Aligned Diffusion Schrödinger Bridges

Vignesh Ram Somnath[*1,2] **Matteo Pariset**[*1,3] **Ya-Ping Hsieh**[1]
Maria Rodriguez Martinez[2] **Andreas Krause**[1] **Charlotte Bunne**[1]

[1]Department of Computer Science, ETH Zürich
[2]IBM Research Zürich
[3]Department of Computer Science, EPFL

Abstract

Diffusion Schrödinger bridges (DSB) have recently emerged as a powerful framework for recovering stochastic dynamics via their marginal observations at different time points. Despite numerous successful applications, existing algorithms for solving DSBs have so far failed to utilize the structure of *aligned* data, which naturally arises in many biological phenomena. In this paper, we propose a novel algorithmic framework that, for the first time, solves DSBs while respecting the data alignment. Our approach hinges on a combination of two decades-old ideas: The classical Schrödinger bridge theory and Doob's *h-transform*. Compared to prior methods, our approach leads to a simpler training procedure with lower variance, which we further augment with principled regularization schemes. This ultimately leads to sizeable improvements across experiments on synthetic and real data, including the tasks of predicting conformational changes in proteins and temporal evolution of cellular differentiation processes.

1 INTRODUCTION

The task of transforming a given distribution into another lies at the heart of many modern machine learning applications such as single-cell genomics (Tong et al., 2020; Schiebinger et al., 2019; Bunne et al., 2022a), meteorology (Fisher et al., 2009), and robotics (Chen et al., 2021a). To this end, diffusion Schrödinger bridges (De Bortoli et al., 2021; Chen et al., 2022a; Vargas et al., 2021; Liu et al., 2022b) have recently emerged as a powerful paradigm due to their ability to generalize prior deep diffusion-based models, notably score matching with Langevin dynamics (Song and Ermon, 2019; Song et al., 2021) and denoising diffusion probabilistic models (Ho et al., 2020), which have achieved the state-of-the-art on many generative modeling problems.

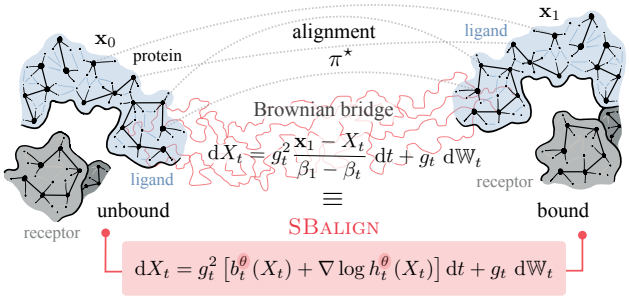

Figure 1: Overview of SBALIGN: In biological tasks such as protein docking, one is naturally provided with *aligned* data in the form of unbound and bound structures of participating proteins. Our goal is to therefore recover a stochastic trajectory from the unbound ($\mathbf{x}_0$) to the bound ($\mathbf{x}_1$) structure. To achieve this, we connect the characterization of an SDE conditioned on $\mathbf{x}_0$ and $\mathbf{x}_1$ (utilizing the Doob's *h-transform*) with that of a Brownian bridge between $\mathbf{x}_0$ and $\mathbf{x}_1$ (classical Schrödinger bridge theory). We show that this leads to a simpler training procedure with lower variance and strong empirical results.

Despite the wide success of DSBs solvers, a significant limitation of existing frameworks is that they fail to capture the *alignment* of data: If $\hat{\mathbb{P}}_0, \hat{\mathbb{P}}_1$ are two (empirical) distributions between which we wish to interpolate, then a tacit assumption in the literature is that the dependence of $\hat{\mathbb{P}}_0$ and $\hat{\mathbb{P}}_1$ is unknown and somehow has to be recovered. Such an assumption, however, ignores important scenarios where the data is *aligned*, meaning that the samples from $\hat{\mathbb{P}}_0$ and $\hat{\mathbb{P}}_1$ naturally come in pairs $(\mathbf{x}_0^i, \mathbf{x}_1^i)_i^N$, which is common in many biological phenomena. Proteins, for instance, undergo conformational changes upon interactions with other biomolecules (protein docking, see Fig. 1). The goal is to model conformational changes by recovering a (stochastic) trajectory $\mathbf{x}_t$ based on the positions observed at

[*]Equal contribution.

Proceedings of the 39th Conference on Uncertainty in Artificial Intelligence (UAI 2023), PMLR 216:1985–1995.

two-time points $(\mathbf{x}_0, \mathbf{x}_1)$. Failing to incorporate this alignment would mean that we completely ignore information on the correspondence between the initial and final points of the molecules, resulting in a much harder problem than necessary. Beyond, the recent use of SBs has been motivated by an important task in molecular biology: Cells change their molecular profile throughout developmental processes (Schiebinger et al., 2019; Bunne et al., 2022b) or in response to perturbations such as cancer drugs (Lotfollahi et al., 2019; Bunne et al., 2021). As most measurement technologies are destructive assays, i.e., the same cell cannot be observed twice nor fully profiled over time, these methods aim at reconstructing cell dynamics from *unpaired* snapshots. Recent developments in molecular biology, however, aim at overcoming this technological limitation. For example, Chen et al. (2022b) propose a transcriptome profiling approach that preserves cell viability. Weinreb et al. (2020) capture cell differentiation processes by clonally connecting cells and their progenitors through barcodes (see Appendix).

Motivated by these observations, the goal of this paper is to propose a novel algorithmic framework for solving DSBs with (partially) *aligned* data. Our approach is in stark contrast to existing works which, due to the lack of data alignment, all rely on some variants of *iterative proportional fitting* (IPF) (Fortet, 1940; Kullback, 1968) and are thus prone to numerical instability. On the other hand, via a combination of the original theory of Schrödinger bridges (Schrödinger, 1931; Léonard, 2013) and the key notion of Doob's h-*transform* (Doob, 1984; Rogers and Williams, 2000), we design a novel loss function that completely bypasses the IPF procedure and can be trained with much lower variance.

To summarize, we make the following contributions:

- To our best knowledge, we consider, for the first time, the problem of interpolation with *aligned* data. We rigorously formulate the problem in the DSB framework.
- Based on the theory of Schrödinger bridges and h-transform, we derive a new loss function that, unlike prior work on DSBs, does not require an IPF-like procedure to train. We also propose principled regularization schemes to further stabilize training.
- We describe how interpolating aligned data can provide better reference processes for use in classical DSBs, paving the way to hybrid aligned/non-aligned Schrödinger bridges (SBs).
- We evaluate our proposed framework on both synthetic and real data. For experiments utilizing real data, we consider two tasks where such aligned data is naturally available. The first is the task of developmental processes in single-cell biology, and the second involves protein docking. For the protein docking task, a comprehensive treatment is elusive, owing to lack of appropriate datasets. Instead, we consider two associated subproblems: (i) modeling conformational changes between unbound and bound states of a protein, and (ii) rigid protein docking, i.e., identifying the best relative orientation. Our method demonstrates a considerable improvement over prior methods across various metrics, thereby substantiating the importance of taking the data alignment into account.

Related work. Solving DSBs is a subject of significant interest in recent years and has flourished in a number of different algorithms (De Bortoli et al., 2021; Chen et al., 2022a; Vargas et al., 2021; Bunne et al., 2023; Liu et al., 2022a). However, all these previous approaches focus on *unaligned* data, and therefore the methodologies all rely on IPF and are hence drastically different from ours. In the experiments, we will demonstrate the importance of considering the alignment of data.

An important ingredient in our theory is Doob's h-transform, which has recently also been utilized by Liu et al. (2023) to solve the problem of constrained diffusion. However, their fundamental motivation is different from ours. Liu et al. (2023) focus on learning the drift of the diffusion model and the h-transform *together*, whereas ours is to read off the drift *from* the h-transform with the help of *aligned data*. Consequently, there is no overlap between the two algorithms and their intended applications.

To the best of our knowledge, the concurrent work of Tong et al. (2023) is the only existing framework that can tackle aligned data, which, however, is not their original motivation. In the context of solving DSBs, their algorithm can be seen as learning a vector field that generates the correct *marginal* probability (cf. Tong et al., 2023, Proposition 4.3). Importantly, this is different from our aim of finding the *pathwise* optimal solution of DSBs: If $(\mathbf{x}^i_{0,\text{test}})_{i=1}^m$ is a test data set for which we wish to predict their destinations, then the framework of Tong et al. (2023) can only ensure that the marginal distribution $(\mathbf{x}^i_{1,\text{test}})_{i=1}^m$ is correct, whereas ours is capable of predicting that $\mathbf{x}^i_{1,\text{test}}$ is precisely the destination of $\mathbf{x}^i_{0,\text{test}}$ for each i. This latter property is highly desirable in tasks like ML-accelerated protein docking.

To solve aligned SB problems, we rely on mixtures of diffusion processes. Like in Peluchetti (2023), we construct them from pairings and define an associated training objective inspired by score-based modeling. However, we represent the learned drift as a sum of the solution to an SB problem (b) and a pairing-related term ($\nabla \log h$). We parametrize the second part of the drift with neural networks, unlike Schauer et al. (2017) which use an auxiliary (simpler) process.

2 BACKGROUND

Problem formulation. Suppose that we are given access to i.i.d. *aligned* data $(\mathbf{x}^i_0, \mathbf{x}^i_1)_{i=1}^N$, where the marginal distri-

bution of $\mathbf{x}_0^i$'s is $\hat{\mathbb{P}}_0$ and of $\mathbf{x}_1^i$'s is $\hat{\mathbb{P}}_1$. Typically, we view $\hat{\mathbb{P}}_0$ as the empirical marginal distribution of a stochastic process at time $t = 0$, and $\hat{\mathbb{P}}_1$ as the empirical marginal observed at $t = 1$. The goal is to reconstruct the stochastic process $\mathbb{P}_t$ based on $(\mathbf{x}_0^i, \mathbf{x}_1^i)_{i=1}^N$, i.e., to transform $\hat{\mathbb{P}}_0$ into $\hat{\mathbb{P}}_1$.

Such a task is ubiquitous in biological applications. For instance, understanding how proteins dock to other biomolecules is of significant interest in biology and has become a topic of intense study in recent years (Ganea et al., 2022; Tsaban et al., 2022; Corso et al., 2023). In the protein docking task, $\mathbf{x}_0^i$ represents the 3D structures of the unbound proteins, while $\mathbf{x}_1^i$ represents the 3D structure of the bound complex. Reconstructing a stochastic process that diffuses $\mathbf{x}_0^i$'s to $\mathbf{x}_1^i$'s is tantamount to recovering the energy landscape governing the docking process. Similarly, in molecular dynamics simulations, we have access to trajectories $\left(\mathbf{x}_t^i\right)_{t\in[0,1]}$, where $\mathbf{x}_0^i$ and $\mathbf{x}_1^i$ represent the initial and final positions of the i-th molecule respectively. Any learning algorithm using these simulations should be able to respect the provided alignment.

Diffusion Schrödinger bridges. To solve the interpolation problem, in Section 3, we will invoke the framework of DSBs, which are designed to solve interpolation problems with *unaligned* data. More specifically, given two marginals $\hat{\mathbb{P}}_0$ and $\hat{\mathbb{P}}_1$, the DSB framework proceeds by first choosing a reference process $\mathbb{Q}_t$ using prior knowledge, for instance a simple Brownian motion, and then solve the entropy-minimization problem over all stochastic processes $\mathbb{P}_t$:

$$\min_{\mathbb{P}_0=\hat{\mathbb{P}}_0,\, \mathbb{P}_1=\hat{\mathbb{P}}_1} D_{\mathrm{KL}}(\mathbb{P}_t \| \mathbb{Q}_t). \tag{SB}$$

Despite the fact that many methods exist for solving (SB) (De Bortoli et al., 2021; Chen et al., 2022a; Vargas et al., 2021; Bunne et al., 2023), none of these incorporate data *alignment*. This can be seen by inspecting the objective (SB), in which the coupling information $(\mathbf{x}_0^i, \mathbf{x}_1^i)$ is completely lost as only its individual marginals $\hat{\mathbb{P}}_0, \hat{\mathbb{P}}_1$ play a role therein. Unfortunately, it is well-known that tackling the marginals separately necessitates a forward-backward learning process known as the *iterative proportional fitting* (IPF) procedure (Fortet, 1940; Kullback, 1968), which constitutes the primary reason of high variance training, thereby confronting DSBs with numerical and scalability issues. Our major contribution is therefore to devise the first algorithmic framework that solves the interpolation problem with aligned data *without* resorting to IPF.

3 ALIGNED DIFFUSION SCHRÖDINGER BRIDGES

In this section, we derive a novel loss function for DSBs with aligned data by combining two classical notions: The theory of Schrödinger bridges (Schrödinger, 1931; Léonard, 2013; Chen et al., 2021b) and Doob's h-transform (Doob, 1984; Rogers and Williams, 2000). We then describe how solutions to DSBs with aligned data can be leveraged in the context of classical DSBs.

3.1 LEARNING ALIGNED DIFFUSION SCHRÖDINGER BRIDGES

Static SB and aligned data. Our starting point is the simple and classical observation that (SB) is the continuous-time analogue of the *entropic optimal transport*, also known as the *static* Schrödinger bridge problem (Léonard, 2013; Chen et al., 2021b; Peyré and Cuturi, 2019):

$$\pi^\star := \operatorname*{argmin}_{\mathbb{P}_0=\hat{\mathbb{P}}_0,\, \mathbb{P}_1=\hat{\mathbb{P}}_1} D_{\mathrm{KL}}(\mathbb{P}_{0,1} \| \mathbb{Q}_{0,1}) \tag{1}$$

where the minimization is over all *couplings* of $\hat{\mathbb{P}}_0$ and $\hat{\mathbb{P}}_1$, and $\mathbb{Q}_{0,1}$ is simply the joint distribution of $\mathbb{Q}_t$ at $t = 0, 1$. In other words, if we denote by $\mathbb{P}_t^\star$ the stochastic process that minimizes (SB), then the joint distribution $\mathbb{P}_{0,1}^\star$ necessarily coincides with the $\pi^\star$ in (1). Moreover, since in DSBs, the data is always assumed to arise from $\mathbb{P}_t^\star$, we see that:

> The *aligned* data $(\mathbf{x}_0^i, \mathbf{x}_1^i)_{i=1}^N$ constitutes samples of $\pi^\star$.

This simple but crucial observation lies at the heart of all derivations to come.

Our central idea is to represent $\mathbb{P}_t^\star$ via two different, but equivalent, characterizations, both of which involve $\pi^\star$: That of a *mixture* of reference processes with pinned end points, and that of conditional *stochastic differential equations* (SDEs).

$\mathbb{P}_t^\star$ from $\pi^\star$: $\mathbb{Q}_t$ with pinned end points. For illustration purposes, we will assume that the reference process $\mathbb{Q}_t$ is a Brownian motion with diffusion coefficient g_t:*

$$\mathrm{d}\mathbb{Q}_t = g_t \, \mathrm{d}\mathbb{W}_t. \tag{2}$$

In this case, it is well-known that $\mathbb{Q}_t$ *conditioned* to start at $\mathbf{x}_0$ and end at $\mathbf{x}_1$ can be written in another SDE (Mansuy and Yor, 2008; Liu et al., 2023):

$$\mathrm{d}X_t = g_t^2 \frac{\mathbf{x}_1 - X_t}{\beta_1 - \beta_t} \, \mathrm{d}t + g_t \, \mathrm{d}\mathbb{W}_t \tag{3}$$

where $X_0 = \mathbf{x}_0$ and

$$\beta_t := \int_0^t g_s^2 \, \mathrm{d}s. \tag{4}$$

*Extension to more involved reference processes is conceptually straightforward but notationally clumsy. Furthermore, reference processes of the form (2) are dominant in practical applications (Song et al., 2021; Bunne et al., 2023), so we omit the general case.

We call the processes in (3) the *scaled Brownian bridges* as they generalize the classical Brownian bridge, which corresponds to the case of $g_t \equiv 1$.

The first characterization of $\mathbb{P}_t^\star$ is then an immediate consequence the following classical result in Schrödinger bridge theory: Draw a sample $(\mathbf{x}_0, \mathbf{x}_1) \sim \pi^\star$ and connect them via (3). The resulting path is a sample from $\mathbb{P}_t^\star$ (Léonard, 2013; Chen et al., 2021b). In other words, $\mathbb{P}_t^\star$ is a *mixture* of scaled Brownian bridges, with the mixing weight given by $\pi^\star$.

$\mathbb{P}_t^\star$ from $\pi^\star$: SDE representation. Another characterization of $\mathbb{P}_t^\star$ is that it is itself given by an SDE of the form (Léonard, 2013; Chen et al., 2021b)

$$\mathrm{d}X_t = g_t^2 b_t(X_t)\,\mathrm{d}t + g_t\,\mathrm{d}\mathbb{W}_t. \tag{5}$$

Here, $b_t : \mathbb{R}^d \to \mathbb{R}^d$ is a time-dependent drift function that we wish to learn.

Now, by Doob's h-transform, we know that the SDE (5) *conditioned* to start at $\mathbf{x}_0$ and end at $\mathbf{x}_1$ is given by another SDE (Doob, 1984; Rogers and Williams, 2000):

$$\mathrm{d}X_t = g_t^2[b_t(X_t) + \nabla \log h_t(X_t)]\,\mathrm{d}t + g_t\,\mathrm{d}\mathbb{W}_t \tag{6}$$

where $h_t(\mathbf{x}) := \mathbb{P}(X_1 = \mathbf{x}_1 | X_t = \mathbf{x})$ is the *Doob's h function*. Notice that we have suppressed the dependence of h_t on $\mathbf{x}_0$ and $\mathbf{x}_1$ for notational simplicity.

Loss function. Since both (3) and (6) represent $\mathbb{P}_t^\star$, the solution of the DSBs, the two SDEs must coincide. In other words, suppose we parametrize b_t as b_t^θ, then, by matching terms in (3) and (6), we can learn the optimal parameter $\theta^\star$ via optimization of the loss function

$$L(\theta) := \mathbb{E}\left[\int_0^1 \left\| \frac{\mathbf{x}_1 - X_t}{\beta_1 - \beta_t} - \nabla \log h_t^\theta(X_t) \right\|^2 \mathrm{d}t\right] \tag{7}$$

where h_t^θ depends on b_t^θ as well as the drawn samples $(\mathbf{x}_0, \mathbf{x}_1)$. This is the case since h_t is defined as an expectation using trajectories sampled under b_t^θ with given endpoints. Therefore, assuming that, for each θ, we can compute h_t^θ *based only on* b_t^θ, we can then backprop through (7) and optimize it using any off-the-shelf algorithm.

A slightly modified (7). Even with infinite data and a neural network with sufficient capacity, the loss function defined in (7) does not converge to 0. For the purpose of numerical stability, we instead propose to modify (7) to:

$$L(\theta) := \mathbb{E}\left[\int_0^1 \left\| \frac{\mathbf{x}_1 - X_t}{\beta_1 - \beta_t} - \left(b_t^\theta + \nabla \log h_t^\theta(X_t)\right) \right\|^2 \mathrm{d}t\right] \tag{8}$$

which is clearly equivalent to (7) at the true solution of b_t. Notice that (8) bears a similar form as the popular score-matching objective employed in previous works (Song and Ermon, 2019; Song et al., 2021):

$$L(\theta) := \mathbb{E}\left[\int_0^1 \left\| \nabla \log p(\mathbf{x}_t|\mathbf{x}_0) - s^\theta(X_t, t) \right\|^2 \mathrm{d}t\right], \tag{9}$$

where the term $\frac{\mathbf{x}_1 - X_t}{\beta_1 - \beta_t}$ is akin to $\nabla \log p(\mathbf{x}_t|\mathbf{x}_0)$, while $\left(b_t^\theta + \nabla \log h_t^\theta(X_t)\right)$ corresponds to $s^\theta(X_t, t)$.

Algorithm 1 SBALIGN

Input: Aligned data $(\mathbf{x}_0^i, \mathbf{x}_1^i)_{i=1}^N$, learning rates $\gamma_\theta, \gamma_\phi$, number of iterations K

Initialize $\theta \leftarrow \theta_0$, $\phi \leftarrow \phi_0$.
for $k = 1$ **to** K **do**
  Draw a mini-batch of samples from $(\mathbf{x}_0^i, \mathbf{x}_1^i)_{i=1}^N$
  Compute empirical average of (12) with mini-batch.
  Update $\phi \leftarrow \phi - \gamma_\phi \nabla L(\theta, \phi)$
  Update $\theta \leftarrow \theta - \gamma_\theta \nabla L(\theta, \phi)$
end for

Computing h_t^θ. Inspecting h_t in (6), we see that, given $(\mathbf{x}_0, \mathbf{x}_1)$, it can be written as the conditional expectation of an indicator function:

$$h_t(\mathbf{x}) = \mathbb{P}(X_1 = \mathbf{x}_1 | X_t = \mathbf{x}) = \mathbb{E}\left[\mathbb{1}_{\{\mathbf{x}_1\}} | X_t = \mathbf{x}\right] \tag{10}$$

where the expectation is over (5). Functions of the form (10) lend itself well to computation since it solves simulating the *unconditioned* paths. Furthermore, in order to avoid overfitting on the given samples, it is customary to replace the "hard" constraint $\mathbb{1}_{\{\mathbf{x}_1\}}$ by its *smoothed* version (Zhang and Chen, 2022; Holdijk et al., 2022):

$$h_{t,\tau}(\mathbf{x}) := \mathbb{E}\left[\exp\left(-\frac{1}{2\tau}\|X_1 - \mathbf{x}_1\|^2\right) | X_t = \mathbf{x}\right]. \tag{11}$$

Here, τ is a regularization parameter that controls how much we "soften" the constraint, and we have $\lim_{\tau \to 0} h_{t,\tau} = h_t$.

Although the computation of (11) can be done via a standard application of the Feynman–Kac formula (Rogers and Williams, 2000), an altogether easier approach is to parametrize $h_{t,\tau}$ by a second neural network m^ϕ and perform alternating minimization steps on b_t^θ and m^ϕ. This choice reduces the variance in training, since it avoids the sampling of unconditional paths described by (5) (see Appendix for a detailed explanation).

Regularization. Since it is well-known that $\nabla \log h_t$ typically explodes when $t \to 1$ (Liu et al., 2023), it is important to regularize the behavior of m^ϕ for numerical stability, especially when $t \to 1$. Moreover, in practice, it is desirable to learn a drift b_t^θ that respects the data alignment *in expectation*: If $(\mathbf{x}_0, \mathbf{x}_1)$ is an input pair, then multiple runs of the SDE (5) starting from $\mathbf{x}_0$ should, on average, produce

samples that are in the proximity of $\mathbf{x}_1$. This observation implies that we should search for drifts whose corresponding h-transforms are diminishing.

A simple way to simultaneously achieve the above two requirements is to add an ℓ^2-regularization term, resulting in the loss function:

$$L(\theta, \phi) := \mathbb{E}\left[\int_0^1 \left\| \frac{\mathbf{x}_1 - X_t}{\beta_1 - \beta_t} - \left(b_t^\theta + m^\phi(X_t)\right)\right\|^2 + \lambda_t \|m^\phi(\mathbf{x}_t)\|^2 \, \mathrm{d}t\right] \tag{12}$$

where λ_t can either be constant or vary with time. The overall algorithm is depicted in Algorithm 1.

3.2 PAIRED SCHRÖDINGER BRIDGES AS PRIOR PROCESSES

Our algorithm finds solutions to SBs on aligned data by relying on samples drawn from the (optimal) coupling $\pi^\star$. This is what differentiates it from classical SBs –which instead only consider samples from $\hat{\mathbb{P}}_0$ and $\hat{\mathbb{P}}_1$– and plays a critical role in avoiding IPF-like iterates. However, SBALIGN reliance on samples from $\pi^\star$ may become a limitation, when the available information on alignments is insufficient.

If the number of pairings is limited, it is unrealistic to hope for an accurate solution to the aligned SB problem. However, the interpolation between $\hat{\mathbb{P}}_0$ and $\hat{\mathbb{P}}_1$ learned by SBALIGN can potentially be leveraged as a starting point to obtain a better reference process, which can then be used when solving a classical SB on the same marginals. In other words, the drift $b_t^{\text{aligned}}(X_t)$ learned through SBALIGN can be used *as is* to construct a data-informed alternative $\tilde{\mathbb{Q}}$ to the standard Brownian motion, defined by paths:

$$\tilde{X}_t = b_t^{\text{aligned}}(\tilde{X}_t)dt + g_t dW_t$$

Intuitively, solving a standard SB problem with $\tilde{\mathbb{Q}}$ as reference is beneficial because the (imperfect) coupling of marginals learned by SBALIGN ($\tilde{\mathbb{Q}}_{01}$) is, in general, closer to the truth than $\mathbb{Q}_{01}$.

Improving reference processes through pre-training or data-dependent initialization has been previously considered in the literature. For instance, both De Bortoli et al. (2021) and Chen et al. (2022a) use a pre-trained reference process for challenging image interpolation tasks. This approach, however, relies on DSBs trained using the classical score-based generative modeling objective between a Gaussian and the data distribution. It, therefore, pre-trains the reference process on a related –but different– process, i.e., the one mapping Gaussian noise to data rather than $\hat{\mathbb{P}}_0$ to $\hat{\mathbb{P}}_1$. An alternative, proposed by Bunne et al. (2023) draws on the closed-form solution of SBs between two Gaussian distributions, which are chosen to approximate $\hat{\mathbb{P}}_0$ and $\hat{\mathbb{P}}_1$, respectively. Unlike our method, these alternatives construct prior drifts by falling back to simpler and related tasks, or approximations of the original problem. We instead propose to shape a coarse-grained description of the drift based on alignments sampled directly from $\pi_{01}^\star$.

4 EXPERIMENTS

In this section, we evaluate SBALIGN in different settings involving 2-dimensional synthetic datasets, the task of reconstructing cellular differentiation processes, as well as predicting the conformation of a protein structure and its ligand formalized as rigid protein docking problem.

4.1 SYNTHETIC EXPERIMENTS

We run our algorithm on two synthetic datasets (Figures in Appendix), and compare the results with classic diffusion Schrödinger bridge models, i.e., the forward-backward SB formulation proposed by Chen et al. (2022a), herein referred to as FBSB. We equip the baseline with prior knowledge, as elaborated below, to further challenge SBALIGN.

Moon dataset. The first synthetic dataset (Fig. 2a-c) consists of two distributions, each supported on two semi-circles ($\hat{\mathbb{P}}_0$ drawn in *blue* and $\hat{\mathbb{P}}_1$ in *red*). $\hat{\mathbb{P}}_1$ was obtained from $\hat{\mathbb{P}}_0$ by applying a clockwise rotation around the center, i.e., by making points in the upper blue arm correspond to those in the right red one. This transformation is clearly not the most likely one under the assumption of Brownian motion of particles and should therefore not be found as the solution of a classical SB problem. This is confirmed by FBSB trajectories (Fig. 2a), which tend to map points to their closest neighbor in $\hat{\mathbb{P}}_1$ (e.g., some points in the upper arm of $\hat{\mathbb{P}}_0$ are brought towards the left rather than towards the right). While being a minimizer of (SB), such a solution completely disregards our prior knowledge on the alignment of particles, which is instead reliably reproduced by the dynamics learned by SBALIGN (Fig. 2c).

One way of encoding this additional information on the nature of the process is to modify $\mathbb{Q}_t$ by introducing a clockwise radial drift, which describes the prior tangential velocity of particles moving circularly around the center. Solving the classical SB with this updated reference process indeed generates trajectories that respect most alignments (Fig. 2b), but requires a hand-crafted expression of the drift that is only possible in very simple cases.

T dataset. In most real-world applications, it is very difficult to define an appropriate reference process $\mathbb{Q}_t$, which respects the known alignment without excessively distorting the trajectories from a solution to (SB). This is already visible in simple examples like (Fig. 2d-f), in which the value

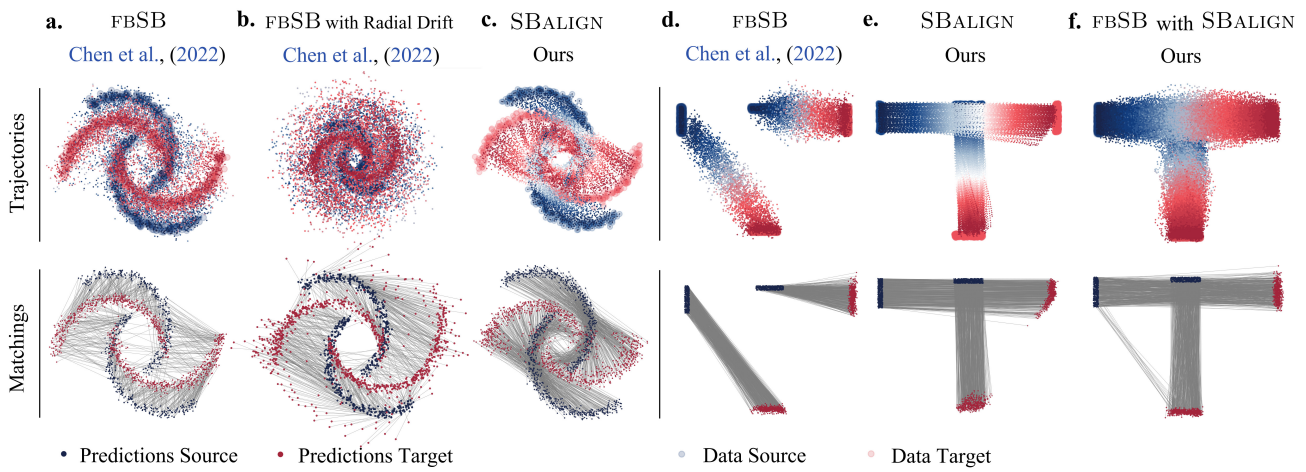

Figure 2: Experimental results on the Moon dataset (**a-c**) and T-dataset (**d-f**). The top row shows the trajectory sampled using the learned drift, and the bottom row shows the matching based on the learnt drift. Compared to other baselines, SBALIGN is able to learn an appropriate drift respecting the true alignment. (**f**) further showcases the utility of SBALIGN's learnt drift as a suitable reference process to improve other training methods.

of good candidate prior drifts at a specific location needs to vary wildly in time. In this dataset, $\hat{\mathbb{P}}_0$ and $\hat{\mathbb{P}}_1$ are both bi-modal distributions, each supported on two of the four extremes of an imaginary T-shaped area. We target alignments that connect the two arms of the T as well as the top cloud with the bottom one. We succeed in learning them with SBALIGN (Fig. 2e) but unsurprisingly fail when using the baseline FBSB (Fig. 2d) with a Brownian motion prior.

In this case, however, attempts at designing a better reference drift for FBSB must take into account the additional constraint that the horizontal and vertical particle trajectories intersect (see Fig. 2e), i.e., they cross the same area at times t_h and t_v (with $t_h > t_v$). This implies that the drift b_t, which initially points downwards (when $t < t_v$), should swiftly turn rightwards (for $t > t_h$). Setting imprecise values for one of t_h and t_v when defining custom reference drifts for classical SBs would hence not lead to the desired result and, worse, would actively disturb the flow of the other particle group.

As described in § 3.2, in presence of hard-to-capture requirements on the reference drift, the use of SBALIGN offers a remarkably easy and efficient way of learning a parameterization of it. For instance, when using the drift obtained by SBALIGN as reference drift for the computation of the SB baseline (FBSB), we find the desired alignments (Fig. 2f).

4.2 CELL DIFFERENTIATION

Biological processes are determined through heterogeneous responses of single cells to external stimuli, i.e., developmental factors or drugs. Understanding and predicting the dynamics of single cells subject to a stimulus is thus crucial to enhance our understanding of health and disease and the focus of this task. Most single-cell high-throughput technologies are destructive assays —i.e., they destroy cells upon measurement— allowing us to only measure *unaligned* snapshots of the evolving cell population. Recent methods address this limitation by proposing (lower-throughput) technologies that keep cells alive after transcriptome profiling (Chen et al., 2022b) or that genetically tag cells to obtain a clonal trace upon cell division (Weinreb et al., 2020).

Dataset. To showcase SBALIGN's ability to make use of such (partial) alignments when inferring cell differentiation processes, we take advantage of the genetic barcoding system developed by Weinreb et al. (2020). With a focus on fate determination in hematopoiesis, Weinreb et al. (2020) use expressed DNA barcodes to clonally trace single-cell transcriptomes over time. The dataset consists of two snapshots: the first, recorded on day 2, when most cells are still undifferentiated (see Fig. 4a), and a second, on day 4, comprising many different mature cell types (see Fig. 4b). Using SBALIGN as well as the baseline FSSB, we attempt to reconstruct cell evolution between day 2 and day 4, all while capturing the heterogeneity of emerging cell types. More details on the dataset can be found in the Appendix.

Baselines. We benchmark SBALIGN against previous DSBs such as (Chen et al., 2022a, FBSB) and also use it to learn a prior reference process. Cell division processes and subsequently the propagation of the barcodes are naturally very noisy. While this genetic annotation provides some form of assignment, it does not capture the full developmental process. We thus test SBALIGN in a setting where it learns a prior from such partial alignments and, plugged into FBSB, is fine-tuned on the full dataset.

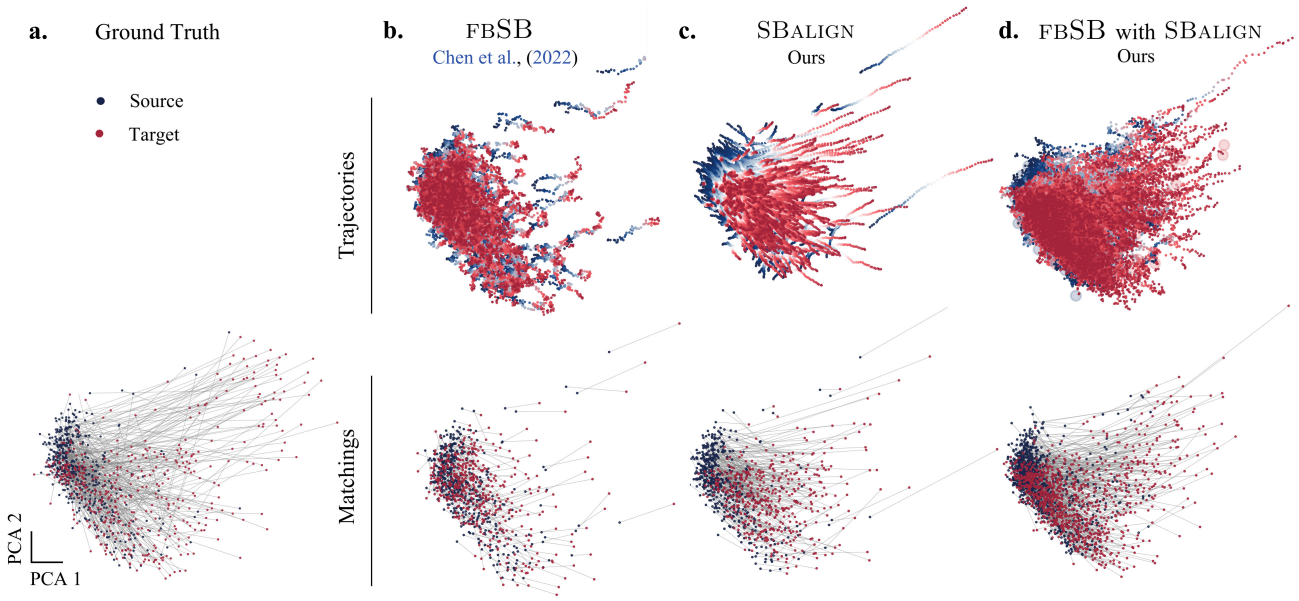

Figure 3: Cell differentiation trajectories based on (**a**) the ground truth and (**b-d**) learned drifts. SBALIGN is able to learn an appropriate drift underlying the true differentiation process while respecting the alignment. (**d**) Using the learned drift from SBALIGN as a reference process helps improve the drift learned by other training methods.

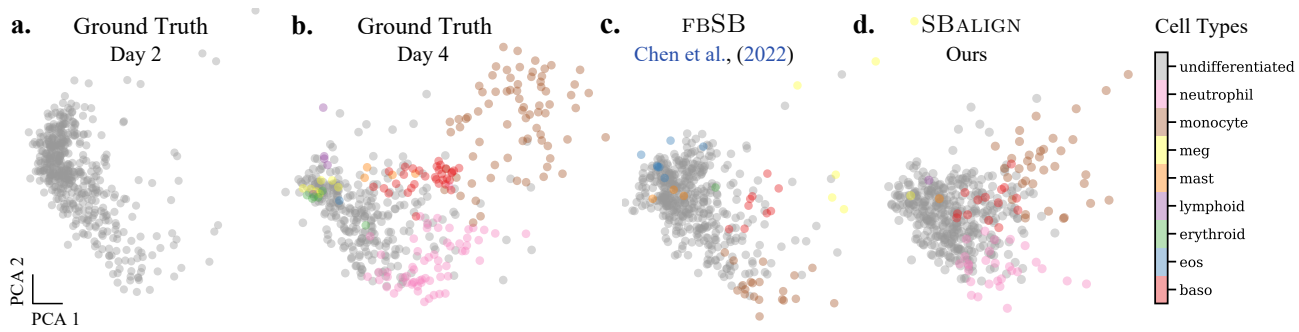

Figure 4: Cell type prediction on the differentiation dataset. All distributions are plotted on the first two principal components. **a-b:** Ground truth cell types on day 2 and day 4 respectively. **c-d:** FBSB and SBALIGN cell type predictions on day 4. SBALIGN is able to better model the underlying differentiation processes and capture the diversity in cell types.

Evaluation metrics. To assess the performance of SBALIGN and the baselines, we monitor several metrics, which include distributional distances, i.e., MMD (Gretton et al., 2012) and W_ϵ (Cuturi, 2013), as well as average (perturbation scores), i.e., ℓ_2(PS) (Bunne et al., 2022a) and RMSD. Moreover, we also train a simple neural network-based classifier to annotate the cell type on day 4 and we report the accuracy of the predicted vs. actual cell type for all the models (more details in the Appendix).

Results. SBALIGN finds matchings between cell states on days 2 and 4 (Fig. 3c, bottom) which resemble the observed ones (Fig. 3a) but also reconstructs the entire evolution path of transcriptomic profiles (Fig. 3c, top). It outperforms the baseline FBSB (Tab. 1) in all metrics: Remarkably, our method exceeds the performances of the baseline also on distributional metrics and not uniquely on alignment-based ones. We also leverage SBALIGN predictions to recover the type of cells at the end of the differentiation process (Fig. 4d): We train a classifier on differentiated cells observed on day 4, and subsequently classify our predictions. While capturing the overall differentiation trend, SBALIGN (as well as FBSB) struggles to isolate rare cell types. Lastly, we employ SBALIGN to learn a prior process from noisy alignments based on genetic barcode annotations. When using this reference process within FBSB, we learn an SB which compensates for inaccuracies stemming from the stochastic nature of cell division and barcode redistribution and which achieves better scores on distributional metrics (Tab. 1). Additional results can be found in the Appendix.

Table 1: **Cell differentiation prediction results.** Means and standard deviations (in parentheses) of distributional metrics (MMD, W_ϵ), alignment-based metrics (ℓ_2, RMSD), and cell type classification accuracy.

Methods	Cell Differentiation				
	MMD ↓	W_ε ↓	ℓ_2(PS) ↓	RMSD ↓	Class. Acc. ↑
FBSB	1.55e-2 (0.03e-2)	12.50 (0.04)	4.08 (0.04)	9.64e-1 (0.02e-1)	56.2% (0.7%)
FBSB with SBALIGN	5.31e-3 (0.25e-3)	10.54 (0.08)	0.99 (0.12)	9.85e-1 (0.07e-1)	47.0% (1.5%)
SBALIGN	1.07e-2 (0.01e-2)	11.11 (0.02)	1.24 (0.02)	9.21e-1 (0.01e-1)	56.3% (0.7%)

4.3 PROTEIN DOCKING

Proteins are dynamic, flexible biomolecules that form complexes upon interaction with other biomolecules. This is a central step in many biological processes, namely signal transduction, DNA replication, and repair. The formation of complexes is guided by appropriate energetics, best orienting the participating proteins relative to each other, along with a dynamic alteration in structure (conformational changes). Modelling this process is thus a central problem in biology and could allow one to engineer protein interactions for desired responses. In (*computational*) protein docking, the goal is to predict the 3D structure of the bound (docked) state of a protein pair, given the unbound states of the corresponding proteins. These proteins are denoted (arbitrarily) as the ligand and receptor respectively.

A comprehensive treatment of the protein docking problem is still elusive, owing to the lack of high-quality large datasets comprising 3D structures of participating proteins in the unbound and bound states. We tackle, instead, two related subproblems: (i) prediction of conformational changes between unbound and bound states of proteins and (ii) identification of the best orientation between interacting proteins, modeled as rigid bodies. This separation into related subproblems was also adopted in (Dominguez et al., 2003), one of the earliest works for the full protein docking problem.

4.3.1 Conformational Changes in Proteins

In this task, we are interested in predicting the 3D structure of the bound state of a protein, given the 3D structure in the unbound state. While it is possible to frame this problem as a (*conditional*) point cloud translation, an approach using Schrödinger bridges is more natural since it leverages the flexibility of proteins and accounts for the underlying stochasticity in the conformational change process.

Dataset. The task of modeling conformational changes starting from a given protein structure is largely unexplored, mainly due to the lack of high-quality large datasets. Here we utilize the recently proposed D3PM dataset (Peng et al., 2022) that provides protein structures before (*apo*) and after (*holo*) binding, covering various types of protein motions. We generate samples by collecting Protein Data Bank (PDB) entries containing the same protein bound to different biomolecules and applying additional quality-control criteria. We only focus on protein pairs where the provided Root Mean Square Deviation (RMSD) of the Cα carbon atoms between unbound and bound 3D structures is > 3.0Å, which amounts to 2370 examples in the D3PM dataset.

For each pair of structures, we first identify common residues and compute the RMSD between Cα carbon atoms of the common residues after superposition using the Kabsch (Kabsch, 1976) algorithm. The pair is accepted only if the relative error between the computed and provided Cα RMSD is less than 0.1. The rationale here is to only retain examples where we can reconstruct the provided RMSD values. The resulting dataset has 1591 examples, which is then divided into a train/valid/test split of 1291/150/150 examples respectively (see Appendix).

Baselines. Since the goal of the task is to predict 3D structures, our model must satisfy the relevant SE(3) symmetries of rotation and translations. To this end, we evaluate SBALIGN against the EGNN model (Satorras et al., 2021), which satisfies the SE(3) symmetries and is a popular architecture used in many point-cloud transformation tasks (Satorras et al., 2021; Hoogeboom et al., 2022).

Table 2: **Conformational changes results.** RMSD between predicted and true structures in the bound state. The first term (parentheses) refers to the number of poses sampled, and the second term refers to the number of simulation steps.

Methods	D3PM Test Set					
	RMSD (Å)			% RMSD (Å) $< \tau$		
	Median	Mean	Std	$\tau = 2$	$\tau = 5$	$\tau = 10$
EGNN	19.99	21.37	8.21	1%	1%	3%
SBALIGN (10, 10)	3.80	4.98	3.95	0%	69%	93%
SBALIGN (10, 100)	3.81	5.02	3.96	0%	70%	93%

Results. To evaluate our model, we report (Tab. 2) summary statistics of the RMSD between the Cα carbon atoms of the predicted structure and the ground truth, and the fraction of predictions with RMSD values $< 2.0, 5.0$ and 10.0Å. SBALIGN outperforms EGNN by a large margin and is able to predict almost 70% examples with an RMSD$<$ 5Å. One of the drawbacks attributed to diffusion models is their slow sampling speed, owing to multiple function calls to a neural network. Remarkably, our model is able to achieve impressive performance with just 10 steps of simulation. We leave it to future work to explore the tradeoff between sampling speed and quality of the predicted conformations.

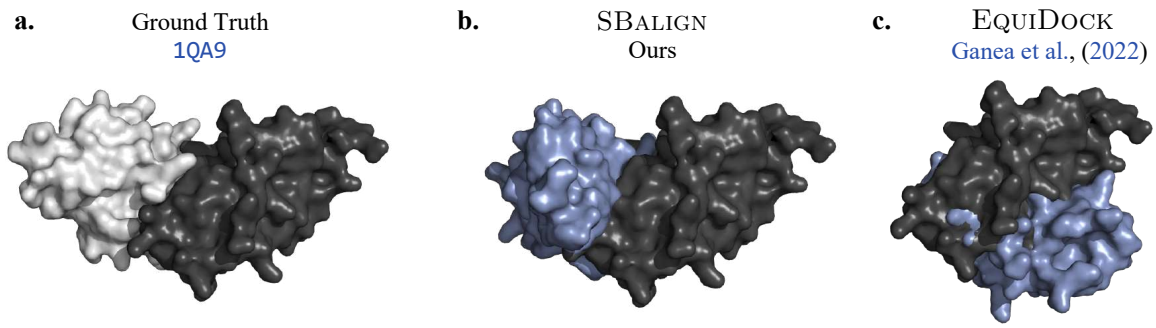

Figure 5: Ground truth and predicted bound structures for the complex with PDB ID: 1QA9. SBALIGN is able to find the true binding interface compared to EQUIDOCK

Table 3: **Rigid docking results.** Complex and interface RMSD between predicted and true bound structures (after Kabsch alignment). Comparison with values reported in (Ganea et al., 2022) can be found in the Appendix.

	DB5.5 Test Set					
	Complex RMSD			Interface RMSD		
Methods	Median	Mean	Std	Median	Mean	Std
EQUIDOCK	14.12	14.73	5.31	11.97	13.23	4.93
SBALIGN	6.59	6.69	2.04	7.69	8.11	2.39

4.3.2 Rigid Protein Docking

In this task, we want to identify the best relative orientation between the two proteins, modeled as rigid bodies.

Experimental setup. Our setup follows a similar convention as EQUIDOCK (Ganea et al., 2022). To summarize, the unbound structure of the ligand is derived by applying a random rotation and translation to the corresponding bound structure, while the receptor is held fixed w.l.o.g. Applying a different rotation and translation to each ligand can however result in a different Brownian bridge for each complex, resulting in limited meaningful signal for learning b_t^θ. To avoid this, we sample a rotation and translation at the start of training and apply the same rotation and translation to all complexes across training, validation, and testing (more details in the Appendix).

Dataset. We evaluate our method on the DB5.5 dataset (Vreven et al., 2015) which is a standard choice for protein-protein docking but contains only 253 complexes. We use the same splits as EquiDock (Ganea et al., 2022) –containing 203/25/25 complexes in the training, validation and test sets respectively– and show the results in Tab. 3. For ligands in the test set, we generate the corresponding unbound versions by applying the rotation and translation sampled during training. We compare our method to EQUIDOCK as well as to traditional docking software (see Appendix for details).

Evaluation metrics. We report two metrics, Complex RMSD and Interface RMSD. Following (Ganea et al., 2022), we first superimpose the ground truth and the predicted complex structures using the Kabsch algorithm (Kabsch, 1976), and then calculate Complex RMSD. A similar procedure is used for computing Interface RMSD, but only using the residues from the two proteins that are within 8 Å of each other (see Appendix for more details).

Results. SBALIGN considerably outperforms EQUIDOCK across all metrics (Table 3). An example of docked structures, in direct comparison with EQUIDOCK is displayed in Fig. 5, with more visualizations & results in the Appendix.

Future outlook. In this section, we presented a proof of concept application of SBALIGN for the subproblems associated with the protein docking task. While SBALIGN provides a principled method to model conformational changes, our setup for rigid protein docking is limited by utilizing the same rotation and translation across training and testing. A combination of SBALIGN for conformational change modeling, with more recent methods for rigid-protein docking (Ketata et al., 2023) can provide a complete solution for the protein docking task, which we leave to future work.

5 CONCLUSION

In this paper, we propose a new framework to tackle the interpolation task with aligned data via diffusion Schrödinger bridges. Our central contribution is a novel algorithmic framework derived from the Schrödinger bridge theory and Doob's h-transform. Via a combination of the two notions, we derive novel loss functions which, unlike all prior methods for solving diffusion Schrödinger bridges, do not rely on the iterative proportional fitting procedure and are hence numerically stable. We verify our proposed algorithm on various synthetic and real-world tasks and demonstrate noticeable improvement over the previous state-of-the-art, thereby substantiating the claim that data alignment is a highly relevant feature that warrants further research.

Acknowledgements

This publication was supported by the NCCR Catalysis (grant number 180544), a National Centre of Competence in Research funded by the Swiss National Science Foundation as well as the European Union's Horizon 2020 research and innovation programme 826121. We thank Caroline Uhler for introducing us to the dataset by Weinreb et al. (2020), which was instrumental in this research.

References

Charlotte Bunne, Stefan G Stark, Gabriele Gut, Jacobo Sarabia del Castillo, Kjong-Van Lehmann, Lucas Pelkmans, Andreas Krause, and Gunnar Ratsch. Learning Single-Cell Perturbation Responses using Neural Optimal Transport. *bioRxiv*, 2021.

Charlotte Bunne, Andreas Krause, and Marco Cuturi. Supervised Training of Conditional Monge Maps. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2022a.

Charlotte Bunne, Laetitia Meng-Papaxanthos, Andreas Krause, and Marco Cuturi. Proximal Optimal Transport Modeling of Population Dynamics. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 25, 2022b.

Charlotte Bunne, Ya-Ping Hsieh, Marci Cuturi, and Andreas Krause. The Schrödinger Bridge between Gaussian Measures has a Closed Form. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2023.

Tianrong Chen, Guan-Horng Liu, and Evangelos A Theodorou. Likelihood Training of Schrödinger Bridge using Forward-Backward SDEs Theory. In *International Conference on Learning Representations (ICLR)*, 2022a.

Wanze Chen, Orane Guillaume-Gentil, Pernille Yde Rainer, Christoph G Gäbelein, Wouter Saelens, Vincent Gardeux, Amanda Klaeger, Riccardo Dainese, Magda Zachara, Tomaso Zambelli, et al. Live-seq enables temporal transcriptomic recording of single cells. *Nature*, 608, 2022b.

Yongxin Chen, Tryphon T Georgiou, and Michele Pavon. Optimal Transport in Systems and Control. *Annual Review of Control, Robotics, and Autonomous Systems*, 4, 2021a.

Yongxin Chen, Tryphon T Georgiou, and Michele Pavon. Stochastic Control Liaisons: Richard Sinkhorn Meets Gaspard Monge on a Schrödinger Bridge. *SIAM Review*, 63(2), 2021b.

Gabriele Corso, Hannes Stärk, Bowen Jing, Regina Barzilay, and Tommi Jaakkola. Diffusion Steps, Twists, and Turns for Molecular Docking. In *International Conference on Learning Representations (ICLR)*, 2023.

Marco Cuturi. Sinkhorn Distances: Lightspeed Computation of Optimal Transport. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 26, 2013.

Valentin De Bortoli, James Thornton, Jeremy Heng, and Arnaud Doucet. Diffusion Schrödinger Bridge with Applications to Score-Based Generative Modeling. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 35, 2021.

Cyril Dominguez, Rolf Boelens, and Alexandre MJJ Bonvin. Haddock: a protein- protein docking approach based on biochemical or biophysical information. *Journal of the American Chemical Society*, 125(7):1731–1737, 2003.

Joseph Doob. *Classical Potential Theory and Its Probabilistic Counterpart*, volume 549. Springer, 1984.

Mike Fisher, Jorge Nocedal, Yannick Trémolet, and Stephen J Wright. Data assimilation in weather forecasting: a case study in PDE-constrained optimization. *Optimization and Engineering*, 10(3), 2009.

Robert Fortet. Résolution d'un systeme d'équations de M. Schrödinger. *J. Math. Pure Appl. IX*, 1, 1940.

Octavian-Eugen Ganea, Xinyuan Huang, Charlotte Bunne, Yatao Bian, Regina Barzilay, Tommi S. Jaakkola, and Andreas Krause. Independent SE(3)-Equivariant Models for End-to-End Rigid Protein Docking. In *International Conference on Learning Representations (ICLR)*, 2022.

Arthur Gretton, Karsten M Borgwardt, Malte J Rasch, Bernhard Schölkopf, and Alexander Smola. A Kernel Two-Sample Test. *Journal of Machine Learning Research*, 13, 2012.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising Diffusion Probabilistic Models. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.

Lars Holdijk, Yuanqi Du, Ferry Hooft, Priyank Jaini, Bernd Ensing, and Max Welling. Path Integral Stochastic Optimal Control for Sampling Transition Paths. *arXiv preprint arXiv:2207.02149*, 2022.

Emiel Hoogeboom, Vıctor Garcia Satorras, Clément Vignac, and Max Welling. Equivariant diffusion for molecule generation in 3d. In *International Conference on Machine Learning*, pages 8867–8887. PMLR, 2022.

Wolfgang Kabsch. A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography*, 32(5), 1976.

Mohamed Amine Ketata, Cedrik Laue, Ruslan Mammadov, Hannes Stärk, Menghua Wu, Gabriele Corso, Céline Marquet, Regina Barzilay, and Tommi S Jaakkola. Diffdock-pp: Rigid protein-protein docking with diffusion models. *arXiv preprint arXiv:2304.03889*, 2023.

Solomon Kullback. Probability densities with given marginals. *The Annals of Mathematical Statistics*, 39 (4):1236–1243, 1968.

Christian Léonard. A survey of the Schrödinger problem and some of its connections with optimal transport. *arXiv preprint arXiv:1308.0215*, 2013.

Guan-Horng Liu, Tianrong Chen, Oswin So, and Evangelos Theodorou. Deep Generalized Schrödinger Bridge. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2022a.

Guan-Horng Liu, Tianrong Chen, Oswin So, and Evangelos A Theodorou. Deep Generalized Schrödinger Bridge. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2022b.

Xingchao Liu, Lemeng Wu, Mao Ye, and qiang Liu. Learning Diffusion Bridges on Constrained Domains. *International Conference on Learning Representations (ICLR)*, 2023.

Mohammad Lotfollahi, F Alexander Wolf, and Fabian J Theis. scGen predicts single-cell perturbation responses. *Nature Methods*, 16(8), 2019.

Roger Mansuy and Marc Yor. *Aspects of Brownian motion*. Springer Science & Business Media, 2008.

Stefano Peluchetti. Diffusion bridge mixture transports, schrödinger bridge problems and generative modeling, 2023.

Cheng Peng, Xinben Zhang, Zhijian Xu, Zhaoqiang Chen, Yanqing Yang, Tingting Cai, and Weiliang Zhu. D3pm: a comprehensive database for protein motions ranging from residue to domain. *BMC bioinformatics*, 23(1):1–11, 2022.

Gabriel Peyré and Marco Cuturi. Computational Optimal Transport. *Foundations and Trends in Machine Learning*, 11(5-6), 2019.

L Chris G Rogers and David Williams. *Diffusions, Markov Processes and Martingales: Volume 2, Itô Calculus*, volume 2. Cambridge University Press, 2000.

Victor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E(n)-equivariant graph neural networks. *arXiv preprint arXiv:2102.09844*, 2021.

Moritz Schauer, Frank van der Meulen, and Harry van Zanten. Guided proposals for simulating multi-dimensional diffusion bridges. *Bernoulli*, 23(4A), 2017.

Geoffrey Schiebinger, Jian Shu, Marcin Tabaka, Brian Cleary, Vidya Subramanian, Aryeh Solomon, Joshua Gould, Siyan Liu, Stacie Lin, Peter Berube, et al. Optimal-Transport Analysis of Single-Cell Gene Expression Identifies Developmental Trajectories in Reprogramming. *Cell*, 176(4), 2019.

Erwin Schrödinger. *Über die Umkehrung der Naturgesetze*. Verlag der Akademie der Wissenschaften in Kommission bei Walter De Gruyter u. Company, 1931.

Yang Song and Stefano Ermon. Generative Modeling by Estimating Gradients of the Data Distribution. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2019.

Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-Based Generative Modeling through Stochastic Differential Equations. In *International Conference on Learning Representations (ICLR)*, volume 9, 2021.

Alexander Tong, Jessie Huang, Guy Wolf, David Van Dijk, and Smita Krishnaswamy. TrajectoryNet: A Dynamic Optimal Transport Network for Modeling Cellular Dynamics. In *International Conference on Machine Learning (ICML)*, 2020.

Alexander Tong, Nikolay Malkin, Guillaume Huguet, Yanlei Zhang, Jarrid Rector-Brooks, Kilian Fatras, Guy Wolf, and Yoshua Bengio. Conditional Flow Matching: Simulation-Free Dynamic Optimal Transport. *arXiv preprint arXiv:2302.00482*, 2023.

Tomer Tsaban, Julia K Varga, Orly Avraham, Ziv Ben-Aharon, Alisa Khramushin, and Ora Schueler-Furman. Harnessing protein folding neural networks for peptide-protein docking. *Nature Communications*, 13(1):176, 2022.

Francisco Vargas, Pierre Thodoroff, Neil D Lawrence, and Austen Lamacraft. Solving Schrödinger Bridges via Maximum Likelihood. *Entropy*, 23(9), 2021.

Thom Vreven, Iain H Moal, Anna Vangone, Brian G Pierce, Panagiotis L Kastritis, Mieczyslaw Torchala, Raphael Chaleil, Brian Jiménez-García, Paul A Bates, Juan Fernandez-Recio, et al. Updates to the integrated protein–protein interaction benchmarks: docking benchmark version 5 and affinity benchmark version 2. *Journal of Molecular Biology*, 427(19), 2015.

Caleb Weinreb, Alejo Rodriguez-Fraticelli, Fernando D Camargo, and Allon M Klein. Lineage tracing on transcriptional landscapes links state to fate during differentiation. *Science*, 367, 2020.

Qinsheng Zhang and Yongxin Chen. Path Integral Sampler: A Stochastic Control Approach For Sampling. In *International Conference on Learning Representations (ICLR)*, 2022.