagation, making domain adaptation essential.

Geometry. Riemannian geometry not only demonstrates superior expressiveness compared to its Euclidean counterparts [Sun *et al.*, 2022b], but also introduces the concept of curvature to describe graph structure [Sun *et al.*, 2023b], shedding light on quantifying the difference among social networks and enabling domain adaptation. Also, we note that recent efforts report the effectiveness of Riemannian geometry in information diffusion [Sun *et al.*, 2024a]. However, its application to the inverse problem of source localization has not yet been explored. Nevertheless, calculating Riemannian divergence or heat kernel is computationally expensive [Sun

*Corresponding Author: Li Sun.

Preprint – IJCAI 2025: This is the accepted version made available for conference attendees.
Do not cite. The final version will appear in the IJCAI 2025 proceedings.

et al., 2024b], and training models on the manifold remains challenging.

In light of the aforementioned issues, we are motivated to propose the problem of *Transferable Source Localization*, enabling the knowledge transfer to new graphs, and to approach this problem from a fresh geometric perspective.

In this paper, we present a novel structural Riemannian bridge matching model (termed as **`Trace`**) for transferable source localization. The key innovation is that we leverage the unified construction of **Riemannian manifold** to explore the transferability, and formulate a **structural Schrödinger Bridge** on the manifold along with a new learning algorithm. Specifically, to address the first issue of error accumulation, we make the first attempt, to the best of our knowledge, to introduce the Schrödinger Bridge [Bortoli *et al.*, 2021] to source localization. It directly models the map between the source and final distributions via a Stochastic Differential Equation (SDE), and we establish a structural Schrödinger bridge with a Wiener process on the manifold. In particular, the proposed bridge is governed by the functional curvature, which is designed to encapsulate the structural information and is learnable with given structure, addressing the second issue of transferability. Note that, learning Schrödinger bridges on the manifold is challenging (the third issue). To fill this gap, we propose a simple yet effective algorithm, geodesic bridge matching. Concretely, we discover the optimal projection for Riemannian measures with theoretical guarantees, so that the Schrödinger bridge is learned through geodesic bridge matching, avoiding the calculation of Riemannian heat kernel. Finally, `Trace` integrates the advantages of Schrödinger bridge and Riemannian geometry to directly infer the sources. We emphasize that the Riemannian manifold endows our model with the transferability. With the underlying dynamics learned from previous observations, we fine-tune the functional curvature on the new graph to conduct knowledge transfer, taking the graph structure into account.

Contribution Highlights. Main contributions are four-fold:

- **Problem.** We consider the geometric effect on source localization and for the first time pose the transferable source localization problem, emphasizing its transferability for practical usage, to the best of our knowledge.
- **Riemannian Model.** We propose a novel approach, `Trace`, in which the structural Riemannian Schrödinger bridge is established for the initial-terminal distribution map, equipped with functional curvature to enhance knowledge transfer through domain adaptation.
- **Learning Algorithm.** We design a new learning algorithm for Riemannian Schrödinger bridges, called geodesic bridge matching, in which we prove the optimal projection holds for Riemannian measures, thereby avoiding the need for the expensive iterative procedure.
- **Extensive Experiments.** We evaluate the superiority of `Trace` on both synthetic and real datasets, analyze each proposed component with ablation study, and examine its transferability among different datasets.

2 Preliminaries and Notations

Manifold and Curvature. In Riemannian geometry, a graph structure is related to certain Riemannian manifold $(\mathcal{M}, \mathfrak{g})$, a smooth manifold $\mathcal{M}$ attached to a Riemannian metric $\mathfrak{g}$. Each point on the manifold $\boldsymbol{x} \in \mathcal{M}$ is associated with a tangent space $\mathcal{T}_{\boldsymbol{x}}\mathcal{M}$ where the metric $\mathfrak{g}$ is defined. Given two points and $\mathfrak{g}$, geodesic is the curve that connects them with minimal integral length. The exponential and logarithmic map are inherited from the Lie algebra for the projection between tangent space and the manifold. Riemannian measure is defined over the infinitesimal on the manifold. *The notion of curvature κ describes how a smooth manifold deviates from being flat and, accordingly, it offers a geometric perspective to quantify graph structures.* There exist three types of constant curvature spaces: negative curvature hyperbolic space, positive curvature hyperbolic space, and the flat Euclidean space, a special case of Riemannian geometry.

Source Localization and Generative Models. In this paper, we recast source localization as a generative task and the idea is that we leverage the distribution of final observations to generate the initial source distribution, while capturing the social dynamics in the generative model. Recently, score-based (diffusion-based) models formulate Stochastic Differential Equations (SDEs) to generate data distribution from a simple one such as Gaussian [Song *et al.*, 2021], while flow-based models use Ordinary Differential Equations (ODEs) to perform such probability pushforward [Lipman *et al.*, 2023]. Though both models achieve notable success, we highlight that they fall short in transforming arbitrary distributions, i.e., source and final distributions, or require auxiliary time information to complete source localization [Huang *et al.*, 2023]. In other words, a promising generative localizer remains largely under-explored.

Problem Statement of Transferable Source Localization. We consider a graph $G(\mathcal{V}, \mathcal{E})$ defined over node set $\mathcal{V}$ and edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. The adjacency matrix is denoted as $\boldsymbol{A}$ and graph Laplacian is given as $\boldsymbol{L} = \boldsymbol{I} - \boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{A}\boldsymbol{D}^{-\frac{1}{2}}$, where $\boldsymbol{D}$ is the diagonal degree matrix. For source localization, only the final snapshot is accessible and the status of each node is described in $\boldsymbol{y} \in \{0,1\}^{|\mathcal{V}|}$, where $\boldsymbol{y}_i = 1$ indicates that node i is infected, otherwise, $\boldsymbol{y}_i = 0$. The sources are collected in $\boldsymbol{s} \in \{0,1\}^{|\mathcal{V}|}$, and $\boldsymbol{s}_i = 1$ iff node i is one of the sources responsible for $\boldsymbol{y}$. Given the graph and the final snapshot, the problem of source localization aims to seek $\mathcal{F}(G, \boldsymbol{y})$ that is able to infer the sources $\boldsymbol{s}$. Different from previous works, in *Transferable Source Localization*, we are dedicated to design $\mathcal{F}$ that offers the transferability among different graphs.

3 Methodology: `Trace`

We present a novel structural Riemannian bridge matching model (termed as `Trace`) for transferable source localization. The key novelty lies in that we formulate the *Structural Riemannian Schrödinger Bridge* along with a *New Learning Algorithm*, while exploring the advantage of Riemannian manifold for *Transferability*. The overall architecture is sketched in Fig. 1, and we start with the Schrödinger bridge in the heart of our model.

Preprint – IJCAI 2025: This is the accepted version made available for conference attendees.
Do not cite. The final version will appear in the IJCAI 2025 proceedings.

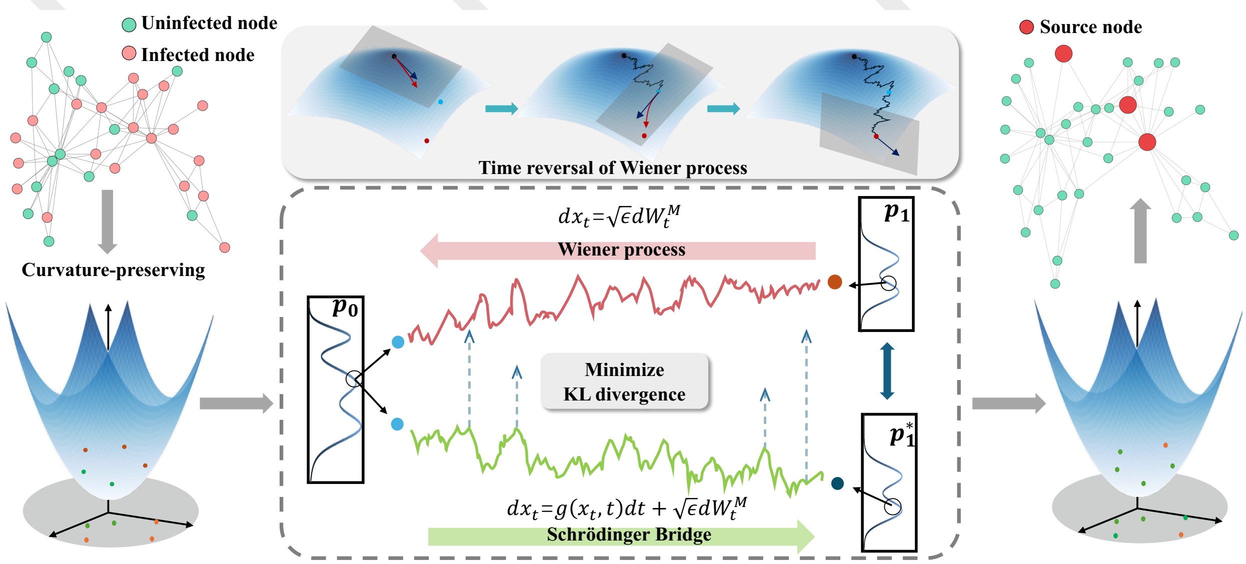


Figure 1: Overall architecture of structural Riemannian Schrödinger bridge and its training algorithm with the optimal projection.

3.1 Structural Riemannian Schrödinger Bridge

To the best of our knowledge, we for the first time introduce the Schrödinger bridge to the problem of source localization. The advantage is that we can directly model the distribution map between source and final states, thereby mitigating error accumulation. In a nutshell, we establish the Riemannian Schrödinger bridge that is equipped with a functional curvature depicting the structural information of the graph.

Schrödinger Bridge. It connects two arbitrary distributions with a Stochastic Differential Equation (SDE). To be specific, we consider a measure space $\mathcal{P}(C[0,1],\mathbb{R}^d)$, which collects the continuous and finite density p_t, where $t \in [0,1]$.[1] The density at point $\boldsymbol{x} \in \mathbb{R}^d$ of time t is denoted as $p_t(\boldsymbol{x})$, and d denotes the dimensionality. For a **process** T in measure space $\mathcal{P}$, the joint distribution at $t = 0, 1$ is $\pi^T \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$, and $T_{|\boldsymbol{x}_0,\boldsymbol{x}_1}$ denotes the distribution conditioned on T's value $\boldsymbol{x}_0, \boldsymbol{x}_1$ at $t = 0, 1$, respectively. We select a reference process $W \in \mathcal{P}(C[0,1],\mathbb{R}^d)$ which is typically set as the Wiener process of Brownian motion. In particular, it satisfies the SDE of $d\boldsymbol{x}_t = \sqrt{\epsilon}dW_t$, where $\epsilon > 0$ is the volatility. Given two arbitrary distributions $p_0, p_1 \in \mathcal{P}$, the (dynamic) Schrödinger bridge problem is formulated as the optimization,

$$T^\star = \arg\min\{KL(T\|W) : T_0 = p_0, T_1 = p_1\}. \quad (1)$$

The optimal process $T^\star$ is referred to as the Schrödinger bridge [Schrödinger, 1932; Föllmer, 1988]. In other words, Schrödinger bridge is the process that admits the distribution p_0, p_1 at $t = 0, 1$ and is closest to the reference process in term of Kullback-Leibler divergence.

Structure-aware Bridge in Riemannian manifold. Here, we formulate the optimal process $T^\star$ on the manifold with the Riemannian Wiener process. *We consider a smooth manifold $\mathcal{M}^{\kappa,d}$ whose curvature encapsulates the structural information of graph G.* Let $\mathcal{P}^{\mathcal{M}}(C[0,1],\mathcal{M}^{\kappa,d})$ denote the Riemannian measure space, and its Wiener process is written as $d\boldsymbol{x}_t = \sqrt{\epsilon}dW_t^{\mathcal{M}}$. The Riemannian Schrödinger bridge is to find the minimizer of Eq. 2 in $\mathcal{P}^{\mathcal{M}}$ pinned down at p_0 and p_1. Different from the Euclidean version, a process on the manifold is described as $d\boldsymbol{x}_t = f(\boldsymbol{x}_t, t)dt + \sqrt{\epsilon}dW_t^{\mathcal{M}}, \boldsymbol{x}_t \in \mathcal{M}^{\kappa,d}$, and its time reversal is derived as

$$d\boldsymbol{y}_t = \left(-f(\boldsymbol{y}_t, t) + \epsilon\nabla_{\boldsymbol{y}_t} \log p_{1-t}(\boldsymbol{y}_t)\right)dt + \sqrt{\epsilon}dW_t^{\mathcal{M}}, \quad (2)$$

where $(\boldsymbol{y}_t)_{t>0} = (\boldsymbol{x}_{1-t})_{t\in[0,1]}$ is the time reversal of $\boldsymbol{x}_t$. $\forall f$ over $\mathcal{M}^{\kappa,d}$, $\nabla_{\boldsymbol{x}_t} f \in \mathcal{T}_{\boldsymbol{x}_t}\mathcal{M}^{\kappa,d}$ is Riemannian gradient. $g(\boldsymbol{y}_t, t) = -f(\boldsymbol{y}_t, t) + \epsilon\nabla_{\boldsymbol{y}_t} \log p_{1-t}(\boldsymbol{y}_t)$ is known as drift.

In fact, the optimal $T^\star$ is unique and can be expressed as $T^\star = \int_{\mathbb{R}^d\times\mathbb{R}^d} W_{|\boldsymbol{x}_0,\boldsymbol{x}_1} d\pi^\star(\boldsymbol{x}_0, \boldsymbol{x}_1)$, where the joint marginal distribution $\pi^\star$ is the solution of the Entropic Optimal Transport (EOT) problem [Léonard, 2014] as follows,

$$\min_{\pi\in\Pi(p_0,p_1)} \int_{\mathcal{M}\times\mathcal{M}} \frac{\|\boldsymbol{x}_0-\boldsymbol{x}_1\|_{\mathfrak{g}}^2}{2}\pi(\boldsymbol{x}_0,\boldsymbol{x}_1)d\boldsymbol{x}_0 d\boldsymbol{x}_1 - \epsilon H(\pi). \quad (3)$$

$\Pi(p_0, p_1)$ is the set of transport plan over the joint $\mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ whose marginals are p_0 and p_1, and $\pi^\star$ is named as EOT plan. $d_{\mathfrak{g}}$ is the metric on the manifold, and $H(\pi)$ is the entropy over π. However, the EOT plan $\pi^\star$ is unavailable. *Hence, we are interested in learning the optimal drift $g^\star$ of Riemannian Schrödinger bridge $T^\star$.*

Functional Curvature. Distinguishing from the existing Schrödinger bridges [Jo and Hwang, 2024; Thornton *et al.*, 2022], the proposed Riemannian Schrödinger bridge $T^\star$ is governed by the functional curvature $\kappa_{\boldsymbol{\theta}}$, a key ingredient in our design. As the curvature summarizes the graph structure, we propose to parametrize $\kappa_{\boldsymbol{\theta}}$ with Laplacian $\boldsymbol{L}$. Denote the tuple of $\boldsymbol{L}$'s dominant eigenvalues as $\boldsymbol{l}$. The functional curvature is given as $\kappa_{\boldsymbol{\theta}} = \boldsymbol{l}^\top\boldsymbol{\theta}$. In **`Trace`**, we learn its parameter in the self-supervised fashion, and the supervision is given by the Parallelogram Law in Riemannian geometry [Gu *et al.*, 2019]. Concretely, given a geodesic triangle abc on the manifold, we have $\gamma_{abc} = d_{\mathfrak{g}}(a,m)^2 + \frac{d_{\mathfrak{g}}(b,c)^2}{4} - \frac{d_{\mathfrak{g}}(a,b)^2+d_{\mathfrak{g}}(a,c)^2}{2}$,

[1] In this paper, we utilize $t = 0$ to refer to the initial of the process, while $t = 1$ stands for its terminal.

Preprint – IJCAI 2025: This is the accepted version made available for conference attendees.
Do not cite. The final version will appear in the IJCAI 2025 proceedings.

where $d_{\mathfrak{g}}$ is the distance metric, and m is the midpoint of bc. Accordingly, the objective of curvature learning is derived as

$$\mathcal{L}_{\text{manifold}} = \|\kappa_{\boldsymbol{\theta}} - \mathbb{E}_{m \in \mathcal{V}}[\gamma_{ma|bc}]\|^2, \ \gamma_{ma|bc} = \frac{\gamma_{abc}}{d_{\mathfrak{g}}(a,m)}, \quad (4)$$

where a, b, c and m in the manifold, and their encodings are introduced in Sec. 3.2.

3.2 Graph Encoding with Functional Curvature

Having the bridge between the initial and terminal distributions at hand, we specify its two ends in this part. As in Fig. 1, the initial distribution is defined as the final observation while the sources are resided in the terminal. This allows us to utilize the final observation to generate the sources in line with the setting of source localization. We first elaborate on the model space of Riemannian geometry.

Unified Construction of Riemannian Manifold. Riemannian manifold offers an elegant unified construction for the constant curvature space of different curvatures. We introduce a model space $\mathcal{M}^{\kappa,d}$ to unify hyperbolic and hyperspherical spaces. Specifically, a $d-$dimensional manifold $\mathcal{M}^{\kappa,d}$ with curvature κ is embedded in $\mathbb{R}^{d+1}$ space,

$$\mathcal{M}^{\kappa,d} = \{\boldsymbol{x} = [x_t \ \boldsymbol{x}_s^\top]^\top | \langle \boldsymbol{x}, \boldsymbol{x} \rangle_\kappa = 1/\kappa, x_t > 0, \kappa \neq 0\}, \ (5)$$

where $x_t \in \mathbb{R}$ and $\boldsymbol{x}_s \in \mathbb{R}^d$ are the time and space dimension, respectively. The curvature-aware inner product is given as,

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_\kappa = \boldsymbol{x}\mathbf{R}\boldsymbol{y}, \quad \mathbf{R} = \operatorname{diag}(sgn(\kappa), 1, \cdots, 1), \quad (6)$$

where sgn is the sign function, and the diagonal $\mathbf{R}$ gives the Riemannian metric in $\mathbb{R}^{d+1}$ space. Note that, $\mathcal{M}^{\kappa,d}$ recovers the Lorentz model of hyperbolic space with $\kappa < 0$, and shifts to Spherical model of hyperspherical space when $\kappa > 0$.

Curvature-preserving Encoding. The samples are given by the proposed curvature-preserving encoding, where we adopt the operations without coupling logarithmic and exponential maps as in [Sun *et al.*, 2025a], and the advantage is investigated in Ablation Study. Specifically, dimension transformation is done by matrix-left-multiplication written as,

$$\text{Transform}_\kappa(\boldsymbol{W}, \boldsymbol{x}) = \begin{bmatrix} \sqrt{-sgn(\kappa)\|\boldsymbol{W}\boldsymbol{x}\|^2 + 1/|\kappa|} \\ \boldsymbol{W}\boldsymbol{x} \end{bmatrix}, \ (7)$$

parameterized by $\boldsymbol{W}$. $\|\cdot\|$ denotes the L2 norm. We formulate the graph convolution as the geometric midpoint on the manifold. Given a set of points $\{(\boldsymbol{x}_i, w_i)\}_{i \in \Omega}$ each attached a weight w_i, the graph convolution over Ω takes the form of

$$\text{Conv}_\kappa(\{(\boldsymbol{x}_i, w_i)\}_{i\in\Omega}) = \frac{1}{\sqrt{|\kappa|}} \sum_{i=1}^{|\Omega|} \frac{w_i}{|\|\sum_{j=1}^{|\Omega|} w_j \boldsymbol{x}_j\|_\kappa|} \boldsymbol{x}_i, \ (8)$$

where $|\|\boldsymbol{a}\|_\kappa| = \sqrt{-\langle \boldsymbol{a}, \boldsymbol{a} \rangle_\kappa}$ is the modulus of the imaginary. Note that, both operations preserve the curvature.

Theorem 1 (Curvature-preserving). *$\forall \boldsymbol{x} \in \mathcal{M}^{\kappa,d_1}$ and any $w \in \mathbb{R}$, $\boldsymbol{W} \in \mathbb{R}^{d_2 \times d_1}$, we have $\text{Conv}_\kappa(\{(\boldsymbol{x}_i, w_i)\}_{i\in\Omega}) \in \mathcal{M}^{\kappa,d_1}$ and $\text{Transform}_\kappa(\boldsymbol{W}, \boldsymbol{x}) \in \mathcal{M}^{\kappa,d_2}$ hold for either positive or negative value of κ.*

We perform dimension transformation and graph convolution to obtain the node encodings. With the curvature-preserving encoding, the structural information is delivered to the manifold where the proposed Schrödinger bridge is established.

3.3 Learning Schrödinger Bridge on Manifolds

Challenges. Popular learning algorithms, such as Iterative Proportional Fitting (IPF) [Bortoli *et al.*, 2021] and Iterative Markovian Fitting (IMF) [Shi *et al.*, 2023], adopt an iterative procedure, which *tends to result in the optimization error in each iteration* [Gushchin *et al.*, 2024]. In addition, *calculating the divergence or heat kernel on the manifold is rather expensive*, and the estimators are typically applied [Sun *et al.*, 2024b]. Therefore, learning Riemannian Schrödinger bridges largely remains open.

Theoretical Results on Riemannian Optimal Projection. To fill this gap, we design a new learning algorithm built upon Riemannian optimal projection. The core contribution is that we discover the optimal projection for Riemannian measures.

We start with IMF algorithm in Euclidean space, a widely used successor of the IPF. Concretely, IMF first introduces the reciprocal projection $\text{proj}_R(T)$ that gives the reciprocal process T_π which is equal to a mixture of Wiener processes with given coupling π. Second, the Markovian projection $\text{proj}_M(T)$ is defined by the SDE in [Shi *et al.*, 2023]. Given p_1 and p_0, IMF alternates between the projections iteratively, and the sequence of $\{T^{2n+1} = \text{proj}_M(T^{2n}), T^{2n+2} = \text{proj}_R(T^{2n+1})\}$ is shown to converge to $T^\star$. Note that, the Markovian projection is problematic, which requires to solve the optimization over $\mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ in every odd iteration.

Instead of the Markovian projection, we propose a novel *Riemannian optimal projection*, generalizing the Euclidean one [Gushchin *et al.*, 2024], and prove its optimality for Riemannian measures. To be specific, we **first** introduce the concept of Riemannian Schrödinger family $\mathcal{S}$, collecting the stochastic process of Schrödinger bridge S pinned down at p_{s0} and p_{s1} in the Riemannian measure space $\mathcal{P}^{\mathcal{M}}(C[0,1], \mathcal{M}^{\kappa,d})$. **Second**, for any given $\pi \in \Pi(p_0, p_1)$ and process T, the Riemannian reciprocal process is written as $T_\pi = \text{proj}_{R^{\mathcal{M}}}(T) = \int_{\mathcal{M}\times\mathcal{M}} W^{\mathcal{M}}_{|\boldsymbol{x}_0,\boldsymbol{x}_1} d\pi(\boldsymbol{x}_0, \boldsymbol{x}_1)$ with Riemannian Wiener process. **Third**, the proposed optimal projection is given as the solution to the following equation,

$$\text{proj}_{\mathcal{S}^{\mathcal{M}}}(T_\pi) = \arg_{S \in \mathcal{S}} \min KL(T_\pi || S). \quad (9)$$

The intuition behind this is to find the Schrödinger bridge S which is closest to the Riemannian reciprocal process T_π in terms of KL divergence.

Interestingly, we discover that the result of Riemannian optimal projection is the Riemannian Schrödinger bridge connecting $p_0, p_1 \in \mathcal{P}^{\mathcal{M}}$. We show the following theorem holds for Riemannian measures. (The proof is in Appendix B.)

Theorem 2 (Optimality). *Given the initial and terminal distributions $p_0, p_1 \in \mathcal{P}^{\mathcal{M}}$, a joint distribution $\pi \in \Pi(p_0, p_1)$, Riemannian reciprocal process T_π is given accordingly. Riemannian optimal projection of T_π leads to Riemannian Schrödinger Bridge T^* between the distributions p_0 and p_1,* $\text{proj}_{\mathcal{S}^{\mathcal{M}}}(T_\pi) = \arg_{S \in \mathcal{S}} \min KL(T_\pi || S) = T^*$.

In light of the optimization in Eq. 9, we construct the Riemannian Schrödinger family $\mathcal{S}$ with tractable KL divergence. With the adjusted potential v, we consider the joint distribution $\pi_v(\boldsymbol{x}_0, \boldsymbol{x}_1) = p_0(\boldsymbol{x}_0)\pi_v(\boldsymbol{x}_1|\boldsymbol{x}_0)$, where $\pi_v(\boldsymbol{x}_1|\boldsymbol{x}_0) =$

Preprint – IJCAI 2025: This is the accepted version made available for conference attendees.
Do not cite. The final version will appear in the IJCAI 2025 proceedings.

Algorithm 1 Geodesic Random Walk

Input: length N, initial value $\boldsymbol{x}$, curvature κ, drift f, volatility σ.
Output: samples $\{\boldsymbol{x}_i^\gamma\}_{i=0}^N$

Set step size $\gamma \leftarrow T/N$.
for $i = 0$ **to** $N-1$ **do**
  $\bar{\boldsymbol{z}}_{i+1} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}d)$, $\boldsymbol{z}_{i+1} \leftarrow \mathrm{Transport}^{\kappa}_{\boldsymbol{x}_i^\gamma}(\bar{\boldsymbol{z}}_{i+1})$;
  $\boldsymbol{w}_{i+1} \leftarrow \gamma f(i\gamma, \boldsymbol{x}_i^\gamma) + \sqrt{\gamma}\sigma(i\gamma, \boldsymbol{x}_i^\gamma)\boldsymbol{z}_{i+1}$;
  $\boldsymbol{x}_{i+1}^\gamma \leftarrow \mathrm{Exp}^{\kappa}_{\boldsymbol{x}_i^\gamma}(\boldsymbol{w}_{i+1})$;
end for

$\frac{\exp(\langle \boldsymbol{x}_0, \boldsymbol{x}_1\rangle_\kappa/\epsilon)v(\boldsymbol{x}_1)}{\int_{\mathcal{M}}\exp(\langle \boldsymbol{x}_0, \boldsymbol{x}_1\rangle_\kappa/\epsilon)v(\boldsymbol{x}_1)d\boldsymbol{x}_1}$. Then, the process $S_{|\boldsymbol{x}_0,\boldsymbol{x}_1} = W^{\mathcal{M}}_{|\boldsymbol{x}_0,\boldsymbol{x}_1}$ is the Schrödinger bridge between p_0 and $p_v(\boldsymbol{x}_1) = \int_{\mathcal{M}} \pi_v(\boldsymbol{x}_0, \boldsymbol{x}_1)d\boldsymbol{x}_0$ on the manifold. The corresponding KL divergence exists closed-form expression regarding Riemannian reciprocal process T_π, detailed in Appendix B. Note that, *the advantage is that we no longer need the iterative procedure but conduct only one step of the proposed projection.*

Algorithm. *Consequently, it enables a new objective function for learning Riemannian Schrödinger bridges.* Riemannian Wiener process (also known as Brownian motion) can be simulated by geodesic random walk [Bortoli *et al.*, 2022], whose procedure is described in Algo. 1. Transport, Exp and Log denote the parallel transport, exponential map and logarithmic map, respectively. With the time reversal process described in the SDE, $d\boldsymbol{x}_t = \frac{\mathrm{Log}_{\boldsymbol{x}_t}(\boldsymbol{x}_1)}{1-t}dt + \sqrt{\epsilon}dW_t^{\mathcal{M}}$, we derive the objective function as follows.

Theorem 3 (Objective Function). *Given $p_0, p_1 \in \mathcal{P}^{\mathcal{M}}$ and Riemannian Wiener process, we have $T^\star$ described in $d\boldsymbol{x}_t = g_{\boldsymbol{\theta}}(\boldsymbol{x}_t, t)dt + \sqrt{\epsilon}dW_t^{\mathcal{M}}$ is the Riemannian Schrödinger bridge connecting p_0 and p_1, if the drift $g_{\boldsymbol{\theta}}$ is the minimzer of the following objective,*

$$\mathcal{L}_{\mathrm{bridge}} = \mathbb{E}_{\substack{(\boldsymbol{x}_0,\boldsymbol{x}_1)\sim\pi, t\sim U_{[0,1]} \\ \boldsymbol{x}_t \sim W_{|\boldsymbol{x}_0,\boldsymbol{x}_1}}}\left[\left\| g_{\boldsymbol{\theta}}(\boldsymbol{x}_t, t) - \frac{\mathrm{Log}^{\kappa}_{\boldsymbol{x}_t}(\boldsymbol{x}_1)}{1-t}\right\|_\kappa^2\right], \quad (10)$$

where $U_{[0,1]}$ denotes the uniform distribution over $[0, 1]$.

The strength lies in that Eq. 10 offers a simple yet effective mean squared error objective, avoiding the expensive Riemannian divergence or heat kernel. Next, we develop the geometric understanding with the notion of geodesic.

Geodesic. *Given $\boldsymbol{x}_0, \boldsymbol{x}_1$ in the geodesically complete manifold $(\mathcal{M}^\kappa, \mathfrak{g})$, the geodesics $\boldsymbol{x}_t = \mathrm{Exp}^{\kappa}_{\boldsymbol{x}_0}(t\,\mathrm{Log}^{\kappa}_{\boldsymbol{x}_0}(\boldsymbol{x}_1))$ is the shortest curve connecting them under $\mathfrak{g}$.*

The intuition behind is that the initial and terminal distributions approach each other along the shortest path. The remaining issue is on the form of $g_{\boldsymbol{\theta}}$. We parameterize the adjusted potential v with a mixture of wrapped Gaussian on the manifold $\mathcal{N}_{\mathcal{M}}(\boldsymbol{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Accordingly, the drift takes the form of $g_{\boldsymbol{\theta}}(\boldsymbol{x}_t, t) = \epsilon\nabla_{\boldsymbol{x}_t}\log\left(\mathcal{N}_{\mathcal{M}}(\boldsymbol{x}_t|0, \epsilon(1-t)\boldsymbol{I}_d)v(\boldsymbol{x}_t, t)\right)$, where $v(\boldsymbol{x}_t, t)$ is given by the aforementioned mixture. Note that, g is easy to compute with the closed form and diagonal covariance. In summary, the overall training procedure is given in Algo. 2. Finally, Trace *integrates the advantages of Schrödinger bridge and Riemannian geometry to directly infer the sources without any iterative procedure.*

Algorithm 2 Training Trace.

Input: Graph G, The set of the sources and final infection pairs.
Output: Learned parameters of the transferable localizer.

Perform normalization in light of transferability;
repeat
  Sample triangles over G for the functional curvature;
  Sample batch $\{\boldsymbol{x}_0, \boldsymbol{x}_1\} \sim \pi$ and $\{t_n\} \sim U_{[0,1]}$;
  Sample batch $\{\boldsymbol{x}_t\} \sim W_{[\boldsymbol{x}_0,\boldsymbol{x}_1]}$ by Algo. 1;
  Compute the overall loss $\mathcal{J} = \alpha\mathcal{L}_{\mathrm{manifold}} + (1-\alpha)\mathcal{L}_{\mathrm{bridge}}$;
  Update parameters via the gradient descent method;
until converged

Connection with Recent Advances. In Riemannian geometry, [Thornton *et al.*, 2022] and [Jo and Hwang, 2024] follow the iterative procedure in spirit of IPF and IMF, respectively, while we conduct the one step projection. [Chen and Lipman, 2024] discusses the pushforward from a Gaussian but we are interested in a map of arbitrary distributions. In Euclidean space, [Liu *et al.*, 2023] conducts such map without randomness, while we consider the randomness in the social dynamic process. LightSB-M [Gushchin *et al.*, 2024] can be regarded as a special case of ours, given that $\mathrm{Log}_{\boldsymbol{x}_t}(\boldsymbol{x}_1)$ converges to $\boldsymbol{x}_1 - \boldsymbol{x}_t$ in the limit at zero curvature.

On Transferability. It is the Riemannian manifold that endows Trace with transferability, as it offers a unified construction for different curvatures, corresponding to diverse graph structures. In the transfer setting, we leverage the pretrained parameters of Schrödinger bridge to capture underlying social dynamics. Its functional curvature is fine-tuned on the new graph, so that we adapt the knowledge according to the structural information encapsulated in the curvature.

4 Experiment

4.1 Experimental Settings

Datasets and Baselines. The experiments are conducted on both synthetic and real datasets. The synthetic datasets consist of Jazz, Network Science, Cora-ML, and Facebook [Ling *et al.*, 2022]. We follow [Ling *et al.*, 2022] for diffusion simulation, and the ratio of the sources is set as 10%. The four real datasets include Android, Christianity, Douban, and Twitter of real information cascades [Huang *et al.*, 2023]. We follow [Huang *et al.*, 2023] to define the sources and final observations, and the time ratio of the sources is 5%. The 7 baselines are Netsleuth [Prakash *et al.*, 2012], LPSI [Wang *et al.*, 2017], DDMSL [Yan *et al.*, 2023], TGASI [Hou *et al.*, 2023], SLVAE [Ling *et al.*, 2022], IVGD [Wang *et al.*, 2022] and GCNSI [Dong *et al.*, 2019]. Transferability in source localization is under-explored, and our Trace is proposed to bridge this gap. (Datasets and baselines are detailed in Appendix D.)

Evaluation Metrics. We evaluate the comparison methods with three popular metrics, i.e., AUC, PR and F1 score [Ling *et al.*, 2022; Hou *et al.*, 2023]. Each case undergoes 10 independent runs for statistical robustness. We report the mean value with the standard deviation.

Preprint – IJCAI 2025: This is the accepted version made available for conference attendees.
Do not cite. The final version will appear in the IJCAI 2025 proceedings.

Method	Jazz			Network Science			Cora-ML			Facebook		
	AUC	PR	F1	AUC	PR	F1	AUC	PR	F1	AUC	PR	F1
SLVAE	0.9447	0.8040	0.7701	**0.9711**	0.6541	0.6099	0.8686	0.6695	0.6112	0.5010	0.2000	0.3300
IVGD	0.9375	0.8094	0.7877	0.9606	0.7540	0.7583	0.8280	**0.9381**	0.8233	OOM	OOM	OOM
DDMSL	0.9101	0.8028	0.7506	0.9227	0.7353	0.6198	0.8980	0.8670	0.8225	0.8531	0.8192	0.7931
TGASI	0.9212	0.7821	0.7156	0.9448	0.6397	0.6288	0.8962	0.7631	0.7497	0.9312	0.8223	0.7367
GCNSI	0.6834	0.5622	0.3701	0.5433	0.1613	0.1047	0.5321	0.1725	0.1158	0.6671	0.6654	0.6021
LPSI	0.5553	0.2500	0.2819	0.5614	0.1362	0.2072	0.4986	0.1072	0.1752	0.2710	0.0561	0.1050
Netsleuth	0.5043	0.4351	0.4001	0.5432	0.2642	0.1625	0.5521	0.1627	0.1635	0.5841	0.3753	0.1642
Trace	**0.9454**	**0.8131**	**0.8012**	0.9612	**0.7846**	**0.7656**	**0.9053**	0.8789	**0.8256**	**0.9731**	**0.8742**	**0.8550**

Table 1: Performance on synthetic datasets of Jazz, Network Science, Cora-ML and Facebook. Best results are in **boldfaced**.

Method	Android			Christianity			Douban			Twitter		
	AUC	PR	F1	AUC	PR	F1	AUC	PR	F1	AUC	PR	F1
SLVAE	0.5532	0.4235	0.4563	0.6024	0.4375	0.4162	0.5323	0.4863	0.4632	0.5235	0.4582	0.4132
IVGD	0.5462	0.4542	0.4375	0.5143	0.4213	0.4798	0.4532	0.4163	0.3630	0.4736	0.4113	0.3812
DDMSL	0.5485	0.4790	0.4438	0.5194	0.4727	0.4389	0.4483	0.4232	0.3891	0.4823	0.4076	0.3893
TGASI	0.5033	0.3512	0.2976	0.4723	0.3433	0.2512	0.5102	0.4621	0.3622	0.3813	0.2672	0.2048
GCNSI	0.4458	0.0222	0.0417	0.4193	0.0265	0.0411	0.4503	0.0201	0.0363	0.4771	0.0130	0.0270
LPSI	0.4382	0.0181	0.0387	0.4192	0.0305	0.0455	0.4468	0.0128	0.0309	0.4790	0.0189	0.0250
Netsleuth	0.4531	0.0392	0.0367	0.4241	0.0182	0.0351	0.4444	0.0106	0.0101	0.4640	0.0423	0.0204
Trace	**0.5732**	**0.5011**	**0.4651**	**0.6057**	**0.5012**	**0.4851**	**0.5815**	**0.4997**	**0.4864**	**0.5622**	**0.4712**	**0.4432**

Table 2: Performance on real-world datasets of Android, Christianity, Douban and Twitter. Best results are in **boldfaced**.

Model Configuration and Reproducibility. In light of transferability, Trace performs a graph normalization process. Concretely, the input features are given by the eigenvectors of K largest eigenvalues in graph Laplacian, so that the feature dimension of different datasets is normalized to the predefined K. Correspondingly, the functional curvature $\kappa_{\boldsymbol{\theta}}$ is fed by the tuple of K largest eigenvalues. Then, we use the exponential map with $\kappa_{\boldsymbol{\theta}}$ to obtain the initial encoding on the manifold.

Variant	Facebook		Android	
	AUC	F1	AUC	F1
$w/oFCurv$	83.13$\pm$1.78	77.57$\pm$1.27	43.42$\pm$0.57	37.21$\pm$0.67
$w/oFRiem$	90.32$\pm$0.44	83.21$\pm$0.61	54.21$\pm$0.72	44.32$\pm$0.92
$w/oMPara$	93.44$\pm$0.29	83.81$\pm$0.48	56.81$\pm$0.86	46.03$\pm$1.32
Trace	97.31$\pm$0.23	85.50$\pm$0.45	57.32$\pm$0.69	46.51$\pm$1.03

Table 3: Ablation study in terms of AUC (%) and F1(%).

4.2 Empirical Results & Discussion

Main Results on Real and Synthetic Datasets. The performance on synthetic and real datasets are collected in Table 1 and Table 2, respectively. On synthetic datasets, we generally obtain the best performance except two cases of runner up. On the real datasets, infection observations are fed into the localization methods and the infection simulation process is disabled. IVGD runs into Out-Of-Memory (OOM) error on Facebook dataset. Comparing to the synthetic datasets, all models have experienced the significant decline on real datasets given the complex dynamics in reality. The proposed Trace consistently achieves the best results in terms of AUC, PR and F1 score in Table 2. In addition, to evaluate the efficiency, we show the training time in each epoch in Fig. 3. Trace exhibits satisfactory efficiency compared to the baselines. The following part demonstrates how the proposed components contribute to Trace.

Ablation Study. We study the effectiveness of (a) structure-awareness of Schrödinger bridge, (b) fully Riemannian operations, and (c) the new learning algorithm. In Trace, the Schrödinger bridge is governed by $\kappa_{\boldsymbol{\theta}}$ that describes the graph structure. Accordingly, $w/oFCurv$ variant gives a counterpart regardless of structural information. Concretely, we conduct logarithmic map on the encoding and replace our design with the usual Euclidean Schrödinger bridge [Thornton *et al.*, 2022]. As for Riemannian operations, $w/oFRiem$ variant leverages an additional tangent space with logarithmic and exponential maps. To evaluate the proposed learning algorithm, $w/oMPara$ variant utilizes a usual parameterized drift of [Thornton *et al.*, 2022] and employs the traditional IMF procedure. The results are summarized in Table 3, and the main findings are listed as follows: 1) Trace consistently outperforms $w/oFCurv$, and it suggests the importance of structural information in modeling the distribution map for source localization. 2) Comparing to $w/oFRiem$, we receive performance gain with the proposed operations. 3) $w/oMPara$ exhibits inferior accuracy. It is due to the optimization error in IMF procedure, and a similar result has been reported in Euclidean space [Gushchin *et al.*, 2024]. More importantly, the proposed algorithm demonstrates impressive efficiency, e.g., its training time is less than $1/12$ of IMF on Douban as in Fig. 2.

On Geometry. Trace is presented as a Riemannian model, while existing methods perform localization in the Euclidean geometry. This part evaluates the impact of representation space with geometric ablation. Hyperbolic and hyperspherical Trace are created by setting the functional curvature as -1 and 1, respectively. Without loss of generality, the Euclidean counterpart is instantiated with corresponding operations. The comparison results are given in Fig. 4. It

Preprint – IJCAI 2025: This is the accepted version made available for conference attendees.
Do not cite. The final version will appear in the IJCAI 2025 proceedings.

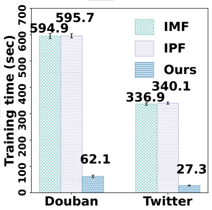


Figure 2: IMF/IPF

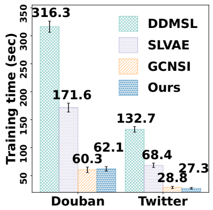


Figure 3: Time

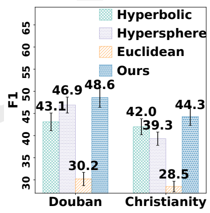
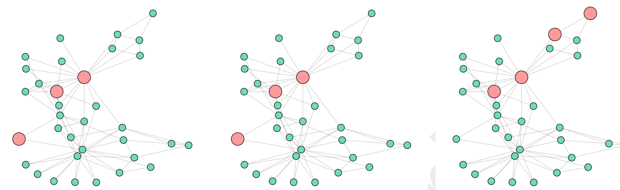


Figure 4: Geometry



(a) Ground truth (b) Trace (c) TGASI

Figure 5: Visualization on Karate dataset.

Pre-training	Testing Datasets			
	Android	Christianity	Douban	Twitter
Android	**46.51**$\pm$0.56	46.81$\pm$0.75	45.33$\pm$0.33	42.45$\pm$0.58
Christianity	44.56$\pm$0.67	**48.51**$\pm$0.42	45.87$\pm$0.43	41.89$\pm$0.92
Douban	45.11$\pm$0.27	47.31$\pm$0.59	**48.64**$\pm$0.51	43.17$\pm$0.37
Twitter	45.23$\pm$0.63	46.78$\pm$0.54	47.51$\pm$0.53	**44.32**$\pm$0.74

Table 4: Performance in the transfer setting with different pre-training datasets in terms of F1 score (%).

shows that the geometry plays a significant role in source localization, as evidenced by the performance across different geometries. Note that, the learnable curvature potentially enhances accuracy through improved geometric matching and offers the advantage for transferability, as in the next part.

Transferability and Impact of Pre-training Dataset. We demonstrate the transferability of the proposed model. Tables 1 and 2 collect the results of training and testing models on the same dataset. Here, `Trace` is first pre-trained on one dataset, and then the model is fine-tuned and tested on another. We do experiments on different pre-training datasets for comprehensive evaluation. F1 scores are reported in Table 4, which shows the effectiveness of transferring the pre-trained model to new graphs. Additionally, `Trace` exhibits robustness over different pre-training datasets. We argue that the Schrödinger bridge on the manifold captures the underlying social dynamics, while the learnable curvature adapts knowledge according to the graph structure.

Case Study & Visualization. We conduct a case study on Karate dataset, and visualize the inference results in Fig. 5, where the inferred sources are marked in red color. Compared to TGASI in Fig. 5(c), `Trace` shows a more precise alignment with the actual source nodes in Fig. 5(a).

5 Related Work

Source Localization. Early practices leverage statistic metrics, e.g., Jordan centrality, to infer the sources [Zhu *et al.*, 2017]. Nowadays, the deep localization models can be roughly categorized into two groups. The first group primarily focuses on static or temporal characteristics of the sources [Cheng *et al.*, 2024; Hou *et al.*, 2023; Wang *et al.*, 2023; Hou *et al.*, 2024], while the second group models the backward dynamics along with the forward information diffusion [Wang *et al.*, 2022; Ling *et al.*, 2022; Huang *et al.*, 2023; Yan *et al.*, 2023; Zhang *et al.*, 2024]. We are in line with the latter, and the Schrödinger bridge presents the duality via forward and backward SDEs. Recently, [Wang *et al.*, 2024] jointly exploits co-related networks to improve localization accuracy, while [Ling *et al.*, 2024] studies a different problem of source localization from one network to another. Distinguishing from prior Euclidean models, our Riemannian `Trace` is endowed with the transferability.

Generative Models & Schrödinger Bridge. Diffusion- and flow-based models emerge as popular generative models, described as SDEs [Song *et al.*, 2021] and probability ODEs [Lipman *et al.*, 2023], respectively. Both of them conduct data generation from a simple distribution, e.g., Gaussian, while we are devoted to bridge the source and final distributions. We notice that rectified flow [Liu *et al.*, 2023] builds such distribution map without randomness. However, the randomness is often nested in the information propagation process. Hence, we consider the Schrödinger bridge that models the map between two arbitrary distributions [Shi *et al.*, 2023; Gushchin *et al.*, 2024; Liu *et al.*, 2024]. It is noteworthy that structural information has not yet been encapsulated into the Schrödinger bridge, and we offer an alternative to fill this gap.

Riemannian Graph Learning. Compared to the traditional Euclidean space, Riemannian manifolds show superiority in graph clustering [Sun *et al.*, 2024c; Sun *et al.*, 2024b; Sun *et al.*, 2023b], node classification [Sun *et al.*, 2022b; Sun *et al.*, 2024d; Sun *et al.*, 2024e; Sun *et al.*, 2025a; Fu *et al.*, 2023; Li *et al.*, 2022; Wei *et al.*, 2024], graph structural learning [Sun *et al.*, 2023a] as well as modeling graph dynamics [Sun *et al.*, 2023c; Sun *et al.*, 2022a; Sun *et al.*, 2021]. Recently, ODEs and SDEs are extended to Riemannian manifolds for graph generation [Sun *et al.*, 2025b; Sun *et al.*, 2024a; Fu *et al.*, 2024]. We notice that [Jo and Hwang, 2024] build a Riemannian Schrödinger bridge. Besides the different designs on curvature, another key contribution is that we design a new learning algorithm more efficient than previous ones.

6 Conclusion

This paper presents the Riemannian source localizer named `Trace` with the transferability to different graphs. Specifically, the structural Schrödinger bridge governed by functional curvature is established to model the direct transformation between source and final distributions. Furthermore, we design a new learning algorithm that conducts the efficient geodesic bridge matching with the provable optimal projection for Riemannian measures. Extensive experiments show the effectiveness and transferability of the proposed model.

Preprint – IJCAI 2025: This is the accepted version made available for conference attendees.
Do not cite. The final version will appear in the IJCAI 2025 proceedings.

Acknowledgments

This work is supported in part by National Natural Science Foundation of China under Grant 62202164, and the Fundamental Research Funds for the Central Universities 2025MS024. Philip S. Yu is supported in part by NSF under grants III-2106758, and POSE-2346158. The corresponding author is Li Sun.

References

[Bortoli *et al.*, 2021] Valentin De Bortoli, James Thornton, Jeremy Heng, and Arnaud Doucet. Diffusion schrödinger bridge with applications to score-based generative modeling. In *Proceedings of the 34th NeurIPS*, pages 17695–17709, 2021.

[Bortoli *et al.*, 2022] Valentin De Bortoli, Emile Mathieu, Michael J. Hutchinson, James Thornton, Yee Whye Teh, and Arnaud Doucet. Riemannian score-based generative modelling. In *Proceedings of the 35th NeurIPS*, 2022.

[Chen and Lipman, 2024] Ricky T. Q. Chen and Yaron Lipman. Flow matching on general geometries. In *Proceedings of the 12th ICLR*, 2024.

[Cheng *et al.*, 2024] Le Cheng, Peican Zhu, Keke Tang, Chao Gao, and Zhen Wang. GIN-SD: source detection in graphs with incomplete nodes via positional encoding and attentive fusion. In *Proceedings of the 38th AAAI*, pages 55–63. AAAI Press, 2024.

[Dong *et al.*, 2019] Ming Dong, Bolong Zheng, Nguyen Quoc Viet Hung, Han Su, and GuoHui Li. Multiple rumor source detection with graph convolutional networks. In *Proceedings of the 28th CIKM*, pages 569–578. ACM, 2019.

[Föllmer, 1988] H. Föllmer. Random fields and diffusion processes. *École d'Été de Probabilités de Saint-Flour XV–XVII*, pages 101–203, 1988.

[Fu *et al.*, 2023] Xingcheng Fu, Yuecen Wei, Qingyun Sun, Haonan Yuan, Jia Wu, Hao Peng, and Jianxin Li. Hyperbolic geometric graph representation learning for hierarchy-imbalance node classification. In *Proceedings of the ACM Web Conference (WWW)*, pages 460–468. ACM, 2023.

[Fu *et al.*, 2024] Xingcheng Fu, Yisen Gao, Yuecen Wei, Qingyun Sun, Hao Peng, Jianxin Li, and Xianxian Li. Hyperbolic geometric latent diffusion model for graph generation. In *Proceedings of the 41st ICML*. OpenReview.net, 2024.

[Gu *et al.*, 2019] Albert Gu, Frederic Sala, Beliz Gunel, and Christopher Ré. Learning mixed-curvature representations in products of model spaces. In *Proceedings of the 7th ICLR*, 2019.

[Gushchin *et al.*, 2024] Nikita Gushchin, Sergei Kholkin, Evgeny Burnaev, and Alexander Korotin. Light and optimal schrödinger bridge matching. In *Proceedings of the 41st ICML*. OpenReview.net, 2024.

[Hou *et al.*, 2023] Dongpeng Hou, Zhen Wang, Chao Gao, and Xuelong Li. Sequential attention source identification based on feature representation. In *Proceedings of the 32th IJCAI*, pages 4794–4802. ijcai.org, 2023.

[Hou *et al.*, 2024] Dongpeng Hou, Yuchen Wang, Chao Gao, Xianghua Li, and Zhen Wang. New localization frameworks: User-centric approaches to source localization in real-world propagation scenarios. In *Proceedings of the 33rd ACM CIKM*, pages 839–848. ACM, 2024.

[Huang *et al.*, 2023] Bosong Huang, Weihao Yu, Ruzhong Xie, Jing Xiao, and Jin Huang. Two-stage denoising diffusion model for source localization in graph inverse problems. In *Proceedings of the ECML/PKDD*, volume 14171, pages 325–340. Springer, 2023.

[Jo and Hwang, 2024] Jaehyeong Jo and Sung Ju Hwang. Generative modeling on manifolds through mixture of riemannian diffusion processes. In *Proceedings of the 41st ICML*, 2024.

[Léonard, 2014] C. Léonard. A survey of the schrödinger problem and some of its connections with optimal transport. *Discrete & Continuous Dynamical Systems-A*, 34(4):1533–1574, 2014.

[Li *et al.*, 2022] Jianxin Li, Xingcheng Fu, Qingyun Sun, Cheng Ji, Jiajun Tan, Jia Wu, and Hao Peng. Curvature graph generative adversarial networks. In *Proceedings of the ACM Web Conference (WWW)*, pages 1528–1537. ACM, 2022.

[Ling *et al.*, 2022] Chen Ling, Junji Jiang, Junxiang Wang, and Liang Zhao. Source localization of graph diffusion via variational autoencoders for graph inverse problems. In *Proceedings of the 28th SIGKDD*, pages 1010–1020. ACM, 2022.

[Ling *et al.*, 2024] Chen Ling, Tanmoy Chowdhury, Jie Ji, Sirui Li, Andreas Züfle, and Liang Zhao. Source localization for cross network information diffusion. In *Proceedings of the 30th SIGKDD*, pages 5419–5429. ACM, 2024.

[Lipman *et al.*, 2023] Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matthew Le. Flow matching for generative modeling. In *Proceedings of the 11th ICLR*. OpenReview.net, 2023.

[Liu *et al.*, 2023] Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. In *Proceedings of the 11th ICLR*, 2023.

[Liu *et al.*, 2024] Guan-Horng Liu, Yaron Lipman, Maximilian Nickel, Brian Karrer, Evangelos A. Theodorou, and Ricky T. Q. Chen. Generalized schrödinger bridge matching. In *Proceedings of the 12th ICLR*. OpenReview.net, 2024.

[Prakash *et al.*, 2012] B. Aditya Prakash, Jilles Vreeken, and Christos Faloutsos. Spotting culprits in epidemics: How many and which ones? In *Proceedings of the 12th ICDM*, pages 11–20. IEEE, 2012.

Preprint – IJCAI 2025: This is the accepted version made available for conference attendees.
Do not cite. The final version will appear in the IJCAI 2025 proceedings.

[Schrödinger, 1932] E. Schrödinger. Sur la théorie relativiste de l'électron et l'interprétation de la mécanique quantique. *Annales de l'Institut Henri Poincaré*, 2(4):269–310, 1932.

[Shi *et al.*, 2023] Yuyang Shi, Valentin De Bortoli, Andrew Campbell, and Arnaud Doucet. Diffusion schrödinger bridge matching. In *Proceeding of the 36th NeurIPS*, 2023.

[Song *et al.*, 2021] Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *Proceedings of the 9th ICLR*. OpenReview.net, 2021.

[Sun *et al.*, 2021] Li Sun, Zhongbao Zhang, Jiawei Zhang, Feiyang Wang, Hao Peng, Sen Su, and Philip S. Yu. Hyperbolic variational graph neural network for modeling dynamic graphs. In *Proceedings of the 35th AAAI*, pages 4375–4383, 2021.

[Sun *et al.*, 2022a] Li Sun, Junda Ye, Hao Peng, and Philip S. Yu. A self-supervised riemannian GNN with time varying curvature for temporal graph learning. In *Proceedings of the 31st CIKM*, pages 1827–1836, 2022.

[Sun *et al.*, 2022b] Li Sun, Zhongbao Zhang, Junda Ye, Hao Peng, Jiawei Zhang, Sen Su, and Philip S. Yu. A self-supervised mixed-curvature graph neural network. In *Proceedings of the 36th AAAI*, pages 4146–4155, 2022.

[Sun *et al.*, 2023a] Li Sun, Zhenhao Huang, Hua Wu, Junda Ye, Hao Peng, Zhengtao Yu, and Philip S. Yu. Deepricci: Self-supervised graph structure-feature co-refinement for alleviating over-squashing. In *Proceedings of the 23rd ICDM*, pages 558–567, 2023.

[Sun *et al.*, 2023b] Li Sun, Feiyang Wang, Junda Ye, Hao Peng, and Philip S. Yu. Congregate: contrastive graph clustering in curvature spaces. In *Proceedings of the 32nd IJCAI*, pages 2296–2305, 2023.

[Sun *et al.*, 2023c] Li Sun, Junda Ye, Hao Peng, Feiyang Wang, and Philip S. Yu. Self-supervised continual graph learning in adaptive riemannian spaces. In *Proceedings of the 37th AAAI*, pages 4633–4642, 2023.

[Sun *et al.*, 2024a] Li Sun, Jingbin Hu, Mengjie Li, and Hao Peng. R-ode: Ricci curvature tells when you will be informed. In *Proceedings of the ACM SIGIR*, 2024.

[Sun *et al.*, 2024b] Li Sun, Jingbin Hu, Suyang Zhou, Zhenhao Huang, Junda Ye, Hao Peng, Zhengtao Yu, and Philip S. Yu. Riccinet: Deep clustering via a riemannian generative model. In *Proceedings of the ACM Web Conference (WWW)*, pages 4071–4082, 2024.

[Sun *et al.*, 2024c] Li Sun, Zhenhao Huang, Hao Peng, Yujie Wang, Chunyang Liu, and Philip S. Yu. Lsenet: Lorentz structural entropy neural network for deep graph clustering. In *Proceedings of the 41st ICML*, pages 47078–47104, 2024.

[Sun *et al.*, 2024d] Li Sun, Zhenhao Huang, Qiqi Wan, Hao Peng, and Philip S. Yu. Spiking graph neural network on riemannian manifolds. In *Advances in NeurIPS*, 2024.

[Sun *et al.*, 2024e] Li Sun, Zhenhao Huang, Zixi Wang, Feiyang Wang, Hao Peng, and Philip S. Yu. Motif-aware riemannian graph neural network with generative-contrastive learning. In *Proceedings of the 38th AAAI*, pages 9044–9052, 2024.

[Sun *et al.*, 2025a] Li Sun, Zhenhao Huang, Suyang Zhou, Qiqi Wan, Hao Peng, and Philip S. Yu. Riemanngfm: Learning a graph foundation model from riemannian geometry. In *Proceedings of the ACM Web Conference (WWW)*, pages 1154–1165. ACM, 2025.

[Sun *et al.*, 2025b] Li Sun, Ziheng Zhang, Zixi Wang, Yujie Wang, Qiqi Wan, Hao Li, Hao Peng, and Philip S. Yu. Pioneer: Physics-informed riemannian graph ode for entropy-increasing dynamics. In *Proceedings of the 39th AAAI*, 2025.

[Thornton *et al.*, 2022] James Thornton, Michael J. Hutchinson, Emile Mathieu, Valentin De Bortoli, Yee Whye Teh, and Arnaud Doucet. Riemannian diffusion schrödinger bridge. *CoRR*, abs/2207.03024, 2022.

[Wang *et al.*, 2017] Zheng Wang, Chaokun Wang, Jisheng Pei, and Xiaojun Ye. Multiple source detection without knowing the underlying propagation model. In *Proceedings of the 31th AAAI*, pages 217–223. AAAI Press, 2017.

[Wang *et al.*, 2022] Junxiang Wang, Junji Jiang, and Liang Zhao. An invertible graph diffusion neural network for source localization. In *Proceedings of the 22th ACM TheWebConf (WWW)*, pages 1058–1069. ACM, 2022.

[Wang *et al.*, 2023] Zhen Wang, Dongpeng Hou, Chao Gao, Xiaoyu Li, and Xuelong Li. Lightweight source localization for large-scale social networks. In *Proceedings of ACM TheWebConf (WWW)*, pages 286–294. ACM, 2023.

[Wang *et al.*, 2024] Zhen Wang, Dongpeng Hou, Shu Yin, Chao Gao, and Xianghua Li. Joint source localization in different platforms via implicit propagation characteristics of similar topics. In *Proceedings of the 33th IJCAI*, pages 2424–2432. ijcai.org, 2024.

[Wei *et al.*, 2024] Yuecen Wei, Haonan Yuan, Xingcheng Fu, Qingyun Sun, Hao Peng, Xianxian Li, and Chunming Hu. Poincaré differential privacy for hierarchy-aware graph embedding. In *Proceedings of the 38th AAAI*, pages 9160–9168, 2024.

[Yan *et al.*, 2023] Xin Yan, Hui Fang, and Qiang He. Diffusion model for graph inverse problems: Towards effective source localization on complex networks. In *Proceedings of the 36th NeurIPS*, 2023.

[Zhang *et al.*, 2024] Zonghan Zhang, Zijian Zhang, and Zhiqian Chen. Multiple-source localization from a single-snapshot observation using graph bayesian optimization. In *Proceedings of the 38th AAAI*, pages 22538–22546. AAAI Press, 2024.

[Zhu *et al.*, 2017] Kai Zhu, Zhen Chen, and Lei Ying. Catch'em all: Locating multiple diffusion sources in networks with partial observations. In *Proceedings of the 31th AAAI*, pages 1676–1683. AAAI Press, 2017.