Deep Momentum Multi-Marginal Schrödinger Bridge

Anonymous Authors[1]

Abstract

Reconstructing population dynamics using only samples from distributions at coarse time intervals is a crucial challenge. Recent data-driven approaches such as flow-based models or Schrödinger Bridge models have demonstrated appealing performance, yet the inferred sample trajectories either fail to account for the underlying stochasticity or are unnecessarily rigid. In this article, we propose <u>D</u>eep <u>M</u>omentum Multi-Marginal <u>S</u>chrödinger <u>B</u>ridge (DMSB), a novel computational framework that learns the smooth measure-valued spline for stochastic systems without violating the position marginal constraints across time. We first extend the scalable mean matching objective used in the state space SB algorithm into the phase space. We next carefully craft a multi-constraint optimization training method based on Bregman Iteration that enables effective phase space means matching training for the high-dimensional dataset. We demonstrate that the resulting training algorithm significantly outperforms baselines on both synthetic datasets and a real-world single-cell RNA sequence dataset.

1. Introduction

Understanding the population dynamics of time-evolving individuals received considerable attention recently. The associating applications appear in various scientific areas such as estimating cell dynamics (Schiebinger et al., 2019; Yang & Uhler, 2018), predicting meteorological evolution (Fisher et al., 2009) and medical healthcare statistics tracking (Manton et al., 2008). Central to the research along this line is to understand the dynamics and reactions of individuals given static snapshot samples taken at sporadic time points. The difficulty of population measurement is raised for several reasons. One is that tracking individual

for a long period of time is experimentally expensive or impossible. For example, single-cell RNA sequencing will destroy measured cells which makes tracking of cells impossible (Bunne et al., 2022). The task is to find a path measure whose time marginals are consistent with observed snapshots while certain constraints are implied.

One branch of cutting-edge works in this direction focuses on constructing an energy landscape that best aligned with empirical observations in a deterministic manner. Hashimoto et al. (2016) uses neural networks to represent such potential. TrajectoryNet (Tong et al., 2020) learns regularized Neural ODE (Chen et al., 2018a) to encode potential landscape. JKOnet (Bunne et al., 2022) leverages the special structure of Input Convex Neural Network and integrates it with JKO scheme (Jordan et al., 1998) to represent time-*invariant* potential energy.

Notably, in the aforementioned work, the trajectory of samples is represented in a deterministic way. Another area of research is focused on adding stochasticity to trajectory modeling, which occurs naturally in many systems (such as biological systems (Kolomgorov et al., 1937)). Koshizuka & Sato (2022); Chizat et al. (2022) employ Schrödinger Bridge (SB), an entropy regularized optimal transport model, to determine the most likely evolution of samples between two marginal distributions when the samples are also affected by environmental stochasticity. Yet, these approaches scale poorly w.r.t. the state dimension due to specialized neural network architectures and computational frameworks.

Recently, Diffusion SB (DSB) (De Bortoli et al., 2021; Chen* et al., 2021) has been applied to the high-dimensional generative modeling tasks as a generalization of Diffusion Models (Song et al., 2020). By leveraging the power of deep neural networks and tailored optimization scheme, DSB is able to simulate SB with significantly less computation complexity compared with aforementioned work thus allowing high-fidelity image data generation. Despite the success of DSB in generative modeling, the application of DSB in trajectory inference is limited. Directly applying DSB to interpolate multi-marginals in a repeated, pairwise fashion will lead to the trajectory with sharp turns, which is unnatural in practice and sensitive to the missing snapshot (see fig.3). Inspired by statistical mechanics of molecular simulation (Tuckerman, 2010; Dockhorn et al., 2021;

[1]Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

Chen et al., 2019; 2018b), momentum Schrödinger Bridge in phase space (mSB) shows unique advantages to depict population trajectory. In mSB, the position variable is augmented with an additional velocity variable. The position and velocity variables are coupled with each other while the noise and the control variable are only injected into the velocity variable and boundary conditions (statistical distributions) are specified *only* in the position space. The velocity variable is an auxiliary variable designed for obtaining a smooth trajectory. The solution of mSB leads to the measure-valued splines which interpolate data measures smoothly (Chen et al., 2018b) which yields better trajectory inference performance. This is due to the fact that mSB will consider the curvature of the trajectory throughout the time but vanilla SB will only take care of pair-wise transportation costs. Because the noise is only injected into the velocity, the trajectory of individual samples will get smoother (Dockhorn et al., 2021).

Despite the advantages of mSB, earlier research (Chen et al., 2019; Haasler et al., 2021) has relied mainly on the Sinkhron algorithm, which is dimensionally cursed because of the discretization of space. In this work, we present a novel computation framework named <u>D</u>eep <u>M</u>omentum Multi-Marginal <u>S</u>chrödinger <u>B</u>ridge (DMSB) to infer trajectory for the high dimensional dataset. We first extend the computationally efficient mean matching technique (Chen* et al., 2021; De Bortoli et al., 2021) to the phase space which enables the efficient computation of mSB with two boundary distribution constraints. We further expand the technique to the scenario of many marginal constraints by presenting a novel training method, which is inspired by celebrated Bregman Iteration scheme (Bregman, 1967). It enables the capability of mean matching objective in the multi-marginal case which opens the door for the application of multi-marginal SB in high-dimensional situations. In contrast to (Dockhorn et al., 2021), statistical data on velocity is not always available in the general situation. We propose a novel Langevin sampler to sample velocity variables *without* prior knowledge. The main contributions of our work are fourfold:

1. We extend the mean matching objective to phase space allowing for scalable mSB computing.

2. We tailor the Bregman Iteration for momentum multi-marginal SB which makes it compatible with the phase space mean matching objective, thus the efficient computation is activated for high dimensional momentum multi-marginal SB.

3. We show how to overcome the challenge of sampling the velocity variable when it is not available, which enhances the applicability of our model.

4. We show the performance of DMSB on toy datasets which contains intricate bifurcations and merge. On realistic high-dimension (100-D) single-cell RNA-seq (scRNA-seq) datasets, DMSB outperforms baselines by a significant margin in terms of the quality of the generated trajectory both visually and quantitatively.

2. Preliminary

2.1. Dynamical Schrödinger Bridge problem

Dynamical Schrödinger Bridge problem has been extensively studied in the past few decades. The objective of the problem is to resolve the following optimization problem:

$$\min_{\pi \in \Pi(\rho_0, \rho_T)} D_{KL}(\pi || \xi), \qquad (1)$$

where $\pi \in \Pi(\rho_0, \rho_T)$ belongs to a set of path measure with its marginal densities at $t = 0$ and T being ρ_0 and ρ_T. ξ is the reference path measure. The optimality of the problem (1) is characterized by a set of PDEs (3).

Theorem 2.1 (Pavon & Wakolbinger (1991)). *The optimal path measure π in the problem (1) is represented by forward and backward stochastic processes*

$$d\mathbf{x}_t = [2 \nabla_\mathbf{x} \log \Psi_t] dt + \sqrt{2} d\mathbf{w}_t, \quad \mathbf{x}_0 \sim \rho_0, \qquad (2a)$$

$$d\mathbf{x}_t = [-2 \nabla_\mathbf{x} \log \widehat{\Psi}_t] dt + \sqrt{2} d\widehat{\mathbf{w}}_t, \mathbf{x}_T \sim \rho_T. \qquad (2b)$$

in which $\Psi, \widehat{\Psi} \in C^{1,2}$ are the solutions to the following coupled PDEs,

$$\frac{\partial \Psi_t}{\partial t} = -\Delta \Psi_t, \quad \frac{\partial \widehat{\Psi}_t}{\partial t} = \Delta \widehat{\Psi}_t \qquad (3)$$

s.t. $\Psi(0, \cdot) \widehat{\Psi}(0, \cdot) = \rho_0(\cdot), \ \Psi(T, \cdot) \widehat{\Psi}(T, \cdot) = \rho_T(\cdot),$

The stochastic process of SB in (2a) and (2b) are equivalent in the sense of $\forall t \in [0, T], p_t^{(2a)} \equiv p_t^{(2b)} \equiv p_t^{SB}$. Specifically, p_t^{SB} stands for the marginal distribution of SB at time t which also represents for the marginal density of stochastic process induced by Eq.2. The potential Ψ_t and $\widehat{\Psi}_t$ explicitly represent for the solution of Fokker-Plank Equation (FPE) and Hamilton–Jacobi–Bellman equation (HJB) after exponential transform (Chen* et al. (2021)) where FPE describes the evolution of samples density and HJB represents for the optimality of Eq.1. Furthermore, the marginal density also obeys a factorization of $p_t^{SB} = \Psi_t \widehat{\Psi}_t$. However, due to the coupling boundary constraints, the solutions of PDEs are hard to obtain even for low-dimensional system (Folland, 2020). Prior work has primarily used the half-bridge optimization technique, also known as iterative proportional fitting (IPF), in which one iteratively solves optimization problem with one of the boundary condition(Vargas, 2021; De Bortoli et al., 2021; Chen* et al., 2021),

$$\pi^{(d+1)} := \arg\min_{\pi \in \Pi(\cdot, \rho_1)} D_{KL}(\pi || \pi^{(d)}),$$

$$\pi^{(d+2)} := \arg\min_{\pi \in \Pi(\rho_0, \cdot)} D_{KL}(\pi || \pi^{(d+1)})$$

with initial ansatz of path measure $\pi^{(0)} := \xi$ (i.e., Chen* et al. (2021) sets ξ as Wiener process from ρ_0). By repeatedly iterating over aforementioned optimizations until the algorithm converges, the SB solution will be attained as $\pi^{SB} \equiv \pi^{(d+1)} \equiv \pi^{(d)}$(Benamou et al., 2015) when $d \to \infty$.

Dai Pra (1991) shows that the drift term in SB problem can also be interpreted as the solution Stochastic Optimal Control (SOC) problem by having $\mathbf{u}^* = 2\nabla_{\mathbf{x}} \log \Psi(t, \mathbf{x}_t)$:

$$\mathbf{u}^*(\mathbf{x}) \in \arg\min_{\mathbf{u} \in \mathcal{U}} \mathbb{E}\left[\int_0^T \frac{1}{2}\|\mathbf{u}_t\|^2 dt\right]$$

$$s.t \quad \begin{cases} d\mathbf{x}_t = \mathbf{u}_t dt + \sqrt{2}d\mathbf{w}_t \\ \mathbf{x}_0 \sim \rho_0, \quad \mathbf{x}_1 \sim \rho_T. \end{cases}$$

This formulation reveals that SB is the straightest path measure between initial and terminal distribution with noise.

2.2. Bregman Iteration for Multiple Constraints

Bregman Iteration (Bregman, 1967) is widely used to solve entropy regularized optimal transport problem (Chen et al., 2019) with multiple constraints. The algorithm can efficiently solve problems in the form of,

$$\inf_{\pi \in \mathcal{K}} KL(\pi|\xi),$$

where $\mathcal{K}$ is the intersection of multiple closed convex constraint sets $\mathcal{K}_l$: $\mathcal{K} = \cap_{l=1}^L \mathcal{K}_l$.

Bregman Projection (BP) is defined as

$$P_{\mathcal{K}_l}^{KL}(\xi) := \arg\min_{\pi \in \mathcal{K}_l} KL(\pi|\xi),$$

and d-th Bregman Iteration (BI) is recursively computing

$$\forall 0 < n \leq L, \quad \pi^{(d,n)} := P_{\mathcal{K}_l^n}^{KL}(\pi^{(d,n-1)}),$$

w.r.t one of the constraint $\mathcal{K}_l^n \in \mathcal{K}$ at iteration n. The initial condition for $(d+1)$-th Bregman iteration is,

$$\pi^{(d+1,0)} = \pi^{(d,L)}$$

Under certain conditions (see e.g., Benamou et al. (2015)), one has that $\pi^{(d,L)}$ converges to the unique solution:

$$\pi^{(d,L)} \to P_{\mathcal{K}}^{KL}(\xi) \quad \text{as} \quad d \to +\infty$$

Remark 2.2. One can notice that the BI will become the aforementioned IPF procedure solving SB problem (1) by defining $L = 2$, $\mathcal{K}_1 = \Pi(\rho_0, \cdot)$, $\mathcal{K}_2 = \Pi(\cdot, \rho_1)$.

3. Momentum Schrödinger Bridge

Notations used in following sections are listed in Table.1.

Table 1. Mathematical notation.

NOTATION	DEFINITIONS
μ	distribution of velocity and position $\mu(\mathbf{x}, \mathbf{v})$
ρ	position distribution $\rho(\mathbf{x})$
γ	velocity Distribution $\gamma(\mathbf{v})$
$\mathbf{x}$	position variable
$\mathbf{v}$	velocity variable
$\mathbf{a}$	acceleration variable
$\mathbf{m}$	concatenation of $[\mathbf{x}, \mathbf{v}]^\top$

3.1. Two Marginals Momentum Schrödinger Bridge

We first introduce the concept of momentum Schroödinger Bridge (mSB) which extends SB problem to the phase space of position and velocity. Even though for this moment boundary distributions depend on both x and v, the goal is still to find transport maps between two distributions that only depend on position $\mathbf{x}$, and velocity $\mathbf{v}$ is an auxiliary variable artificially introduced for obtaining smooth transport. Conceptually, as an entropy regularized optimal transport problem, SB tries to obtain the straightest path between two empirical probability marginals of positions $\mathbf{x}$ with additive noise, but mSB aims at finding the smooth interpolation between two empirical probability marginals of $\mathbf{x}$ (Benamou et al., 2019) conditioned on boundary velocity distributions. Such smooth measure-valued splines are obtained by minimizing $\mathbb{E}\left[\int_0^1 \frac{1}{2}\|\mathbf{a}_t\|^2 dt\right]$ (Chen et al., 2018b) while considering the augmented dynamics:

$$\begin{pmatrix} d\mathbf{x}_t \\ d\mathbf{v}_t \end{pmatrix} = \underbrace{\begin{pmatrix} \mathbf{v}_t \\ \mathbf{0} \end{pmatrix}}_{\boldsymbol{f}(\mathbf{v},t)} dt + \underbrace{\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & g_t \end{pmatrix}}_{\mathbf{g}(t)} \underbrace{\begin{pmatrix} \mathbf{0} \\ \mathbf{a}_t \end{pmatrix}}_{\mathbf{u}(t)} dt + \underbrace{\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & g_t \end{pmatrix}}_{\mathbf{g}(t)} d\mathbf{w}_t,$$

The mSB problem is stated as:

$$\min \int_0^T \int \int \frac{1}{2}\|\mathbf{u}\|_2^2 \mu d\mathbf{x} d\mathbf{v} dt,$$

$$s.t \quad \begin{cases} d\mathbf{m}_t = [\boldsymbol{f}(\mathbf{v}_t, t) + \mathbf{g}(t)\mathbf{u}_t]dt + \mathbf{g}(t)d\mathbf{w}_t \\ \mathbf{m}_0 \sim \mu(0, \mathbf{x}, \mathbf{v}), \quad \mathbf{m}_T \sim \mu(T, \mathbf{x}, \mathbf{v}), \end{cases} \tag{4}$$

where the control is defined as $\mathbf{u}_t = [\mathbf{0}, \mathbf{a}_t]^\top$.

Theorem 3.1 (Optimal forward SDE)**.** *The optimal forward policy of problem (4) is*

$$\mathbf{u}_t^* := \mathbf{Z}_t \equiv \begin{pmatrix} \mathbf{0} \\ \mathbf{z}_t \end{pmatrix} \equiv \begin{pmatrix} \mathbf{0} \\ g\nabla_{\mathbf{v}} \log \Psi_t \end{pmatrix} \tag{5}$$

and Ψ is the solution of a coupled PDE,

$$\begin{cases} \frac{\partial \Psi_t}{\partial t} = -\frac{1}{2}g^2 \Delta_{\mathbf{v}}\Psi_t - \nabla_{\mathbf{x}}\Psi_t^\top \mathbf{v} \\ \frac{\partial \widehat{\Psi}_t}{\partial t} = \frac{1}{2}g^2 \Delta_{\mathbf{v}}\widehat{\Psi}_t - \nabla_{\mathbf{x}}\widehat{\Psi}_t^\top \mathbf{v} \end{cases} \tag{6}$$

$$s.t \quad \Psi_0\widehat{\Psi}_0 = \mu(\mathbf{x}, \mathbf{v}, 0), \quad \Psi_T\widehat{\Psi}_T = \mu(\mathbf{x}, \mathbf{v}, T)$$

The corresponding forward process is represented as:

$$\mathrm{d}\mathbf{m}_t = [\boldsymbol{f} + \mathbf{g}\mathbf{u}_t^*]\,\mathrm{d}t + \mathbf{g}(t)\mathrm{d}\mathbf{w}_t \qquad (7)$$

in which $\boldsymbol{f} = [\mathbf{v}, \mathbf{0}]^\top$.

Proof. see proof in Appendix.A.2 $\qquad\qquad\square$

Due to the special structure of $\boldsymbol{g}(t)$, the mSB admits a simple expression. Thanks to time reversibility of mSB,[1] one can derive a similar result for optimal reverse time SDE driven by $\widehat{\mathbf{z}}_t$ which is presented in Appendix.A.2. However, obtaining the solutions of PDEs (6) is not trivial. As an alternative, one might create the likelihood training objective (Chen* et al., 2021) and optimize it repeatedly as part of the IPF approach to arrive at the mSB solution.

Proposition 3.2. *The log-likelihood at data point $\mathbf{m}_0$ and $\mathbf{m}_T$ can be expressed as*

$$\log p(\mathbf{m}_0, 0) = \mathbb{E}_{\mathbf{m}_t \sim (20)}\left[\log p(\mathbf{m}_T, T)\right] -$$

$$\int_0^T \mathbb{E}_{\mathbf{m}_t \sim (20)}\left[\underbrace{\frac{1}{2}\|\widehat{\mathbf{z}}_t + \mathbf{z}_t - g\nabla_{\mathbf{v}}\log p^{SB}\|^2}_{\textit{mean matching objective}}\right.$$

$$\left. + \frac{1}{2}\|\widehat{\mathbf{z}}_t\|^2 - \frac{1}{2}\|g\nabla_{\mathbf{v}}\log p^{SB} - \widehat{\mathbf{z}}_t\|^2\right]\mathrm{d}t.$$

Result for $\mathbf{m}_T$ can be obtained in the similar derivation (See Appendix A.4).

Proof. see proof in Appendix A.4. $\qquad\square$

The mSB formulation is a desirable approach that pinpoints the smooth path between two joint marginals $p(0, \mathbf{x}_0, \mathbf{v}_0)$ and $p(T, \mathbf{x}_T, \mathbf{v}_T)$ in the sense of eq.4. However, in general applications, only position distributions are available and boundary distributions of velocity $p(\mathbf{v}|\mathbf{x})$ are unknown.

In the next section, we integrate the mSB formulation with Bregman Iteration and derive a new multi-marginal Schrödinger Bridge optimization schedule in which the prior knowledge of velocity can be discarded and the smoothness is ensured implicitly.

4. Deep Momentum Multi-Marginal Schrödinger Bridge (DMSB)

We first state the problem formulation of momentum multi-marginal Schrödinger Bridge (mmmSB). Different from previous two marginals case (§3.1), we consider the scenario where $N+1$ probability measures μ_{t_i} are lying at time t_i while velocity distributions are not necessarily accessible.

---

[1]For completeness, we note that the necessary and sufficient conditions (Millet et al., 1989) for reversing an SDE with a diffusion $\sigma(t, X_t)$ requires all elements of $\nabla \cdot (\sigma\sigma^\top p_t)$ being locally integrable functions, which is trivially satisfied for our case.

Proposition 4.1 (Chen et al. (2019)). *Recalling the relationship between the dynamical SB and SOC in §2.1, problem (eq.4) with mutliple marginal constraints reads:*

$$\min \mathcal{J}(\pi) := \sum_{i=0}^{N-1} KL\left(\pi_{t_i, t_{i+1}} | \xi_{t_i, t_{i+1}}\right), \qquad (8)$$

$$s.t \quad \pi \in \mathcal{K} := \cap_{i=0}^N \mathcal{K}_{t_i}$$

where

$$\mathcal{K}_{t_0} = \left\{\int \pi_{t_0 t_1}\mathrm{d}\mathbf{m}_{t_1} = \mu_{t_0}, \int \mu_{t_0}\mathrm{d}\mathbf{v}_{t_0} = \rho_{t_0}\right\}$$

$$\mathcal{K}_{t_N} = \left\{\int \pi_{t_{N-1}, t_N}\mathrm{d}\mathbf{m}_{t_{N-1}} = \mu_{t_N}, \int \mu_{t_N}\mathrm{d}\mathbf{v}_{t_N} = \rho_{t_N}\right\}$$

$$\mathcal{K}_{t_i} = \left\{\int \pi_{t_i t_{i+1}}\mathrm{d}\mathbf{m}_{t_{i+1}} = \mu_{t_i}, \right.$$

$$\left. \int \pi_{t_{i-1}t_i}\mathrm{d}\mathbf{m}_{t_{i-1}} = \mu_{t_i}, \int \mu_{t_i}\mathrm{d}\mathbf{v}_{t_i} = \rho_{t_i}\right\}, \qquad (9)$$

and $\mathcal{K}$ is the intersection of close convex set of $\mathcal{K}_{t_i}$.

The problem described in Prop.4.1 can be solved by classical Bregman Iteration algorithm. However, due to the curse of dimensionality, the Bregman Projection cannot be applied in high-dimensional cases directly. Meanwhile, The solution of Bregman Projection under the constraint set (9) will lead to unfavorable geometric averaging which prevents the application of the scalable likelihood objective (Prop.3.2). In order to solve mmmSB in high-dimensional applications, we parameterize the forward and backward policies by a pair of neural networks. We further decouple and resemble the constraints by which it enables the scalable likelihood training objective feasible under mmmSB context.

4.1. Parameterization

Motivated by the success of prior work (Chen* et al., 2021), we parameterize path measure π by forward policy $\mathbf{z}_t^\theta$ or backward policy $\widehat{\mathbf{z}}_t^\phi$ combined with certain boundary condition. We adopt Euler–Maruyama discretization scheme in our framework and denote the time step as δ_t. In the following section, we only demonstrate the optimization for $\mathbf{z}_t^\theta$ given reference path measure driven by fixed $\widehat{\mathbf{z}}_t^\phi$. The procedure can be applied in the opposite direction easily due to the time reversibility of SB problem.
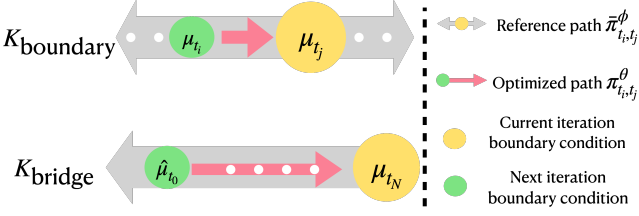
Figure 1. Optimization step w.r.t. different constraints.*Top*: The reference path measure is induced by backward policy $\widehat{\mathbf{z}}_t$ from intermediate marginal μ_{t_j} and forward policy $\mathbf{z}_t$ is optimized accordingly. *Bottom*: The reference path measure is generated from terminal marginal μ_{t_N} by $\widehat{\mathbf{z}}_t$. The forward policy $\mathbf{z}_t$ will be optimized for all $t \in [t_0, t_N]$ and ignore all the intermediate marginals. The next boundary distribution $\hat{\mu}_{t_0}$ is represented by empirical samples from reference path measure.

4.2. Decoupling and Reassembling Constraints

We decompose the constraint set (9) by setting $\mathcal{K}_{t_i} = \cap_{r=0}^{2}\mathcal{K}_{t_i}^r$ where

$$\mathcal{K}_{t_i}^0 = \left\{ \int \pi_{t_i t_{t+1}}\mathbf{dm}_{t_{i+1}} = \hat{\mu}_{t_i}, \int \hat{\mu}_{t_i}\mathbf{dv}_{t_i} = \rho_{t_i} \right\}$$

$$\mathcal{K}_{t_i}^1 = \left\{ \int \pi_{t_{i-1} t_i}\mathbf{dm}_{t_{i-1}} = \mu_{t_i}, \int \mu_{t_i}\mathbf{dv}_{t_i} = \rho_{t_i} \right\} \quad (10)$$

$$\mathcal{K}_{t_i}^2 = \left\{ \int \pi_{t_i t_{t+1}}\mathbf{dm}_{t_{i+1}} = \int \pi_{t_{i-1} t_i}\mathbf{dm}_{t_{i-1}} \right\}.$$

One can notice that the $\mathcal{K}_{t_i}^0$ and $\mathcal{K}_{t_i}^1$ share similar structure as simpler boundary marginal conditions $\mathcal{K}_{t_0}$ and $\mathcal{K}_{t_N}$ after decomposition, hence we can get rid of the notorious geometric averaging (see §4 in Chen et al. (2019)). Hereafter, we classify the constraints into two categories:

$$\mathcal{K}_{\text{boundary}} = \left\{ \cap_{i=1}^{N-1}\mathcal{K}_{t_i}^r \cap \mathcal{K}_{t_0} \cap \mathcal{K}_{t_N} | \forall r \in \{0,1\} \right\}, \quad (11a)$$

$$\mathcal{K}_{\text{bridge}} = \left\{ \cap_{i=1}^{N-1}\mathcal{K}_{t_i}^2 \right\}. \quad (11b)$$

As suggested by the aforementioned Bregman Iteration procedure, we execute optimization w.r.t. (8) while projecting the solution to constraint subset within the $\mathcal{K}_{\text{boundary}}$ and $\mathcal{K}_{\text{bridge}}$ iteratively. We will elaborate on the detailed procedure in the next sections.

4.3. Optimization in set $\mathcal{K}_{\text{boundary}}$

We first demonstrate how to optimize $\mathbf{z}_t^\theta$ w.r.t. objective function (8) given the reference path measure $\bar{\pi}^\phi$ under one subset of $\mathcal{K}_{\text{boundary}}$. The same procedure can be applied to other subsets in $\mathcal{K}_{\text{boundary}}$ due to the constructional similarity.

Proposition 4.2 (Optimality w.r.t. $\mathcal{K}_{\text{boundary}}$). *Given the reference path measure $\bar{\pi}^\phi$ driven by the policy $\widehat{\mathbf{z}}_t^\phi$ from boundary $\mu_{t_{i+1}}$ in the reverse time direction, the optimal path measure in the forward time direction of the following*

problem

$$\min_\pi \mathcal{J}(\pi) := \sum_{i=0}^{N-1} KL\left(\pi_{t_i,t_{i+1}}|\bar{\pi}_{t_i,t_{i+1}}^\phi\right),$$

$$s.t \quad \pi \in \left\{ \int \pi_{t_i t_{i+1}}\mathbf{dm}_{t_{i+1}} = \mu_{t_i}, \int \mu_{t_i}\mathbf{dv}_{t_i} = \rho_{t_i} \right\}$$

is:

$$\pi_{t_i t_{i+1}}^* = \frac{\rho_{t_i}\bar{\pi}_{t_i,t_{i+1}}^\phi}{\int \bar{\pi}_{t_i,t_{i+1}}^\phi \mathbf{dm}_{t_{i+1}}\mathbf{dv}_{t_i}}.$$

When $\pi_{t_i t_{i+1}}^\theta \equiv \pi_{t_i t_{i+1}}^*$, the following equations need to hold $\forall \bar{t} \in [t_i, t_{i+1}]$:

$$\|\mathbf{z}_t^\theta + \widehat{\mathbf{z}}_t^\phi - g\nabla_\mathbf{v}\log p_t^\phi\|_2^2 = 0, \quad (12a)$$

$$p_{t_i}(\mathbf{v}_{t_i}|\mathbf{x}_{t_i}) \equiv q_{t_i}^\phi(\mathbf{v}_{t_i}|\mathbf{x}_{t_i}). \quad (12b)$$

Proof. see proof in A.5. □

Notably, eq.12a corresponds to the aforementioned Mean Matching objective (Prop.3.2). Hence, $\mathbf{z}_t^\theta$ can be obtained from the optimization of $\mathcal{L}_{\text{MM}}$ after discretizing time:

$$\mathcal{L}_{MM} = \mathbb{E}\left[\|\delta_t\mathbf{z}_t^\theta(\mathbf{m}_{t+\delta_t}) + \delta_t\widehat{\mathbf{z}}_{t+\delta_t}^\phi(\mathbf{m}_{t+\delta_t}) - (\mathbf{m}_t + \delta_t\mathbf{z}_t^\theta - \mathbf{m}_{t+\delta_t})\|^2\right].$$

The velocity boundary condition for the reference path measure in the next BP is encoded in $p_{t_i}(\mathbf{v}_{t_i}|\mathbf{x}_{t_i}^\phi)$, but the representation of such conditional distribution is not clear. The training complexity will be prohibitively high if one uses emerging generative models to model the conditional distribution. Alternatively, we leverage the favorable property of SB to parameterize and sample from such distribution.

Proposition 4.3 (Anderson (1982); Nelson (2020)). *If π^θ and $\bar{\pi}^\phi$ shares same path measure, then*

$$\tilde{p}_{t_i}^{\theta,\phi}(\mathbf{v}_{t_i},\mathbf{x}_{t_i}) \equiv q_{t_i}^\phi(\mathbf{v}_{t_i},\mathbf{x}_{t_i}) \propto q_{t_i}^\phi(\mathbf{v}_{t_i}|\mathbf{x}_{t_i}), \quad (13)$$

where

$$\nabla_\mathbf{v}\log\tilde{p}_t^{\theta,\phi} = \left(\mathbf{z}_t^\theta + \widehat{\mathbf{z}}_t^\phi\right)/g. \quad (14)$$

Prop.4.3 suggests that one can use $p_{t_i}(\mathbf{v}_{t_i}|\mathbf{x}_{t_i}^\phi) := \tilde{p}_{t_i}^{\theta,\phi}$ to imply condition (12b) and obtain samples from such distribution by simulating Langevin dynamics using eq.14. However, forward and backward processes will not correspond to the same path measure (Song, 2022) when the discretization is coarse and one uses solely $\mathcal{L}_{\text{MM}}$. Hence, same as (Song, 2022), we leverage one extra regularization term to force the condition in Proposition.4.3 to hold:

$$\mathcal{L}_{reg} = \mathbb{E}\left[\|\delta_{t_i}\mathbf{z}_{t_i}^\theta(\mathbf{m}_{t_i}) + \delta_{t_i}\widehat{\mathbf{z}}_{t_i+\delta_{t_i}}^\phi(\mathbf{m}_{t_i+\delta_{t_i}}) - (\mathbf{m}_{t_i} + \delta_{t_i}\mathbf{z}_{t_i}^\theta - \mathbf{m}_{t_i+\delta_{t_i}})\|^2\right]$$
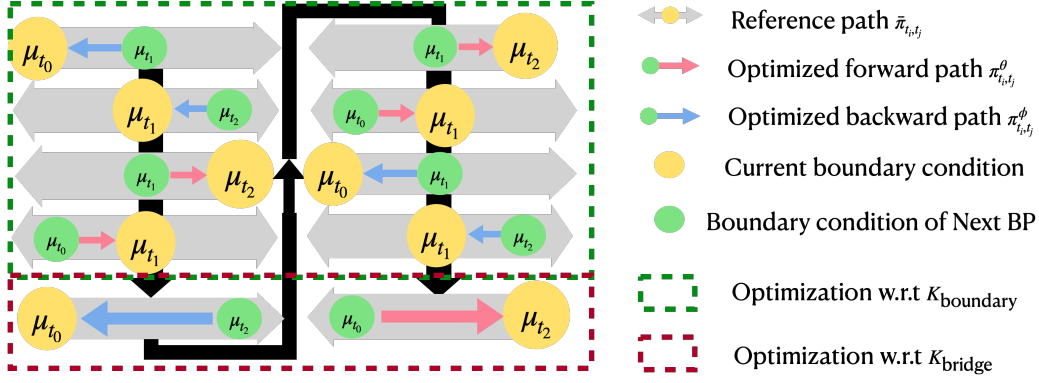
Figure 2. The sketch of training scheme for *two* Bregman Iteration. We first iterate over the constraints set in $\mathcal{K}_{\text{boundary}}$ and end up with $\mathcal{K}_{\text{bridge}}$. Due to the requirement of the mean matching objective, the training direction of the next iteration is flipped. See Appendix.3 for full training scheme.

The ultimate objective function for $t \in [t_i, t_{i+1}]$ reads:

$$\min_{\theta} \mathcal{L}(\theta; \phi) = \alpha \mathcal{L}_{\text{MM}} + (1 - \alpha)\mathcal{L}_{reg}$$

Fig.7 demonstrates the sketch of the training procedure. Given the path measure driven by the opposite direction policy $\hat{\mathbf{z}}_t^\phi$ (gray double arrow with yellow solid circle), we optimize the forward policy $\mathbf{z}_t^\theta$ which leads to the new reference path measure denoted by the green dot attached with red arrow induced by $\mathbf{z}_t^\theta$ from boundary condition. In order to obtain samples from boundary condition $\mu_{t_i}(\mathbf{x}, \mathbf{v})$, one can first sample positions from given distribution $\mathbf{x}_{t_i} \sim \rho_{t_i}(\mathbf{x})$, then sample $\mathbf{v} \sim \tilde{p}_{t_i}^{\theta, \phi}(\mathbf{v}_{t_i} | \mathbf{x}_{t_i})$ condition on $\mathbf{x}_{t_i}$ by simulating Langevin dynamics given the score function (14) of the implicit density $\tilde{p}_{t_i}^{\theta, \phi}$.

Remark 4.4. When the ground truth distributions of velocity γ_t are available, one can simply sample from γ_t instead of simulating Langevin dynamics.

4.4. Optimization in set $\mathcal{K}_{\text{bridge}}$

The formulation of optimization under $\mathcal{K}_{\text{bridge}}$ is similar to the previous section:

Proposition 4.5 (Optimality w.r.t. $\mathcal{K}_{\text{bridge}}$). *Given the reference path measure $\bar{\pi}^\phi$ driven by the policy $\hat{\mathbf{z}}_t^\phi$ from boundary μ_{t_N} in the reverse time direction, the optimal path measure in the forward time direction of the following problem*

$$\min_{\pi} \mathcal{J}(\pi) := \sum_{i=0}^{N-1} KL\left(\pi_{t_i, t_{i+1}}^\theta | \bar{\pi}_{t_i, t_{i+1}}^\phi\right),$$

$$s.t \quad \pi \in \mathcal{K}_{bridge} = \left\{\cap_{i=1}^{N-1} \mathcal{K}_{t_i}^2\right\}$$

is:

$$\pi_{t_0 t_N}^* = \frac{q_{t_0}^\phi \bar{\pi}_{t_0, t_N}^\phi}{\int \bar{\pi}_{t_0, t_N}^\phi \mathbf{dm}_{t_N} \mathbf{dv}_{t_0}}.$$

when $\pi_{t_0 t_N}^\theta \equiv \pi_{t_0 t_N}^*$, the following equations need to hold $\forall t \in [t_0, t_N]$:

$$\|\mathbf{z}_t^\theta + \hat{\mathbf{z}}_t^\phi - g\nabla_{\mathbf{v}}\log p_t^\phi\|_2^2 = 0 \tag{15a}$$

$$p_{t_0}(\mathbf{v}_{t_0}, \mathbf{x}_{t_0}) \equiv q_{t_0}^\phi(\mathbf{v}_{t_0}, \mathbf{x}_{t_0}) \tag{15b}$$

Proof. see proof in A.6 □

Conceptually, the above optimization objective with $\mathcal{K}_{\text{bridge}}$ constrain is aiming to find a *continuous* path measure which is close to nominating path measure $\bar{\pi}$ while any intermediate marginals constraints will not be considered. Eq.15a can be implied by same objective $\mathcal{L}_{\text{MM}}$ between time range $t \in [t_0, t_N]$. The boundary condition of reference path measure in the next iteration $p_{t_0}(\mathbf{v}_{t_0}, \mathbf{x}_{t_0})$ is determined by eq.15b. Fortunately, the empirical samples from this distribution are available, though the analytic representation of the distribution $q_{t_0}^\phi(\mathbf{v}_{t_0}, \mathbf{x}_{t_0})$ is unknown. Hence we can utilize these samples as empirical source from boundary distribution $q_{t_0}^\phi(\mathbf{v}_{t_0}, \mathbf{x}_{t_0})$ for the next iteration. The sketch of the optimization is in fig.7.

4.5. Training Scheme

In this section, we introduce the training scheme to traverse Bregman Iteration (BI). In a single BI, all constraints must only be iterated once. For the sake of $\mathcal{L}_{\text{MM}}$, the reference path measure should be induced in the opposite direction.

We demonstrate an example of 3 marginals case in Fig.2. The training scheme can be extended to general N marginals easily. Fig.2 consists of two BI that differs by the training order. Given the reference path measure, we first run the Bregman Projection (BP) within the subset of $\mathcal{K}_{\text{boundary}}$ sequentially and end up with the constraint $\mathcal{K}_{\text{bridge}}$. Due to the fact that the reference path in the mean matching objective has to be sampled from the opposite time direction, a single BI cannot be recursively repeated because of the conflict of sampling direction.

In order to tackle this difficulty, we reschedule the optimization order in the next BI. Since the last BP in the prior BI is optimizing the backward policy as shown in the bottom left of Fig.2, we define next reference path is driven by $\hat{\mathbf{z}}_t^\phi$ from boundary condition $\hat{\mu}_{t_N} := q_{t_N}^\theta(\mathbf{v}_{t_N}, \mathbf{x}_{t_N})$. Hence, forward policy $\mathbf{z}_t^\theta$ is optimized in the first BP for the next BI as shown in the upper right of Fig.2. The detailed algorithm is in Appendix.3.

5. Experiments

Setups: We test DMSB on 2D synthetic datasets and real-world scRNA-seq dataset (Moon et al., 2019). We parameterize $\mathbf{z}(\cdot, \cdot, \cdot; \theta)$ and $\hat{\mathbf{z}}(\cdot, \cdot, \cdot; \phi)$ with residual-based networks for all datasets. The network adopts position encoding and is trained with AdamW (Loshchilov & Hutter, 2017) on one Nvidia 3090 Ti GPU. We use constant $g(t)$ for simplicity though the framework can adopt time varying function $g(t)$. On all datasets, we set the time horizon $T = t_N = 1 \cdot N$ and interval $\delta_t = 0.01$. The SDEs are solved via Euler-Maruyama method. For scRNA-seq dataset, we split data into train and test subsets (85% and 15%). All the experiment results are simulated by all-step push forward from initial data points at time $t = t_0$.

We choose up-to-date algorithms MIOFlow (Huguet et al., 2022) and NLSB (Koshizuka & Sato, 2022) as our baselines. For both baselines, we follow the same hyperparameters listed in the paper. For MIOFlow, we report the best performance for all experiments w/ and w/o GAE embedding. For NLSB, we enlarge the size of the neural network to the best of our GPU capacity for 100-dimensional scRNA-seq dataset and report the best performance during the training. As MIOFlow and NLSB use Wasserstein distance as training objective directly, we instead use Maximum Mean Discrepancy (MMD) (Gretton et al., 2006) together with visualization as our criterion. The detail of the experiments can be found in Appendix.B.

5.1. 2D Toy Synthetic Datasets

Moving Gaussian: We first validate our algorithm on the task of inferring the path measure of a moving uni-modal Gaussian with time-varying variance. The mean of the Gaussian distribution moves from $[5, 0]$ to $[-5, 0]$ with uniform velocity on a circular path with radius 5 while the variance increase linearly w.r.t time. Fig.3 shows how learnt path (gray) evolves over time given the snapshot (blue). Compared with vanilla SB (Chen* et al., 2021), our method can learn a smoother path which is aligned with the ground truth faithfully. Additionally, DMSB is able to generate smoother trajectories in the word of individual samples due to the fact that noise is only injected into velocity variables.

Petal Synthetic Dataset: The petal dataset (Huguet et al.,
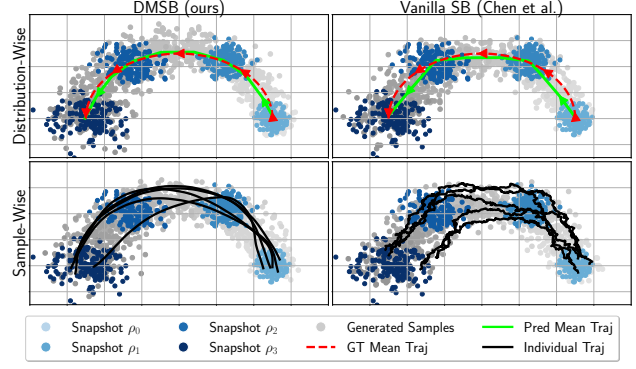

Figure 3. Validation of our DMSB model on semicircle experiment compared with vanilla SB (Chen* et al., 2021) and ground truth. The mean trajectory (Mean Traj) of Vanilla SB is straight between marginals and is sensitive to the missing snapshots since it is looking for the straightest path between each interval while DMSB is regularizing the smoothness of trajectory implicitly. GT and Pred are the abbreviations of Ground Truth and Predicted respectively.
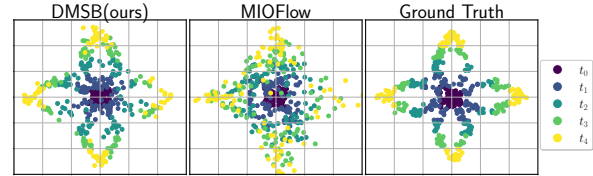

Figure 4. Comparsion with prior work and ground truth on challenging petal dataset. DMSB is able to generate trajectories whose time marginal matches ground truth faithfully and outperform prior work. Time is indicated by colors.

2022) is simple yet challenging, as it mimics natural dynamics arising in cellular differentiation, including bifurcations and merges. We compare our algorithm with MIOFlow in Fig.4. DMSB can infer trajectories aligned with ground truth distribution more faithfully compared with MIOFlow at snapshot timestep.

Gaussian Mixture Model (GMM): We demonstrate DMSB can produce reasonable velocity samples in complex four marginal GMM experiment which consists of population bifurcations and merge. We choose uni-modal Gaussian with mean $[-6, 0]$ and $[6, 0]$ at initial and terminal time steps t_0 and t_3 respectively. Four-modals GMM and eight-modals GMM are placed at intermediate time steps t_1 and t_2. Different from (Bunne et al., 2022), we are aiming to model a time-varying potential field which allows us to infer path trajectory between complicated marginals. Fig.5 demonstrates the evolution of particles in the position and velocity space. DMSB is able to learn non-trivial velocity variables which drive the position variables to interested marginals.
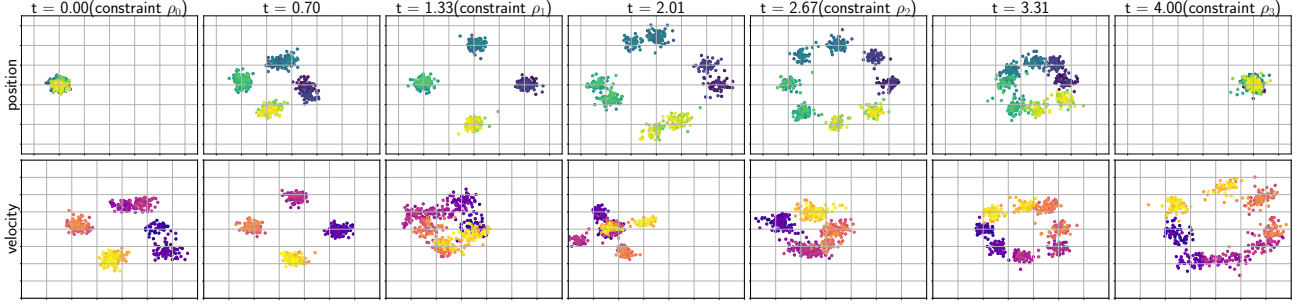
Figure 5. Validation of our DMSB model on complex GMM synthetic dataset which is mimicking the bifurcations and merge behavior. The velocity and position of the same sample correspond to the same shade level. *Upper*: Samples' evolution in the position space. *Bottom*: Samples' evolution in the velocity space.

5.2. scRNA-seq Dataset

The emerging technologies are able to identify high-resolution single-cell data which paves the way to profile the state progression of individual cells across time. However, because the cell population is eliminated after the measurement, one may only gather statistical data for single samples at particular timesteps, which neither preserves any correlations over time nor provides access to the ground truth trajectory. The diversity of embryonic stem cells after their development from embryoid bodies, which comprises mesoderm, endoderm, neuroectoderm, and neural crest in 27 days, is demonstrated by the scRNA-seq dataset. The snapshot of cells are collected between (t_0: day 0 to 3, t_1: day 6 to 9, t_2: day 12 to 15, t_3: day 18 to 21, t_4: day 24 to 27). The snapshot data are prepossessed by the quality control stage and are projected to feature space by principal component analysis (PCA). We inherit processed data from (Tong et al., 2020). We validate our algorithm on 5-dim and 100-dim PCA space to show superior performance on high-dimension problems compared with prior work. We further show that, DMSB is able to estimate better velocity distribution compared with baselines, when the ground truth is absent during training and testing.

We testify the performance of our model by computing MMD with full snapshots and when one of the snapshots is left out (LO). Let $\mathcal{D}_{-t_k}$ be the training dataset without a snapshot at time t_k. Then for experiments of $LO - t_k$, models are trained with D_{-t_k} and evaluated by the trajectory predicted from initial test samples $\mathbf{x}_0$. Table.2 summarizes the Maximum Mean Discrepancy (MMD) between estimated marginal and ground truth marginal at different time steps. The MMD value at t_k is computed between the ground truth samples and the samples predicted from the initial ground-truth samples at t_0 for each k = 1, 2, 3 and 4. Smaller MMD indicates better performance. Table.2 suggests that DMSB outperforms prior work by a large margin in high (100) dimensional scenarios. The visualization (Fig.6) in PCA space further justifies the numerical result and highlights the variety and quality of the samples pro-

duced by DMSB. The results for 5-D PCA experiment can be found in Fig.9 and Table.3.

Fig.5 demonstrates that DMSB is able to simulate reasonable velocity distribution. We further validate such property in 100-D RNAsc experiments in which the ground truth velocity is available. During the training and testing, all the models do not have access to the ground truth velocity. We evaluate the velocity of NLSB, as an SDE model, by its estimated drift term. Because MIOflow w/ GAE simulates trajectories in the latent space, we estimate the velocity by using the forward finite difference technique with discretization after mapping from the latent code to the original space. We run the experiments of 100-D and 5-D RNAsc datasets and average the discrepancy between ground truth velocity and estimated velocity over snapshot time. The numerical values are listed in the Table.4 and Table.5. The plot of velocity and position can be found in Fig.9 and Fig.10. The plot demonstrates that, even though all models are able to learn flawless trajectory, only DMSB can estimate reasonable velocity distribution. This property is held even for 100-D RNA dataset (see Fig.11,Fig.12).

6. Conclusion and Limitations

In this paper, we propose DMSB, a scalable algorithm that learns the trajectory which fits the different marginal distributions over time. We extend the mean matching objective to phase space which enables efficient mSB computing. We propose a novel training scheme to fit the mean matching objective without violating Bregman iteration which is the root of solving mmmSB problem. We demonstrate the superior result of DMSB compared with the existing algorithms.

A main limitation of this work is, the rate of convergence to the actual mmmSB has not been quantified after neural network approximations are introduced. Even though De Bortoli et al. (2021) theoretically analyzed the convergence of mean matching iteration, supporting its outstanding performance (Chen* et al., 2021), the iteration still fails to converge to the actual SB (Fernandes et al., 2021) precisely
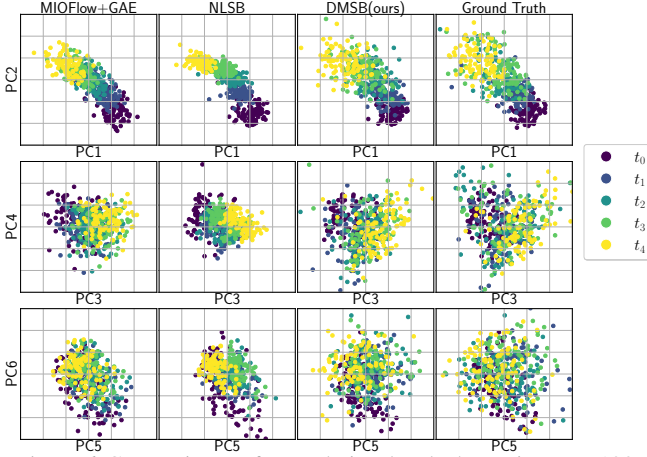
Figure 6. Comparison of population-level dynamics on 100-dimensional PCA space at the moment of observation for scRNA-seq data using MIOFlow, NLSB, and DMSB. We display the plot of the first 6 principle components (PC). Baselines can only learn the trajectory's fundamental trend, whereas DMSB can match the target marginal along the trajectory across different dimensions.

due to practical neural network estimation errors accumulating over BI. In addition, DMSB cannot simulate the process with death and birth of cells which can be potentially described as unbalanced optimal transport (Chen et al., 2022).

Table 2. Numerical result of Maximum Mean Discrepancy on 100 dimensions single-cell RNA-seq dataset and results for leaving out marginals at different observation. Numerical results. DMSB outperforms prior work by a large margin for all the timesteps and all leave-out case.

MMD ↓	t_1	t_2	t_3	t_4	Avg
NLSB	0.49	0.62	0.68	0.88	0.67
MIOFLOW	0.12	0.18	0.21	0.36	0.22
MIOFLOW+GAE	0.38	0.47	0.54	0.70	0.52
DMSB(ours)	**0.03**	**0.03**	**0.02**	**0.03**	**0.03**

MMD (LO t_1) ↓	t_1	t_2	t_3	t_4	Avg
NLSB	0.55	0.62	0.71	0.89	0.69
MIOFLOW	0.19	0.19	0.21	0.28	0.21
MIOFlow+GAE	0.38	0.42	0.49	0.63	0.48
DMSB(ours)	**0.08**	**0.01**	**0.02**	**0.02**	**0.03**

MMD (LO t_2) ↓	t_1	t_2	t_3	t_4	Avg
NLSB	0.39	0.64	0.67	0.84	0.64
MIOFlow	0.12	0.23	0.20	0.28	0.21
MIOFlow+GAE	0.49	0.82	0.70	1.06	0.77
DMSB(ours)	**0.02**	**0.08**	**0.02**	**0.02**	**0.04**

MMD (LO t_3) ↓	t_1	t_2	t_3	t_4	Avg
NLSB	0.46	0.61	0.68	0.82	0.64
MIOFlow	0.16	0.18	0.22	0.25	0.20
MIOFlow+GAE	0.36	0.44	0.51	0.62	0.48
DMSB(ours)	**0.02**	**0.02**	**0.10**	**0.02**	**0.04**

References

Anderson, B. D. Reverse-time diffusion equation models. *Stochastic Processes and their Applications*, 12(3):313–326, 1982.

Benamou, J.-D., Carlier, G., Cuturi, M., Nenna, L., and Peyré, G. Iterative bregman projections for regularized transportation problems. *SIAM Journal on Scientific Computing*, 37(2):A1111–A1138, 2015.

Benamou, J.-D., Gallouët, T. O., and Vialard, F.-X. Second-order models for optimal transport and cubic splines on the wasserstein space. *Foundations of Computational Mathematics*, 19(5):1113–1143, 2019.

Bregman, L. M. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *USSR computational mathematics and mathematical physics*, 7(3): 200–217, 1967.

Bunne, C., Papaxanthos, L., Krause, A., and Cuturi, M. Proximal optimal transport modeling of population dynamics. In *International Conference on Artificial Intelligence and Statistics*, pp. 6511–6528. PMLR, 2022.

Caluya, K. and Halder, A. Wasserstein proximal algorithms for the schrödinger bridge problem: Density control with nonlinear drift. *IEEE Transactions on Automatic Control*, 2021.

Chen*, T., Liu*, G.-H., and Theodorou, E. A. Likelihood training of schrödinger bridge using forward-backward sdes theory. *arXiv preprint arXiv:2110.11291*, 2021.

Chen, T. Q., Rubanova, Y., Bettencourt, J., and Duvenaud, D. K. Neural ordinary differential equations. In *Advances in Neural Information Processing Systems*, pp. 6572–6583, 2018a.

Chen, Y., Conforti, G., and Georgiou, T. T. Measure-valued spline curves: An optimal transport viewpoint. *SIAM Journal on Mathematical Analysis*, 50(6):5947–5968, 2018b.

Chen, Y., Conforti, G., Georgiou, T. T., and Ripani, L. Multi-marginal schrödinger bridges. In *International Conference on Geometric Science of Information*, pp. 725–732. Springer, 2019.

Chen, Y., Georgiou, T. T., and Pavon, M. The most likely evolution of diffusing and vanishing particles: Schrodinger bridges with unbalanced marginals. *SIAM Journal on Control and Optimization*, 60(4):2016–2039, 2022.

Chizat, L., Zhang, S., Heitz, M., and Schiebinger, G. Trajectory inference via mean-field langevin in path space. *arXiv preprint arXiv:2205.07146*, 2022.

Cole, J. D. On a quasi-linear parabolic equation occurring in aerodynamics. *Quarterly of applied mathematics*, 9(3): 225–236, 1951.

Dai Pra, P. A stochastic control approach to reciprocal diffusion processes. *Applied mathematics and Optimization*, 23(1):313–329, 1991.

De Bortoli, V., Thornton, J., Heng, J., and Doucet, A. Diffusion schrödinger bridge with applications to score-based generative modeling. *arXiv preprint arXiv:2106.01357*, 2021.

Dockhorn, T., Vahdat, A., and Kreis, K. Score-based generative modeling with critically-damped langevin diffusion. *arXiv preprint arXiv:2112.07068*, 2021.

Eberhard, H. The partial differential equation ut+ uux= μxx. *Communications on Pure and Applied Mathematics*, 3(3): 201–230, 1950.

Exarchos, I. and Theodorou, E. A. Stochastic optimal control via forward and backward stochastic differential equations and importance sampling. *Automatica*, 87:159–165, 2018.

Fernandes, D. L., Vargas, F., Ek, C. H., and Campbell, N. D. Shooting schrödinger's cat. In *Fourth Symposium on Advances in Approximate Bayesian Inference*, 2021.

Fisher, M., Nocedal, J., Trémolet, Y., and Wright, S. J. Data assimilation in weather forecasting: a case study in pde-constrained optimization. *Optimization and Engineering*, 10(3):409–426, 2009.

Folland, G. B. Introduction to partial differential equations. In *Introduction to Partial Differential Equations*. Princeton university press, 2020.

Gorodetsky, A. A., Karaman, S., and Marzouk, Y. M. Efficient high-dimensional stochastic optimal motion control using tensor-train decomposition. In *Robotics: Science and Systems*, 2015.

Gretton, A., Borgwardt, K., Rasch, M., Schölkopf, B., and Smola, A. A kernel method for the two-sample-problem. *Advances in neural information processing systems*, 19, 2006.

Haasler, I., Ringh, A., Chen, Y., and Karlsson, J. Multimarginal optimal transport with a tree-structured cost and the schrödinger bridge problem. *SIAM Journal on Control and Optimization*, 59(4):2428–2453, 2021.

Hashimoto, T., Gifford, D., and Jaakkola, T. Learning population-level diffusions with generative rnns. In *International Conference on Machine Learning*, pp. 2417–2426. PMLR, 2016.

Huguet, G., Magruder, D. S., Fasina, O., Tong, A., Kuchroo, M., Wolf, G., and Krishnaswamy, S. Manifold interpolating optimal-transport flows for trajectory inference. *arXiv preprint arXiv:2206.14928*, 2022.

Jordan, R., Kinderlehrer, D., and Otto, F. The variational formulation of the fokker–planck equation. *SIAM journal on mathematical analysis*, 29(1):1–17, 1998.

Kolomgorov, A., Petrovskii, I., and Piskunov, N. Study of a diffusion equation that is related to the growth of a quality of matter, and its application to a biological problem, byul. *Mosk. Gos. Univ. Ser. A Mat. Mekh*, 1(1): 26, 1937.

Koshizuka, T. and Sato, I. Neural lagrangian schr\" odinger bridge. *arXiv preprint arXiv:2204.04853*, 2022.

Loshchilov, I. and Hutter, F. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017.

Manton, K. G., Gu, X., and Lowrimore, G. R. Cohort changes in active life expectancy in the us elderly population: Experience from the 1982–2004 national long-term care survey. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 63(5): S269–S281, 2008.

Millet, A., Nualart, D., and Sanz, M. Integration by parts and time reversal for diffusion processes. *The Annals of Probability*, pp. 208–238, 1989.

Moon, K. R., van Dijk, D., Wang, Z., Gigante, S., Burkhardt, D. B., Chen, W. S., Yim, K., Elzen, A. v. d., Hirn, M. J., Coifman, R. R., et al. Visualizing structure and transitions in high-dimensional biological data. *Nature biotechnology*, 37(12):1482–1492, 2019.

Nelson, E. *Dynamical theories of Brownian motion*, volume 106. Princeton university press, 2020.

Pavon, M. and Wakolbinger, A. On free energy, stochastic control, and schrödinger processes. In *Modeling, Estimation and Control of Systems with Uncertainty*, pp. 334–348. Springer, 1991.

Schiebinger, G., Shu, J., Tabaka, M., Cleary, B., Subramanian, V., Solomon, A., Gould, J., Liu, S., Lin, S., Berube, P., et al. Optimal-transport analysis of single-cell gene expression identifies developmental trajectories in reprogramming. *Cell*, 176(4):928–943, 2019.

Song, K.-U. Applying regularized schr\" odinger-bridge-based stochastic process in generative modeling. *arXiv preprint arXiv:2208.07131*, 2022.

Song, Y., Sohl-Dickstein, J., Kingma, D. P., Kumar, A., Ermon, S., and Poole, B. Score-based generative modeling through stochastic differential equations. *arXiv preprint arXiv:2011.13456*, 2020.

Song, Y., Durkan, C., Murray, I., and Ermon, S. Maximum likelihood training of score-based diffusion models. *arXiv e-prints*, pp. arXiv–2101, 2021.

Tong, A., Huang, J., Wolf, G., Van Dijk, D., and Krishnaswamy, S. Trajectorynet: A dynamic optimal transport network for modeling cellular dynamics. In *International conference on machine learning*, pp. 9526–9536. PMLR, 2020.

Tuckerman, M. *Statistical mechanics: theory and molecular simulation*. Oxford university press, 2010.

Vargas, F. Machine-learning approaches for the empirical schrödinger bridge problem. Technical report, University of Cambridge, Computer Laboratory, 2021.

Yang, K. D. and Uhler, C. Scalable unbalanced optimal transport using generative adversarial networks. *arXiv preprint arXiv:1810.11447*, 2018.

Yong, J. and Zhou, X. Y. *Stochastic controls: Hamiltonian systems and HJB equations*, volume 43. Springer Science & Business Media, 1999.

A. Proof in §3

Before stating our proofs, we provide the assumptions used throughout the paper. These assumptions are adopted from stochastic analysis for SGM (Song et al., 2021; Yong & Zhou, 1999; Anderson, 1982), SB (Caluya & Halder, 2021), and FBSDE (Exarchos & Theodorou, 2018; Gorodetsky et al., 2015).

(i) μ_{t_i} with finite second-order moment for all t_i.

(ii) $\boldsymbol{f}$ and g are continuous functions, and $|g(t)|^2 > 0$ is uniformly lower-bounded w.r.t. t.

(iii) $\forall t \in [0, T]$, we have $\nabla_{\mathbf{v}} \log p_t(\mathbf{m}_t, t), \nabla_{\mathbf{v}} \log \Psi(\cdot, \cdot, \cdot), \nabla_{\mathbf{v}} \log \widehat{\Psi}(\cdot, \cdot, \cdot), \mathbf{Z}(\cdot, \cdot, \cdot; \theta)$, and $\widehat{\mathbf{Z}}(\cdot, \cdot, \cdot; \phi)$ Lipschitz and at most linear growth w.r.t. $\mathbf{x}$ and $\mathbf{v}$.

(iv) $\Psi, \widehat{\Psi} \in C^{1,2}$.

(v) $\exists k > 0 : p_t^{SB}(\mathbf{m}) = \mathcal{O}(\exp^{-\|\mathbf{m}\|_k^2})$ as $\mathbf{m} \to \infty$.

Assumptions (i) (ii) (iii) are standard conditions in stochastic analysis to ensure the existence-uniqueness of the SDEs; hence also appear in SGM analysis (Song et al., 2021). Assumption (iv) allows applications of Itô formula and properly defines the backward SDE in FBSDE theory. Finally, assumption (v) assures the exponential limiting behavior when performing integration by parts. w.o.l.g, we denote $\boldsymbol{f} = [\mathbf{v}, \mathbf{0}]^\mathsf{T}$.

Lemma A.1. *The optimization of (eq.4)*

$$\min \int_0^1 \int \frac{1}{2} \|\mathbf{a}\|_2^2 \mu \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{v}\mathrm{d}t, \tag{16}$$

$$s.t \begin{cases} \frac{\partial \mu(\mathbf{m}_t)}{\partial t} = -\nabla_{\mathbf{m}} \cdot \{[(\boldsymbol{f} + g\mathbf{u})] \mu\} + \frac{1}{2} g^2 \Delta_{\mathbf{v}} \mu, \\ \mu_0 = p(0, \mathbf{x}, \mathbf{v}), \quad \mu_1 = p(T, \mathbf{x}, \mathbf{v}), \end{cases} \tag{17}$$

will induce the coupled PDEs,

$$\frac{\partial \mu(\mathbf{m}_t)}{\partial t} = -\nabla_{\mathbf{m}} \cdot \left[\left(\boldsymbol{f} + g^2 \nabla_{\mathbf{v}} \phi \right) \mu \right] + \frac{1}{2} g^2 \Delta_{\mathbf{v}} \mu, \tag{18}$$

$$\frac{\partial \phi(\mathbf{m}_t)}{\partial t} = -\frac{1}{2} \|g \nabla_{\mathbf{v}} \phi\|_2^2 - \mathbf{v}^\mathsf{T} \nabla_{\mathbf{x}} \phi - \frac{1}{2} g^2 \Delta_{\mathbf{v}} \phi, \tag{19}$$

and the optimal control of the problem is

$$\mathbf{a}^* = g \nabla_{\mathbf{v}} \phi.$$

Proof. One can write the Lagrange by introducing lagrangian multiplier ϕ:

$$\begin{aligned}
\mathcal{L}(\mu, \mathbf{a}, \phi) &= \int_0^1 \int_{\mathbb{R}^n \times \mathbb{R}^n} \frac{1}{2} \|\mathbf{a}\|_2^2 \mu \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{v}\mathrm{d}t + \int_0^1 \int_{\mathbb{R}^n \times \mathbb{R}^n} \phi \frac{\partial \mu}{\partial t} \mathrm{d}\mathbf{v}\mathrm{d}\mathbf{x}\mathrm{d}t \\
&\quad + \int_0^1 \int_{\mathbb{R}^n \times \mathbb{R}^n} \phi \left\{ -\frac{1}{2} g^2 \Delta_{\mathbf{m}} \mu + \nabla_{\mathbf{m}} \cdot [(\boldsymbol{f} + g\mathbf{u}) \mu] \right\} \mathrm{d}\mathbf{v}\mathrm{d}\mathbf{x}\mathrm{d}t \\
&= \int_0^1 \int_{\mathbb{R}^n \times \mathbb{R}^n} \frac{1}{2} \|\mathbf{a}\|_2^2 \mu \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{v}\mathrm{d}t - \int_0^1 \int_{\mathbb{R}^n \times \mathbb{R}^n} \mu \frac{\partial \phi}{\partial t} \mathrm{d}\mathbf{v}\mathrm{d}\mathbf{x}\mathrm{d}t \\
&\quad + \int_0^1 \int_{\mathbb{R}^n \times \mathbb{R}^n} \phi \nabla_{\mathbf{m}} \cdot [(\boldsymbol{f} + g\mathbf{u}) \mu] - \phi \left[\frac{1}{2} g^2 \Delta_{\mathbf{m}} \mu \right] \mathrm{d}\mathbf{v}\mathrm{d}\mathbf{x}\mathrm{d}t \\
&= \int_0^1 \int_{\mathbb{R}^n \times \mathbb{R}^n} \frac{1}{2} \|\mathbf{a}\|_2^2 \mu \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{v}\mathrm{d}t - \int_0^1 \int_{\mathbb{R}^n \times \mathbb{R}^n} \mu \frac{\partial \phi}{\partial t} \mathrm{d}\mathbf{v}\mathrm{d}\mathbf{x}\mathrm{d}t \\
&\quad + \int_0^1 \int_{\mathbb{R}^n \times \mathbb{R}^n} -\nabla_{\mathbf{m}} \phi^\mathsf{T} [(\boldsymbol{f} + g\mathbf{u})] \mu - \mu \left[\frac{1}{2} g^2 \Delta_{\mathbf{m}} \phi \right] \mathrm{d}\mathbf{v}\mathrm{d}\mathbf{x}\mathrm{d}t \\
&= \int_0^1 \int_{\mathbb{R}^n \times \mathbb{R}^n} \left\{ \frac{1}{2} \|\mathbf{a}\|_2^2 - \frac{\partial \phi}{\partial t} - \mathbf{v}^\mathsf{T} \nabla_{\mathbf{x}} \phi - \frac{1}{2} g^2 \Delta_{\mathbf{v}} \phi - g \nabla_{\mathbf{v}} \phi^\mathsf{T} \mathbf{a} \right\} \mu \; \mathrm{d}\mathbf{v}\mathrm{d}\mathbf{x}\mathrm{d}t
\end{aligned}$$

By taking the minimization within the bracket, The optimal control is,

$$\mathbf{a}^* = g\nabla_\mathbf{v}\phi$$

By Plugging it back, the optimality of the aforementioned problem is presented as:

$$\frac{\partial\mu(\mathbf{m}_t)}{\partial t} = -\nabla_\mathbf{v}\cdot\left[\left(\boldsymbol{f} + g^2\nabla_\mathbf{v}\phi\right)\mu\right] + \frac{1}{2}g^2\Delta_\mathbf{v}\mu,$$

$$\frac{\partial\phi(\mathbf{m}_t)}{\partial t} = -\frac{1}{2}\|g\nabla_\mathbf{v}\phi\|_2^2 - \mathbf{v}^\mathsf{T}\nabla_\mathbf{x}\phi - \frac{1}{2}g^2\Delta_\mathbf{v}\phi,$$

□

Theorem A.2. *The optimal forward and backward processes are represented as:*

$$\mathrm{d}\mathbf{m}_t = \left[\boldsymbol{f} + g\mathbf{u}_t^{f*}\right]\mathrm{d}t + g(t)\mathrm{d}\mathbf{w}_t \quad \textit{(forward)} \tag{20}$$

$$\mathrm{d}\mathbf{m}_s = \left[\boldsymbol{f} + g\mathbf{u}_s^{b*}\right]\mathrm{d}t + g(t)\mathrm{d}\mathbf{w}_s \quad \textit{(Backward)} \tag{21}$$

in which $\boldsymbol{f} = [\mathbf{v}, \mathbf{0}]^\mathsf{T}$. *Optimal control is expressed as,*

$$\mathbf{u}_t^{f*} := \mathbf{Z}_t \equiv \begin{pmatrix} \mathbf{0} \\ \mathbf{z}_t \end{pmatrix} \equiv \begin{pmatrix} \mathbf{0} \\ g\nabla_\mathbf{v}\log\Psi_t \end{pmatrix} \tag{22}$$

$$\mathbf{u}_t^{b*} := \widehat{\mathbf{Z}}_t \equiv \begin{pmatrix} \mathbf{0} \\ \widehat{\mathbf{z}}_t \end{pmatrix} \equiv \begin{pmatrix} \mathbf{0} \\ g\nabla_\mathbf{v}\log\widehat{\Psi}_t \end{pmatrix} \tag{23}$$

where Ψ *and* $\widehat{\Psi}$ *are the solution of following PDEs,*

$$\boxed{\begin{aligned} \frac{\partial\Psi_t}{\partial t} &= -\frac{1}{2}g^2\Delta_\mathbf{v}\Psi_t - \nabla_\mathbf{x}\Psi_t^\mathsf{T}\mathbf{v} \\ \frac{\partial\widehat{\Psi}_t}{\partial t} &= \frac{1}{2}g^2\Delta_\mathbf{v}\widehat{\Psi}_t - \nabla_\mathbf{x}\widehat{\Psi}_t^\mathsf{T}\mathbf{v} \\ \textit{s.t}\quad &\Psi(\mathbf{x},\mathbf{v},0)\widehat{\Psi}(\mathbf{x},\mathbf{v},0) = p(\mathbf{x},\mathbf{v},0), \quad \Psi(\mathbf{x},\mathbf{v},T)\widehat{\Psi}(\mathbf{x},\mathbf{v},T) = p(\mathbf{x},\mathbf{v},T) \end{aligned}} \tag{24}$$

Proof. By Lemma.A.1, we notice that the optimal control is:

$$\mathbf{a}^* = g\nabla_\mathbf{v}\phi.$$

By leveraging Hopf-Cole (Eberhard, 1950; Cole, 1951) transformation, here we define

$$\Psi = \exp\left(\phi\right),$$

$$\widehat{\Psi} = \mu\exp\left(-\phi\right).$$

Then we can have the following expressions:

$$\nabla \Psi = \exp(\phi)\nabla \phi$$

$$\Delta \Psi = \nabla \cdot (\nabla \Psi)$$

$$= \sum_i \frac{\partial}{\partial \mathbf{m}_i}\left[\exp\left(\phi\right)\nabla \phi\right]$$

$$= \left[\nabla \phi^\mathsf{T}\left(\exp\left(\phi\right)\frac{\partial \phi_i}{\partial \mathbf{m}_i}\right) + \exp\left(\phi\right)\frac{\partial (\nabla \phi)_i}{\partial \mathbf{m}_i}\right]$$

$$= \exp\left(\phi\right)\left[\|\nabla \phi\|_2^2 + \Delta \phi\right]$$

$$\nabla \widehat{\Psi} = \mu \exp\left(-\phi\right)\left(-\nabla \phi\right) + \exp(-\phi)\nabla \mu$$

$$= \exp(-\phi)(-\mu\nabla \phi + \nabla \mu)$$

$$\Delta \widehat{\Psi} = \nabla \cdot \left(\nabla \widehat{\Psi}\right)$$

$$= \sum_i \frac{\partial}{\partial \mathbf{m}_i}\left[\exp\left(-\phi\right)\left(-\mu\nabla \phi + \nabla \mu\right)\right]$$

$$= \sum_i \left[(-\mu\nabla \phi + \nabla \mu)^\mathsf{T}\left(\exp\left(-\phi\right)\frac{-\partial [\nabla \phi]_i}{\partial \mathbf{m}_i}\right)\right.$$

$$\left.+ \exp\left(-\phi\right)\left(\frac{\partial}{\partial \mathbf{m}_i}[\nabla \mu]_i - \mu\frac{\partial}{\partial \mathbf{m}_i}[\nabla \phi]_i - \nabla \phi^\mathsf{T}\frac{\partial}{\partial \mathbf{m}_i}[\nabla \mu]_i\right)\right]$$

$$= \exp\left(-\phi\right)\left[\mu\|\nabla \phi\|_2^2 - \nabla \mu^\mathsf{T}\nabla \phi + \Delta \mu - \mu\Delta \phi - \nabla \phi^\mathsf{T}\nabla \mu\right]$$

$$= \exp\left(-\phi\right)\left[\mu\|\nabla \phi\|_2^2 - 2\nabla \mu^\mathsf{T}\nabla \phi + \Delta \mu - \mu\Delta \phi\right]$$

Thus, we can have.

$$\frac{\partial \Psi}{\partial t} = \exp\left(\phi\right)\frac{\partial \phi}{\partial t}$$

$$= \exp\left(\phi\right)\left(-\frac{1}{2}\|g\nabla_\mathbf{v}\phi\|_2^2 - \mathbf{v}^\mathsf{T}\nabla_\mathbf{x}\phi - \frac{1}{2}g^2\Delta_\mathbf{v}\phi\right)$$

$$= -\frac{1}{2}g^2\Delta_\mathbf{v}\Psi - \nabla_\mathbf{x}\Psi^\mathsf{T}\mathbf{v}$$

$$\frac{\partial \widehat{\Psi}}{\partial t} = \exp\left(-\phi\right)\frac{\partial \mu}{\partial t} - \mu\exp\left(-\phi\right)\frac{\partial \phi}{\partial t}$$

$$= \exp\left(-\phi\right)\left(\frac{\partial \mu}{\partial t} - \mu\frac{\partial \phi}{\partial t}\right)$$

$$= \exp\left(-\phi\right)\left[-\nabla_\mathbf{m} \cdot \{[(\boldsymbol{f} + g\mathbf{u})\,\mathbf{I}_d]\,\mu\} + \frac{1}{2}g^2\Delta_\mathbf{m}\mu + \mu\left(\frac{1}{2}\|g\nabla_\mathbf{v}\phi\|_2^2 + \mathbf{v}^\mathsf{T}\nabla_\mathbf{x}\phi + \frac{1}{2}g^2\Delta_\mathbf{v}\phi\right)\right]$$

$$= \exp\left(-\phi\right)\left[-\nabla_\mathbf{v} \cdot (g^2\mu\nabla_\mathbf{v}\phi) - \mathbf{v}^\mathsf{T}\nabla_\mathbf{x}\mu + \frac{1}{2}g^2\Delta_\mathbf{v}\mu + \frac{\mu}{2}\|g\nabla_\mathbf{v}\phi\|_2^2 + \mu\mathbf{v}^\mathsf{T}\nabla_\mathbf{x}\phi + \mu\frac{1}{2}g^2\Delta_\mathbf{v}\phi\right]$$

$$= \exp\left(-\phi\right)\left[-g^2\nabla_\mathbf{v}\mu^\mathsf{T}\nabla_\mathbf{v}\phi - g^2\mu\Delta_\mathbf{v}\phi - \mathbf{v}^\mathsf{T}\nabla_\mathbf{x}\mu + \frac{1}{2}g^2\Delta_\mathbf{v}\mu + \frac{\mu}{2}\|g\nabla_\mathbf{v}\phi\|_2^2 + \mu\mathbf{v}^\mathsf{T}\nabla_\mathbf{x}\phi + \mu\frac{1}{2}g^2\Delta_\mathbf{v}\phi\right]$$

$$= \exp\left(-\phi\right)\left[-g^2\nabla_\mathbf{v}\mu^\mathsf{T}\nabla_\mathbf{v}\phi - \mu\frac{1}{2}g^2\Delta_\mathbf{v}\phi - \mathbf{v}^\mathsf{T}\nabla_\mathbf{x}\mu + \frac{1}{2}g^2\Delta_\mathbf{v}\mu + \frac{\mu}{2}\|g\nabla_\mathbf{v}\phi\|_2^2 + \mu\mathbf{v}^\mathsf{T}\nabla_\mathbf{x}\phi\right]$$

$$= \exp\left(-\phi\right)\left[\frac{\mu}{2}\|g\nabla_\mathbf{v}\phi\|_2^2 - g^2\nabla_\mathbf{v}\mu^\mathsf{T}\nabla_\mathbf{v}\phi - \mu\frac{1}{2}g^2\Delta_\mathbf{v}\phi + \frac{1}{2}g^2\Delta_\mathbf{v}\mu - \mathbf{v}^\mathsf{T}\nabla_\mathbf{x}\mu + \mu\mathbf{v}^\mathsf{T}\nabla_\mathbf{x}\phi\right]$$

$$= \frac{1}{2}g^2\Delta_\mathbf{v}\hat{\Psi} - \nabla_\mathbf{x}\widehat{\Psi}^\mathsf{T}\mathbf{v}$$

Then we can represent the optimal control as:

$$\mathbf{u}_t^{f*} := \mathbf{Z}_t \equiv \begin{pmatrix} \mathbf{0} \\ \mathbf{z}_t \end{pmatrix} \tag{25}$$

$$\equiv \begin{pmatrix} \mathbf{0} \\ g\nabla_{\mathbf{v}}\phi \end{pmatrix} \stackrel{\text{Hopf-Cole}}{=} \begin{pmatrix} \mathbf{0} \\ g\nabla_{\mathbf{v}}\log\Psi_t \end{pmatrix} \tag{26}$$

Then the solution of such mSB is characterized by the forward SDE:

$$d\mathbf{m}_t = \left[\boldsymbol{f} + g\mathbf{u}_t^{f*}\right]dt + g(t)d\mathbf{w}_t, \tag{27}$$

Due to the structure of Hopf-Cole transform, one can have

$$p_t^{SB} = p_t^{eq.(27)} = \Psi_t\widehat{\Psi}_t \tag{28}$$

According to (Nelson, 2020; Anderson, 1982), the reverse drift of such SDE (eq.27) $\mathbf{u}_t^{b*}$ should admits,

$$\mathbf{u}_t^{f*} + \mathbf{u}_t^{b*} = \mathbf{g}\nabla_{\mathbf{v}}\log p_t^{SB} \tag{29}$$

$$\begin{pmatrix} \mathbf{0} \\ g\nabla_{\mathbf{v}}\log\Psi_t \end{pmatrix} + \mathbf{u}_t^{b*} = \begin{pmatrix} \mathbf{0} \\ g\nabla_{\mathbf{v}}\log\Psi_t + g\nabla_{\mathbf{v}}\log\widehat{\Psi}_t \end{pmatrix} \tag{30}$$

$$\mathbf{u}_t^{b*} = \begin{pmatrix} \mathbf{0} \\ g\nabla_{\mathbf{v}}\log\widehat{\Psi}_t \end{pmatrix} \tag{31}$$

which yields The backward optimal control

$$\mathbf{u}_t^{b*} := \widehat{\mathbf{Z}}_t \equiv \begin{pmatrix} \mathbf{0} \\ \widehat{\mathbf{z}}_t \end{pmatrix} \equiv \begin{pmatrix} \mathbf{0} \\ g\nabla_{\mathbf{v}}\log\widehat{\Psi}_t \end{pmatrix} \tag{32}$$

Thus, the optimal forward and backward process is

$$d\mathbf{m}_t = \left[\boldsymbol{f} + g\mathbf{u}_t^{f*}\right]dt + g(t)d\mathbf{w}_t \tag{33}$$

$$d\mathbf{m}_s = \left[\boldsymbol{f} + g\mathbf{u}_s^{b*}\right]dt + g(t)d\widehat{\mathbf{w}}_s \tag{34}$$

And Ψ and $\widehat{\Psi}$ satisfy following PDEs,

$$\frac{\partial\Psi_t}{\partial t} = -\frac{1}{2}g^2\Delta_{\mathbf{v}}\Psi_t - \nabla_{\mathbf{x}}\Psi_t^{\mathsf{T}}\mathbf{v}$$

$$\frac{\partial\widehat{\Psi}_t}{\partial t} = \frac{1}{2}g^2\Delta_{\mathbf{v}}\widehat{\Psi}_t - \nabla_{\mathbf{x}}\widehat{\Psi}_t^{\mathsf{T}}\mathbf{v}$$

$\square$

Lemma A.3. *By specifying $\boldsymbol{f} = [\mathbf{v}, \mathbf{0}]^{\mathsf{T}}$, The PDE shown in 24 can be represented by following SDEs*

$$\begin{pmatrix} d\mathbf{x} \\ d\mathbf{v} \end{pmatrix} = \begin{pmatrix} \mathbf{v} \\ -g^2\nabla_{\mathbf{v}}\log\Psi \end{pmatrix}dt + \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & g \end{pmatrix}d\mathbf{w} \tag{35}$$

$$d\mathbf{y} = \frac{1}{2}\|\mathbf{z}\|^2 dt + \mathbf{z}^{\mathsf{T}}d\mathbf{w}_t \tag{36}$$

$$d\widehat{\mathbf{y}} = \left[\frac{1}{2}\|\widehat{\mathbf{z}}\|^2 + \mathbf{z}^{\mathsf{T}}\widehat{\mathbf{z}} + \nabla_{\mathbf{v}}\cdot g\widehat{\mathbf{z}}\right]dt + \widehat{\mathbf{z}}^{\mathsf{T}}d\mathbf{w}_t \tag{37}$$

$$\boldsymbol{s.t}: \exp\left(\mathbf{y}_0 + \widehat{\mathbf{y}}_0\right) = p(\mathbf{x}, \mathbf{v}, 0), \quad \exp\left(\mathbf{y}_T + \widehat{\mathbf{y}}_T\right) = p(\mathbf{x}, \mathbf{v}, T) \tag{38}$$

Where:

$$\mathbf{y} \equiv \mathbf{y}(\mathbf{x}_t, \mathbf{v}, t) = \log\Psi(\mathbf{x}_t, \mathbf{v}_t, t), \quad \mathbf{z} \equiv \mathbf{z}(\mathbf{x}_t, \mathbf{v}_t, t) = g\nabla_{\mathbf{v}}\log\Psi(\mathbf{x}_t, \mathbf{v}_t, t)$$

$$\widehat{\mathbf{y}} \equiv \widehat{\mathbf{y}}(\mathbf{x}_t, \mathbf{v}_t, t) = \log\widehat{\Psi}(\mathbf{x}_t, \mathbf{v}_t, t), \quad \widehat{\mathbf{z}} \equiv \widehat{\mathbf{z}}(\mathbf{x}_t, \mathbf{v}_t, t) = g\nabla_{\mathbf{v}}\log\widehat{\Psi}(\mathbf{x}_t, \mathbf{v}_t, t)$$

Proof. One can write

$$\frac{\partial \log \Psi}{\partial t} = \frac{1}{\Psi}\left(-\nabla_{\mathbf{x}}\Psi^{\mathsf{T}}\mathbf{v} - \frac{1}{2}g^2 \Delta_{\mathbf{v}}\Psi\right)$$

$$= -\nabla_{\mathbf{x}}\log\Psi^{\mathsf{T}}\mathbf{v} - \frac{1}{2}g^2\frac{\Delta_{\mathbf{v}}\Psi}{\Psi}$$

$$= -\nabla_{\mathbf{x}}\log\Psi^{\mathsf{T}}\mathbf{v} - \frac{1}{2}g^2\operatorname{Tr}\left[\frac{1}{\Psi}\nabla_{\mathbf{v}}^2\Psi\right]$$

$$\frac{\partial \log \widehat{\Psi}}{\partial t} = \frac{1}{\widehat{\Psi}}\left(-\nabla_{\mathbf{x}}\widehat{\Psi}^{\mathsf{T}}\mathbf{v} + \frac{1}{2}g^2 \Delta_{\mathbf{v}}\widehat{\Psi}\right)$$

$$= -\nabla_{\mathbf{x}}\log\widehat{\Psi}^{\mathsf{T}}\mathbf{v} - \frac{1}{2}g^2\operatorname{Tr}\left[\frac{1}{\widehat{\Psi}}\nabla_{\mathbf{v}}^2\widehat{\Psi}\right]$$

By applying Itô's lemma,

$$\mathrm{d}\log\Psi = \frac{\partial \log \Psi}{\partial t}\mathrm{d}t + \left[\nabla_{\mathbf{x}}\log\Psi^{\mathsf{T}}\mathbf{v} + g^2\|\nabla_{\mathbf{v}}\log\Psi\|_2^2 + \frac{1}{2}g^2\Delta_{\mathbf{v}}\log\Psi\right]\mathrm{d}t + \left[\nabla_{\mathbf{m}}\log\Psi^{\mathsf{T}}\right]g\mathrm{d}\mathbf{w}_t$$

$$= \left[-\nabla_{\mathbf{x}}\log\Psi^{\mathsf{T}}\mathbf{v} - \frac{1}{2}g^2\operatorname{Tr}\left[\frac{1}{\Psi}\nabla_{\mathbf{v}}^2\Psi\right]\right]\mathrm{d}t$$

$$+ \left[\nabla_{\mathbf{x}}\log\Psi^{\mathsf{T}}\mathbf{v} + g^2\|\nabla_{\mathbf{v}}\log\Psi\|_2^2 + \frac{1}{2}g^2\operatorname{Tr}\left[\frac{1}{\Psi}\nabla_{\mathbf{v}}^2\Psi - \frac{1}{\Psi^2}\nabla_{\mathbf{v}}\Psi\nabla_{\mathbf{v}}\Psi^{\mathsf{T}}\right]\right]\mathrm{d}t + g\left[\nabla_{\mathbf{v}}\log\Psi^{\mathsf{T}}\right]\mathrm{d}\mathbf{w}_t$$

$$= \left[g^2\|\nabla_{\mathbf{v}}\log\Psi\|_2^2 - \frac{1}{2}g^2\operatorname{Tr}\left[\frac{1}{\Psi^2}\nabla_{\mathbf{v}}\Psi\nabla_{\mathbf{v}}\Psi^{\mathsf{T}}\right]\right]\mathrm{d}t + g\left[\nabla_{\mathbf{v}}\log\Psi^{\mathsf{T}}\right]\mathrm{d}\mathbf{w}_t$$

$$= \left[\frac{1}{2}g^2\|\nabla_{\mathbf{v}}\log\Psi\|_2^2\right]\mathrm{d}t + g\left[\nabla_{\mathbf{v}}\log\Psi^{\mathsf{T}}\right]\mathrm{d}\mathbf{w}_t$$

Similarly, one can have,

$$\mathrm{d}\log\widehat{\Psi} = \frac{\partial \log \widehat{\Psi}}{\partial t}\mathrm{d}t + \left[\nabla_{\mathbf{x}}\log\widehat{\Psi}^{\mathsf{T}}\mathbf{v} + g^2\nabla_{\mathbf{v}}\log\Psi^{\mathsf{T}}\nabla_{\mathbf{v}}\log\widehat{\Psi} + \frac{1}{2}g^2\Delta_{\mathbf{v}}\log\widehat{\Psi}\right]\mathrm{d}t + \left[\nabla_{\mathbf{m}}\log\widehat{\Psi}^{\mathsf{T}}\right]g\mathrm{d}\mathbf{w}_t$$

$$= \left[-\nabla_{\mathbf{x}}\log\widehat{\Psi}^{\mathsf{T}}\mathbf{v} + \frac{1}{2}g^2\operatorname{Tr}\left[\frac{1}{\widehat{\Psi}}\nabla_{\mathbf{v}}^2\widehat{\Psi}\right]\right]\mathrm{d}t$$

$$+ \left[\nabla_{\mathbf{x}}\log\widehat{\Psi}^{\mathsf{T}}\mathbf{v} + g^2\nabla_{\mathbf{v}}\log\Psi^{\mathsf{T}}\nabla_{\mathbf{v}}\log\widehat{\Psi} + \frac{1}{2}g^2\Delta_{\mathbf{v}}\log\widehat{\Psi}\right]\mathrm{d}t + g\left[\nabla_{\mathbf{v}}\log\widehat{\Psi}^{\mathsf{T}}\right]\mathrm{d}\mathbf{w}_t$$

Noticing:

$$\frac{1}{2}\left[\frac{1}{\widehat{\Psi}}\nabla_{\mathbf{v}}^2\widehat{\Psi} + \nabla_{\mathbf{v}}^2\log\widehat{\Psi}\right] = \operatorname{Tr}\left[\frac{1}{\widehat{\Psi}}\nabla_{\mathbf{v}}^2\widehat{\Psi} - \frac{1}{2}\|\nabla_{\mathbf{v}}\log\widehat{\Psi}\|^2\right]$$

$$= \frac{1}{2}\|\nabla_{\mathbf{v}}\log\widehat{\Psi}\|^2 + \Delta_{\mathbf{v}}\log\widehat{\Psi}$$

Following the above derivation, one can have,

$$
d \log \widehat{\Psi} = \left[-\nabla_{\mathbf{x}} \log \widehat{\Psi}^{\mathsf{T}} \mathbf{v} + \frac{1}{2} g^2 \operatorname{Tr} \left[\frac{1}{\widehat{\Psi}} \nabla_{\mathbf{v}}^2 \widehat{\Psi} \right] \right] dt
$$

$$
+ \left[\nabla_{\mathbf{x}} \log \widehat{\Psi}^{\mathsf{T}} \mathbf{v} + g^2 \nabla_{\mathbf{v}} \log \Psi^{\mathsf{T}} \nabla_{\mathbf{v}} \log \widehat{\Psi} + \frac{1}{2} g^2 \Delta_{\mathbf{v}} \log \widehat{\Psi} \right] dt + g \left[\nabla_{\mathbf{v}} \log \widehat{\Psi}^{\mathsf{T}} \right] d\mathbf{w}_t
$$

$$
= \left[g^2 \nabla_{\mathbf{v}} \log \Psi^{\mathsf{T}} \nabla_{\mathbf{v}} \log \widehat{\Psi} + \frac{1}{2} g^2 \| \nabla_{\mathbf{v}} \log \widehat{\Psi} \|^2 + 2 \frac{1}{2} g^2 \Delta_{\mathbf{v}} \log \widehat{\Psi} \right] dt + g \left[\nabla_{\mathbf{v}} \log \widehat{\Psi}^{\mathsf{T}} \right] d\mathbf{w}_t
$$

By defining

$$
\mathbf{y} \equiv \mathbf{y}(\mathbf{x}_t, \mathbf{v}, t) = \log \Psi(\mathbf{x}_t, \mathbf{v}_t, t), \quad \mathbf{z} \equiv \mathbf{z}(\mathbf{x}_t, \mathbf{v}_t, t) = g \nabla_{\mathbf{v}} \log \Psi(\mathbf{x}_t, \mathbf{v}_t, t)
$$

$$
\widehat{\mathbf{y}} \equiv \widehat{\mathbf{y}}(\mathbf{x}_t, \mathbf{v}_t, t) = \log \widehat{\Psi}(\mathbf{x}_t, \mathbf{v}_t, t), \quad \widehat{\mathbf{z}} \equiv \widehat{\mathbf{z}}(\mathbf{x}_t, \mathbf{v}_t, t) = g \nabla_{\mathbf{v}} \log \widehat{\Psi}(\mathbf{x}_t, \mathbf{v}_t, t)
$$

One can conclude the results.

$$
\begin{pmatrix} d\mathbf{x} \\ d\mathbf{v} \end{pmatrix} = \begin{pmatrix} \mathbf{v} \\ -g^2 \nabla_{\mathbf{v}} \log \Psi \end{pmatrix} dt + \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & g \end{pmatrix} d\mathbf{w}
$$

$$
d\mathbf{y} = \frac{1}{2} \| \mathbf{z} \|^2 dt + \mathbf{z}^{\mathsf{T}} d\mathbf{w}_t
$$

$$
d\widehat{\mathbf{y}} = \left[\frac{1}{2} \| \widehat{\mathbf{z}} \|^2 + \mathbf{z}^{\mathsf{T}} \widehat{\mathbf{z}} + \nabla_{\mathbf{v}} \cdot g\widehat{\mathbf{z}} \right] dt + \widehat{\mathbf{z}}^{\mathsf{T}} d\mathbf{w}_t
$$

$$
\mathbf{s.t} : \exp(\mathbf{y}_0 + \widehat{\mathbf{y}}_0) = p(\mathbf{x}, \mathbf{v}, 0), \quad \exp(\mathbf{y}_T + \widehat{\mathbf{y}}_T) = p(\mathbf{x}, \mathbf{v}, T)
$$

$\square$

Proposition A.4. *The log-likelihood at data point* $\mathbf{m}_0$ *can be expressed as*

$$
\log p(\mathbf{m}_0, 0) = \mathbb{E}_{\mathbf{m}_t \sim (21)} \left[\log p(\mathbf{m}_T, T) \right] - \int_0^T \mathbb{E}_{\mathbf{m}_t \sim (21)} \left[\frac{1}{2} \| \mathbf{z}_t \|^2 dt + \frac{1}{2} \| \widehat{\mathbf{z}}_t \|^2 + \mathbf{z}_t^{\mathsf{T}} \widehat{\mathbf{z}}_t + \nabla_{\mathbf{v}} \cdot g\widehat{\mathbf{z}}_t \right] dt
$$

$$
= \mathbb{E}_{\mathbf{m}_t \sim (21)} \left[\log p(\mathbf{m}_T, T) \right] - \int_0^T \mathbb{E}_{\mathbf{m}_t \sim (21)} \left[\frac{1}{2} \| \mathbf{z}_t \|^2 + \underbrace{\frac{1}{2} \| \widehat{\mathbf{z}}_t - g\nabla_{\mathbf{v}} \log p^{SB} + \mathbf{z}_t \|^2}_{\textit{mean matching objective}} - \frac{1}{2} \| g\nabla_{\mathbf{v}} \log p^{SB} - \mathbf{z}_t \|^2 \right] dt
$$

$$
\log p(\mathbf{m}_T, T) = \mathbb{E}_{\mathbf{m}_t \sim (20)} \left[\log p(\mathbf{m}_0, 0) \right] - \int_0^T \mathbb{E}_{\mathbf{m}_t \sim (20)} \left[\frac{1}{2} \| \mathbf{z}_t \|^2 dt + \frac{1}{2} \| \widehat{\mathbf{z}}_t \|^2 + \mathbf{z}_t^{\mathsf{T}} \widehat{\mathbf{z}}_t + \nabla_{\mathbf{v}} \cdot g\mathbf{z}_t \right] dt
$$

$$
= \mathbb{E}_{\mathbf{m}_t \sim (20)} \left[\log p(\mathbf{m}_0, 0) \right] - \int_0^T \mathbb{E}_{\mathbf{m}_t \sim (20)} \left[\frac{1}{2} \| \widehat{\mathbf{z}}_t \|^2 + \underbrace{\frac{1}{2} \| \widehat{\mathbf{z}}_t - g\nabla_{\mathbf{v}} \log p^{SB} + \mathbf{z}_t \|^2}_{\textit{mean matching objective}} - \frac{1}{2} \| g\nabla_{\mathbf{v}} \log p^{SB} - \widehat{\mathbf{z}}_t \|^2 \right] dt
$$

By maximizing the log-likelihood at time $t = 0$ *then* $t = T$ *iteratively,* $(\mathbf{z}_t, \widehat{\mathbf{z}}_t)$ *will converge to the solution of phase space SB.*

Proof. from Lemma.A.3, one can have:

$$\log p(\mathbf{m}_0, 0) = \mathbb{E}\left[\mathbf{y}_0 + \widehat{\mathbf{y}}_0\right]$$

$$= \mathbb{E}\left[\mathbf{y}_T + \widehat{\mathbf{y}}_T\right] - \int_0^T \mathbb{E}\left[\frac{1}{2}\|\mathbf{z}_t\|^2 \mathrm{d}t + \frac{1}{2}\|\widehat{\mathbf{z}}_t\|^2 + \mathbf{z}_t^\mathsf{T}\widehat{\mathbf{z}}_t + \nabla_\mathbf{v} \cdot g\widehat{\mathbf{z}}_t\right] \mathrm{d}t$$

$$= \mathbb{E}\left[\log p(\mathbf{m}_T, T)\right] - \int_0^T \mathbb{E}\left[\frac{1}{2}\|\mathbf{z}_t\|^2 + \frac{1}{2}\|\widehat{\mathbf{z}}_t\|^2 + \mathbf{z}_t^\mathsf{T}\widehat{\mathbf{z}}_t + \nabla_\mathbf{v} \cdot g\widehat{\mathbf{z}}_t\right] \mathrm{d}t$$

$$= \mathbb{E}\left[\log p(\mathbf{m}_T, T)\right] - \int_0^T \mathbb{E}\left[\frac{1}{2}\|\mathbf{z}_t\|^2 + \frac{1}{2}\|\widehat{\mathbf{z}}_t\|^2 - \widehat{\mathbf{z}}_t^\mathsf{T}\left(g\nabla_\mathbf{v}\log p^{SB}\right) + \mathbf{z}_t^\mathsf{T}\widehat{\mathbf{z}}_t\right] \mathrm{d}t$$

$$= \mathbb{E}\left[\log p(\mathbf{m}_T, T)\right] - \int_0^T \mathbb{E}\left[\frac{1}{2}\|\mathbf{z}_t\|^2 + \frac{1}{2}\|\widehat{\mathbf{z}}_t - g\nabla_\mathbf{v}\log p^{SB} + \mathbf{z}_t\|^2 - \frac{1}{2}\|g\nabla_\mathbf{v}\log p^{SB} - \mathbf{z}_t\|^2\right] \mathrm{d}t$$

A similar result can be obtained for $\log p(\mathbf{m}_T, T)$.

One can notice that the likelihood objective is a continuous time analog of the mean matching objective proposed in (De Bortoli et al., 2021), and iterative optimization between $log p(\mathbf{m}_0, 0)$ and $\log p(\mathbf{m}_T, T)$ are the continuous analog of IPF. Hence, the convergence proof will keep valid (see Proposition 4 in (De Bortoli et al., 2021)). □

Proposition A.5. *Given the reference path measure $\bar{\pi}^\phi$ driven by the policy $\widehat{\mathbf{z}}_t^\phi$ from boundary $\mu_{t_{i+1}}$ in the reverse time direction, the optimal path measure in the forward time direction of the following problem*

$$\min_\pi \mathcal{J}(\pi) := \sum_{i=0}^{N-1} KL\left(\pi_{t_i, t_{i+1}} | \bar{\pi}_{t_i, t_{i+1}}^\phi\right),$$

$$s.t \quad \pi \in \left\{\int \pi_{t_i t_{i+1}} \mathrm{d}\mathbf{m}_{t_{i+1}} = \mu_{t_i}, \int \mu_{t_i} \mathrm{d}\mathbf{v}_{t_i} = \rho_{t_i}\right\}$$

is:

$$\pi_{t_i t_{i+1}}^* = \frac{\rho_{t_i} \bar{\pi}_{t_i, t_{i+1}}^\phi}{\int \bar{\pi}_{t_i, t_{i+1}}^\phi \mathrm{d}\mathbf{m}_{t_{i+1}} \mathrm{d}\mathbf{v}_{t_i}}. \tag{39}$$

When

$$\pi_{t_i t_{i+1}}^\theta \equiv \pi_{t_i t_{i+1}}^* \tag{40}$$

following equivalences have to be hold $\forall t \in [t_i, t_{i+1}]$:

$$\|\mathbf{z}_t^\theta + \widehat{\mathbf{z}}_t^\phi - \nabla_\mathbf{v}\log p_t^{SB}\|_2^2 = 0 \tag{41}$$

$$p_{t_i}(\mathbf{v}_{t_i}|\mathbf{x}_{t_i}^\phi) \equiv q_{t_i}^\phi(\mathbf{v}_{t_i}|\mathbf{x}_{t_i}) \tag{42}$$

Proof. Due to the similarity of optimization for $\mathcal{K}_{\text{boundary}}$, the close form solution of the next path measure is (see §4 in (Chen et al., 2019) for detail):

$$\pi_{t_i t_{i+1}}^* = \frac{\rho_{t_i} \bar{\pi}_{t_i, t_{i+1}}^\phi}{\int \bar{\pi}_{t_i, t_{i+1}}^\phi \mathrm{d}\mathbf{m}_{t_{i+1}} \mathrm{d}\mathbf{v}_{t_i}}.$$

By denoting the transition kernel of parameterized SDE driven by backward policy $\widehat{\mathbf{z}}_t^\phi$ as $q^\phi(\cdot|\cdot)$, and the time range between

t_i and t_{i+1} is discretized into S interval by EM discretization. Then one can get

$$\pi^*_{t_i t_{i+1}} = \frac{\rho_{t_i} \bar\pi^\phi_{t_i, t_{i+1}}}{\int \bar\pi^\phi_{t_i, t_{i+1}} \mathrm{d}\mathbf{m}_{t_{i+1}} \mathrm{d}\mathbf{v}_{t_i}}$$

$$= \frac{p_{t_i}(\mathbf{x}_{t_i}) q^\phi_{t_i}(\mathbf{m}_{t_i}|\mathbf{m}_{t_i+\delta_t}) \cdots q^\phi_{t_{i+1}-\delta_t}(\mathbf{m}_{t_{i+1}-\delta_t}|\mathbf{m}_{t_{i+1}}) \mu_{t_{i+1}}(\mathbf{m}_{t_{i+1}})}{q^\phi_{t_i}(\mathbf{x}_{t_i})}$$

$$= \frac{p_{t_i}(\mathbf{x}_{t_i}) q^\phi_{t_i}(\mathbf{x}_{t_i}, \mathbf{v}_{t_i}|\mathbf{x}_{t_i+\delta_t}, \mathbf{v}_{t_i+\delta_t}) q^\phi_{t_i+\delta_t}(\mathbf{x}_{t_i+\delta_t}, \mathbf{v}_{t_i+\delta_t}) \cdots q^\phi_{t_{i+1}-\delta_t}(\mathbf{m}_{t_{i+1}-\delta_t}|\mathbf{m}_{t_{i+1}}) \mu_{t_{i+1}}(\mathbf{m}_{t_{i+1}})}{q^\phi_{t_i}(\mathbf{x}_{t_i}) q^\phi_{t_i+\delta_t}(\mathbf{x}_{t_i+\delta_t}, \mathbf{v}_{t_i+\delta_t})}$$

$$= \frac{p_{t_i}(\mathbf{x}_{t_i}) q^\phi_{t_i}(\mathbf{x}_{t_i}, \mathbf{v}_{t_i}, \mathbf{x}_{t_i+\delta_t}, \mathbf{v}_{t_i+\delta_t}) \cdots q^\phi_{t_{i+1}-\delta_t}(\mathbf{m}_{t_{i+1}-\delta_t}|\mathbf{m}_{t_{i+1}}) \mu_{t_{i+1}}(\mathbf{m}_{t_{i+1}})}{q^\phi_{t_i}(\mathbf{x}_{t_i}) q^\phi_{t_i+\delta_t}(\mathbf{x}_{t_i+\delta_t}, \mathbf{v}_{t_i+\delta_t})}$$

$$= \frac{p_{t_i}(\mathbf{x}_{t_i}) q^\phi_{t_i}(\mathbf{v}_{t_i}, \mathbf{x}_{t_i+\delta_t}, \mathbf{v}_{t_i+\delta_t}|\mathbf{x}_{t_i}) \cdots q^\phi_{t_{i+1}-\delta_t}(\mathbf{m}_{t_{i+1}-\delta_t}|\mathbf{m}_{t_{i+1}}) \mu_{t_{i+1}}(\mathbf{m}_{t_{i+1}})}{q^\phi_{t_i+\delta_t}(\mathbf{m}_{t_i+\delta_t})}$$

$$= p_{t_i}(\mathbf{x}_{t_i}) q^\phi_{t_i}(\mathbf{v}_{t_i}, \mathbf{x}_{t_i+\delta_t}, \mathbf{v}_{t_i+\delta_t}|\mathbf{x}_{t_i}) \frac{q^\phi_{t_i+\delta_t}(\mathbf{m}_{t_i+\delta_t}|\mathbf{m}_{t_i+2\delta_t}) \cdots q^\phi_{t_{i+1}-\delta_t}(\mathbf{m}_{t_{i+1}-\delta_t}|\mathbf{m}_{t_{i+1}}) \mu_{t_{i+1}}(\mathbf{m}_{t_{i+1}})}{q^\phi_{t_i+\delta_t}(\mathbf{m}_{t_i+\delta_t})}$$

$$= p_{t_i}(\mathbf{x}_{t_i}) q^\phi(\mathbf{v}_{t_i}|\mathbf{x}_{t_i}) q^\phi(\mathbf{m}_{t_i+\delta_t}|\mathbf{m}_{t_i}) \frac{q^\phi_{t_i+\delta_t}(\mathbf{m}_{t_i+\delta_t}|\mathbf{m}_{t_i+2\delta_t}) \cdots q^\phi_{t_{i+1}-\delta_t}(\mathbf{m}_{t_{i+1}-\delta_t}|\mathbf{m}_{t_{i+1}}) \mu_{t_{i+1}}(\mathbf{m}_{t_{i+1}})}{q^\phi_{t_i+\delta_t}(\mathbf{m}_{t_i+\delta_t})}$$

$$= p_{t_i}(\mathbf{x}_{t_i}) q^\phi(\mathbf{v}_{t_i}|\mathbf{x}_{t_i}) q^\phi(\mathbf{m}_{t_i+\delta_t}|\mathbf{m}_{t_i}) \frac{q^\phi_{t_i+\delta_t}(\mathbf{m}_{t_i+2\delta_t}|\mathbf{m}_{t_i+\delta_t}) \cdots q^\phi_{t_{i+1}-\delta_t}(\mathbf{m}_{t_{i+1}-\delta_t}|\mathbf{m}_{t_{i+1}}) \mu_{t_{i+1}}(\mathbf{m}_{t_{i+1}})}{q^\phi_{t_i+\delta_t}(\mathbf{m}_{t_i+2\delta_t})} \quad (43)$$

Doing eq.43 revursively

$$= p_{t_i}(\mathbf{x}_{t_i}) q^\phi(\mathbf{v}_{t_i}|\mathbf{x}_{t_i}) q^\phi(\mathbf{m}_{t_i+\delta_t}|\mathbf{m}_{t_i}) \prod_{s=1}^{S-1} q^\phi_s(\mathbf{m}_{t_i+(s+1)\cdot\delta_t}|\mathbf{m}_{t_i+s\cdot\delta_t})$$

$$= p_{t_i}(\mathbf{x}_{t_i}) q^\phi(\mathbf{v}_{t_i}|\mathbf{x}_{t_i}) \prod_{s=0}^{S-1} q^\phi_s(\mathbf{m}_{t_i+(s+1)\cdot\delta_t}|\mathbf{m}_{t_i+s\cdot\delta_t})$$

According to (De Bortoli et al., 2021), given the policy $\widehat{\mathbf{z}}^\phi_t$, the transition kernel $q^\phi_s(\mathbf{m}_{t_i+(s+1)\cdot\delta_t}|\mathbf{m}_{t_i+s\cdot\delta_t})$ can be estimated by $\widehat{\mathbf{z}}^\phi_t$ (see Proposition 3 in (De Bortoli et al., 2021))and it can be treated as the label for the forward policy $\mathbf{z}^\theta_t$ for all s. Thus, if $\pi^\theta_{t_i t_{i+1}}$ is aligned with $\pi^*_{t_i t_{i+1}}$, then one can construct following objective function for policy $\mathbf{z}^\theta_t$:

$$\mathcal{L}(\theta) = \sum_t \|\underbrace{\mathbf{m}_t + \delta_t \mathbf{Z}^\theta_t(\mathbf{m}_t)}_{①} - (\mathbf{m}_t + \underbrace{\mathbf{m}_{t+\delta_t} + \delta_t \widehat{\mathbf{Z}}^\phi_{t+\delta_t}(\mathbf{m}_{t+\delta_t})}_{②}) - \underbrace{(\mathbf{m}_t + \delta_t \widehat{\mathbf{Z}}^\phi_{t+\delta_t}(\mathbf{m}_t))}_{③})\|^2_2 \quad (44)$$

$$= \sum_t \|\delta_t \mathbf{Z}^\theta_t(\mathbf{m}_t) + \delta_t \widehat{\mathbf{Z}}^\phi_{t+\delta_t}(\mathbf{m}_t) - (\mathbf{m}_{t+\delta_t} - \mathbf{m}_t - \delta_t \widehat{\mathbf{Z}}^\phi_{t+\delta_t}(\mathbf{m}_{t+\delta_t}))\|^2_2 \quad (45)$$

$$\approx \sum_t \|\mathbf{Z}^\theta_t(\mathbf{m}_t) + \widehat{\mathbf{Z}}^\phi_{t+\delta_t}(\mathbf{m}_t) - \nabla_\mathbf{v} \log p^{SB}\|^2_2 \quad (46)$$

due to the special structure of $\mathbf{Z}_t$ and $\widehat{\mathbf{Z}}_t$ $\quad (47)$

$$= \sum_t \|\mathbf{z}^\theta_t(\mathbf{m}_t) + \widehat{\mathbf{z}}^\phi_{t+\delta_t}(\mathbf{m}_t) - \nabla_\mathbf{v} \log p^{SB}\|^2_2 \quad (48)$$

Where ①, ②, ③ corresponds to F_k, B_k, and B_{k+1} in (De Bortoli et al., 2021) respectively. Furthermore, we need to find a density function $p_{t_i}(\mathbf{v}_{t_i}|\mathbf{x}^\phi_{t_i})$ which satisfies

$$p_{t_i}(\mathbf{x}_{t_i}) p_{t_i}(\mathbf{v}_{t_i}|\mathbf{x}^\phi_{t_i}) \equiv p_{t_i}(\mathbf{x}_{t_i}) q^\phi(\mathbf{v}_{t_i}|\mathbf{x}_{t_i})$$

$$p_{t_i}(\mathbf{v}_{t_i}|\mathbf{x}^\phi_{t_i}) \equiv q^\phi_{t_i}(\mathbf{v}_{t_i}|\mathbf{x}_{t_i})$$

to be the new boundary condition. □

Proposition A.6. *Given the reference path measure $\bar{\pi}^\phi$ driven by the policy $\widehat{\mathbf{z}}_t^\phi$ from boundary μ_{t_N} in the reverse time direction, the optimal path measure in the forward time direction of the following problem*

$$\min_\theta \mathcal{J}(\pi^\theta) := \sum_{i=0}^{N-1} KL\left(\pi^\theta_{t_i,t_{i+1}} | \bar{\pi}^\phi_{t_i,t_{i+1}}\right),$$

$$s.t \quad \pi^\theta \in \mathcal{K}_{bridge} = \left\{\cap_{i=1}^{N-1}\mathcal{K}_{t_i}^2\right\}$$

is:

$$\pi_{t_0 t_N} = \frac{q_{t_i}^\phi \bar{\pi}^\phi_{t_0,t_N}}{\int \bar{\pi}^\phi_{t_0,t_N} \mathbf{dm}_{t_N} \mathbf{dv}_{t_0}}. \tag{49}$$

when

$$\pi^\theta_{t_0 t_N} \equiv \pi_{t_0 t_N} \tag{50}$$

following equivalence have to be hold $\forall t \in [t_0, t_N]$:

$$\|\mathbf{z}_t^\theta + \widehat{\mathbf{z}}_t^\phi - \nabla_\mathbf{v} \log p_t^{SB}\|_2^2 = 0 \tag{51}$$

$$p_{t_0}(\mathbf{v}_{t_0}, \mathbf{x}_{t_0}) \equiv q_{t_0}^\phi(\mathbf{v}_{t_0}, \mathbf{x}_{t_0}) \tag{52}$$

Proof. Same proof as A.5. □

B. Experiment Details

Training: We use Exponential Moving Average (EMA) with a decay rate of 0.999. Table.8 details the hyperparameters used for each dataset. The learning rate for all the datasets is set to be 2e-4 and the training batching size is 256. For computation efficiency, we cache large batch size of empirical samples from reference trajectory and sample training batch size from the cache data. The hyperparameters can be found in Table.8.
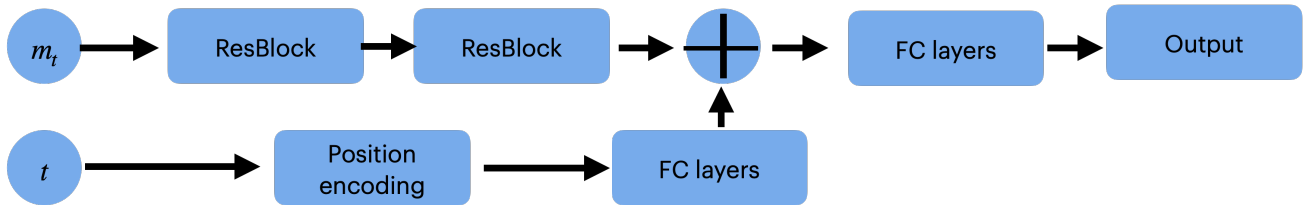

Figure 7. Neural network architecture for all experiments. The network size (# parameters) are varying between different tasks.

Figure 8. Training Hyper-parameters

Dataset	time steps	cache batch size	$g(t)$	# Parameters
Semicircle	400	2000	0.2	1.21M
Petal	400	2000	0.2	1.21M
GMM	400	2000	0.2	1.21M
scRNA (100 dim)	400	4000	0.4	1.34M

Langevin sampling: The Langevin sampling procedure for the velocity is summarized in 2. Given some pre-defined signal-to-noise ratio r (we set snr =0.15 for all experiments), the Langevin noise scale σ at each time step t and each corrector

step i is computed by

$$\sigma_t = \frac{2r^2 g^2 \|\epsilon\|^2}{\|\mathbf{z}(t, \mathbf{m}_t) + \widehat{\mathbf{z}}(t, \mathbf{m}_t)\|^2},\tag{53}$$

Algorithm 1 Sampling Procedure of DMSB

Input: Policies $\mathbf{z}(\cdot, \cdot; \theta)$ and $\widehat{\mathbf{z}}(\cdot, \cdot; \phi)$ Total sampling step $S = \frac{t_N}{\delta_t}$. Data distributions ρ_{t_i}. Initializing velocity distributions $\gamma_{t_i} = \mathcal{N}(0, \mathbf{I})$ if they are not avaliable.
for $s = 0$ **to** $S - 1$ **do**
  if s==0 **then**
    Sample position data $\mathbf{x}_{t_0}$ from ρ_{t_0}.
    if ground truth velocity distribution γ_{t_0} avaliable **then**
      Sample velocity data $\mathbf{v}_{t_0}$ from γ_{t_0}
    else
      Sample velocity data $\mathbf{v}_{t_0}$ by Langevin simulation conditioning on $\mathbf{x}_{t_0}$.(Algorithm.2)
    end if
    $\mathbf{m}_{t_0} = [\mathbf{x}_{t_0}, \mathbf{v}_{t_0}]^\mathsf{T}$
  end if
  Simulating dynamics: $\mathrm{d}\mathbf{m}_t = [\boldsymbol{f}(\mathbf{m}_t, t) + g(t)\mathbf{Z}_t]\,\mathrm{d}t + g(t)\mathrm{d}\mathbf{w}_t$(eq.20)
end for
return $\mathbf{m}_{t \in [t_0, t_N]}$

Algorithm 2 Langevin Sampler at t_i marginal constraint

Input: policies $\mathbf{z}(\cdot, \cdot; \theta)$ and $\widehat{\mathbf{z}}(\cdot, \cdot; \phi)$, Previous timestep predicted velocity $\mathbf{v}_{t_i}$.
Sample position from ground truth $\mathbf{x}_{t_i} \sim \rho_{t_i}$.
for $step = 0$ **to** # Langevin steps **do**
  Sample $\epsilon \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$.
  Construct new $\mathbf{m}_{t_i} = [\mathbf{x}_{t_i}, \mathbf{v}_{t_i}]^\mathsf{T}$
  Compute $\nabla_{\mathbf{v}} \log \tilde{p}_t^{\theta, \phi} \approx [\mathbf{z}(t_i, \mathbf{m}_{t_i}) + \widehat{\mathbf{z}}(t_i, \mathbf{m}_{t_i})]/g$.
  Compute σ_t with (53).
  Langevin Sampling $\mathbf{v}_{t_i} \leftarrow \mathbf{v}_{t_i} + \sigma_{t_i} \nabla_{\mathbf{v}} \log \tilde{p}_{t_i}^{\theta, \phi} + \sqrt{2\sigma_t}\,\epsilon$.
end for
return $\mathbf{m}_{t_i} = [\mathbf{x}_{t_i}, \mathbf{v}_{t_i}]^\mathsf{T}$

Algorithm 3 DMSB Training

Input: $N + 1$ Marginal position distribution $\rho_{t_i}, i \in [0, N]$. Parametrized policies $\mathbf{z}(\cdots ; \theta)$ and $\widehat{\mathbf{z}}(\cdots ; \phi)$. The number of Bregman Iteration B. Initialize postion and velocity at time step $t_i : \widehat{\mathbf{m}}_{t_i} \sim \mathcal{N}(0, \mathbf{I})$ for the last iteration.
if Use ground truth velocity **then**
    set prior velocity: $\gamma_{t_i} = \gamma_{t_i}$
else
    set initial velocity $\gamma_{t_i} = \mathcal{N}(0, \mathbf{I})$
end if
for $b = 0$ **to** $B - 1$ **do**
    for $k = N$ **to** 1 **do**
        $\mathbf{z}^{\phi}, {}_{-} = OptSubSet(t_k, t_{k-1}, \mathbf{z}_{ref} = \mathbf{z}^{\theta}, \mathbf{z}_{opt} = \mathbf{z}^{\phi}, \eta = \phi, \widehat{\mathbf{m}} = \widehat{\mathbf{m}})$
    end for
    for $k = 0$ **to** $N - 1$ **do**
        $\mathbf{z}^{\theta}, {}_{-} = OptSubSet(t_k, t_{k+1}, \mathbf{z}_{ref} = \mathbf{z}^{\phi}, \mathbf{z}_{opt} = \mathbf{z}^{\theta}, \eta = \theta)$
    end for
    $\mathbf{z}^{\phi}, \widehat{\mathbf{m}} = OptSubSet(t_N, t_0, \mathbf{z}_{ref} = \mathbf{z}^{\theta}, \mathbf{z}_{opt} = \mathbf{z}^{\phi}, \eta = \phi)$
    for $k = 0$ **to** $N - 1$ **do**
        $\mathbf{z}^{\theta}, {}_{-} = OptSubSet(t_k, t_{k+1}, \mathbf{z}_{ref} = \mathbf{z}^{\phi}, \mathbf{z}_{opt} = \mathbf{z}^{\theta}, \eta = \theta, \widehat{\mathbf{m}} = \widehat{\mathbf{m}})$
    end for
    for $k = N$ **to** 1 **do**
        $\mathbf{z}^{\phi}, {}_{-} = OptSubSet(t_k, t_{k-1}, \mathbf{z}_{ref} = \mathbf{z}^{\theta}, \mathbf{z}_{opt} = \mathbf{z}^{\phi}, \eta = \phi)$
    end for
    $\mathbf{z}^{\phi}, \widehat{\mathbf{m}} = OptSubSet(t_0, t_N, \mathbf{z}_{ref} = \mathbf{z}^{\phi}, \mathbf{z}_{opt} = \mathbf{z}^{\theta}, \eta = \theta)$
end for

Algorithm 4 Function OptSubSet (Optimization for subsets)

---

input: Initial time t_i and terminal time t_j. Reference path measure boundary condition ρ_{t_i}. Reference path measure driver $\mathbf{z}_{ref}$. Policy being optimized $\mathbf{z}_{opt}$ and corresponding parameter η. Empirical sample form last iteration $\widehat{\mathbf{m}}$.
output: $\mathbf{z}_{opt}$, samples $\widehat{\mathbf{m}}_{t_j}$ from reference path measure.
if $\widehat{\mathbf{m}}$ is None **then**
    Sample position data $\mathbf{x}_{t_i}$ from ρ_{t_i}.
    if velocity distribution γ_{t_i} avaliable **then**
        Sample conditional velocity data $\mathbf{v}_{t_i}$ from γ_{t_i}
    else
        Sample velocity data $\mathbf{v}_{t_i}$ by Langevin simulation conditioning on $\mathbf{x}_{t_i}$.(see Algorithm.2.)
    end if
    $\mathbf{m}_{t_i} = [\mathbf{x}_{t_i}, \mathbf{v}_{t_i}]^{\mathsf{T}}$
else
    $\mathbf{m}_{t_i} = \widehat{\mathbf{m}}$
end if
Sample trajectory $\mathbf{m}_{t \in [t_i, t_j]}$ from $\mathbf{m}_{t_i}$ using $\mathbf{z}_{ref}$
Compute $\mathcal{L} = \alpha \mathcal{L}_{\text{MM}} + (1 - \alpha)\mathcal{L}_{reg}$
update η

Table 3. Numerical result of Wasserstein-1 (W_1) and MMD on 5 dimensions single-cell RNA-seq dataset.

MMD ↓	t_1	t_2	t_3	t_4	Avg
NLSB	0.10	0.08	0.13	0.08	0.10
MIOFLOW	0.29	0.43	0.31	0.09	0.28
DMSB(ours)	**0.08**	**0.05**	**0.06**	**0.04**	**0.06**

W_1 ↓	t_1	t_2	t_3	t_4	Avg
NLSB	0.62	0.74	0.82	0.79	0.74
MIOFLOW	0.67	0.95	0.86	0.68	0.79
DMSB(ours)	**0.58**	**0.69**	**0.71**	**0.68**	**0.67**

Table 4. 100 dimensions single-cell RNA-seq dataset numerical result of velocity estimation based on Energy Distance,MMD and Sliced Wasserstein Distance (SWD). The discrepancy is averaging over all snapshots.

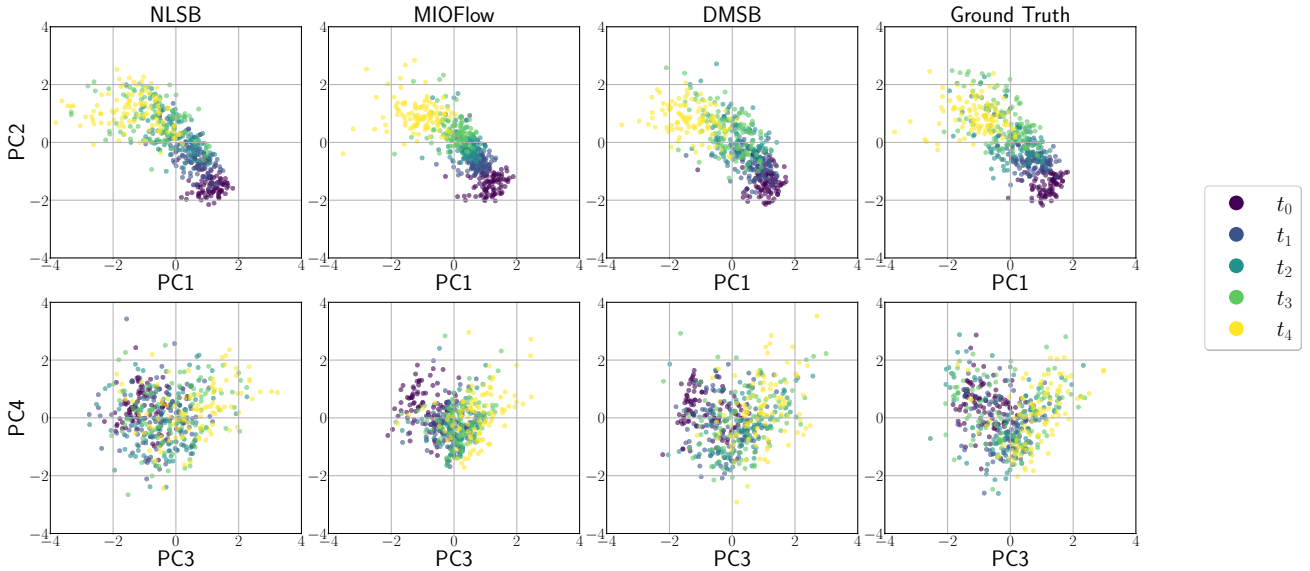| Dim=100	Energy ↓	MMD ↓	SWD ↓
NLSB	1.5	0.7	0.61
MIOFLOW	1.9	1.51	0.9
DMSB(ours)	**0.39**	**0.19**	**0.42**

C. Additional Experiment

Figure 9. Comparison of population-level dynamics on 5-dimensional PCA space at the moment of observation for scRNA-seq data using MIOFlow, NLSB, and DMSB. We display the plot of the first 4 principle components (PC). All method performs well under this experiment setup.

Table 5. 5 dimensions single-cell RNA-seq dataset numerical result of velocity estimation based on Energy Distance,MMD and Sliced Wasserstein Distance (SWD). The discrepancy is averaging over all snapshots.

Dim=5	Energy ↓	MMD ↓	SWD ↓
NLSB	0.44	1.37	0.8
MIOFLOW	0.61	2.10	0.96
DMSB(ours)	**0.39**	**0.85**	**0.76**
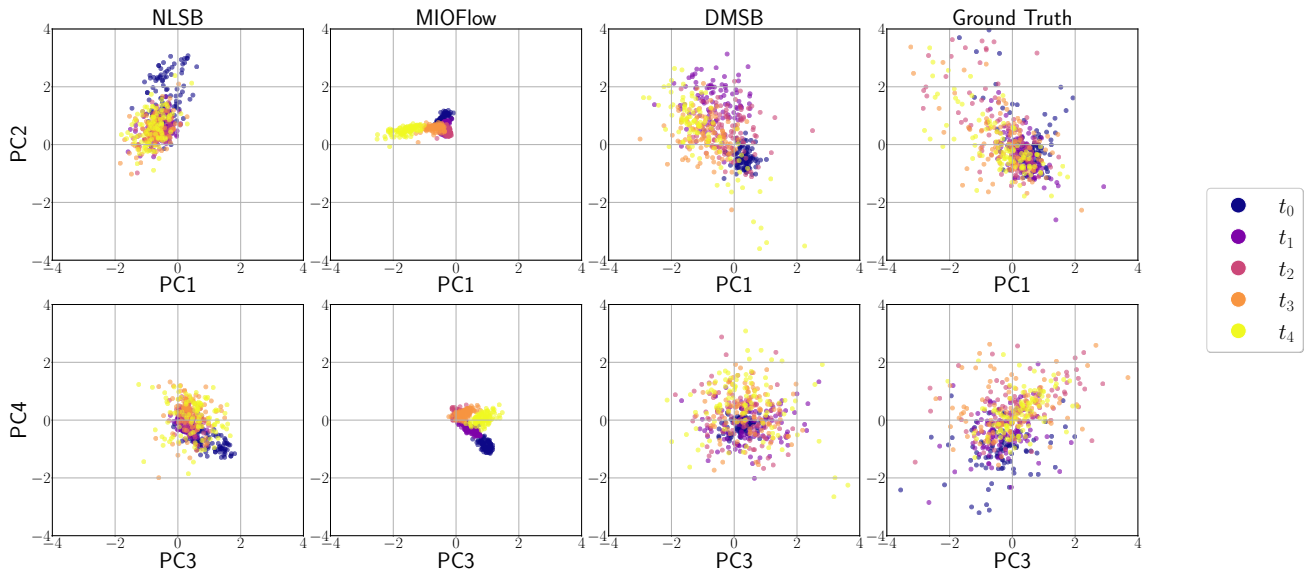
Figure 10. Comparison of estimated velocity on 5-dimensional PCA space at the moment of observation for scRNA-seq data using MIOFlow, NLSB, and DMSB. We display the plot of the first 4 principle components (PC).
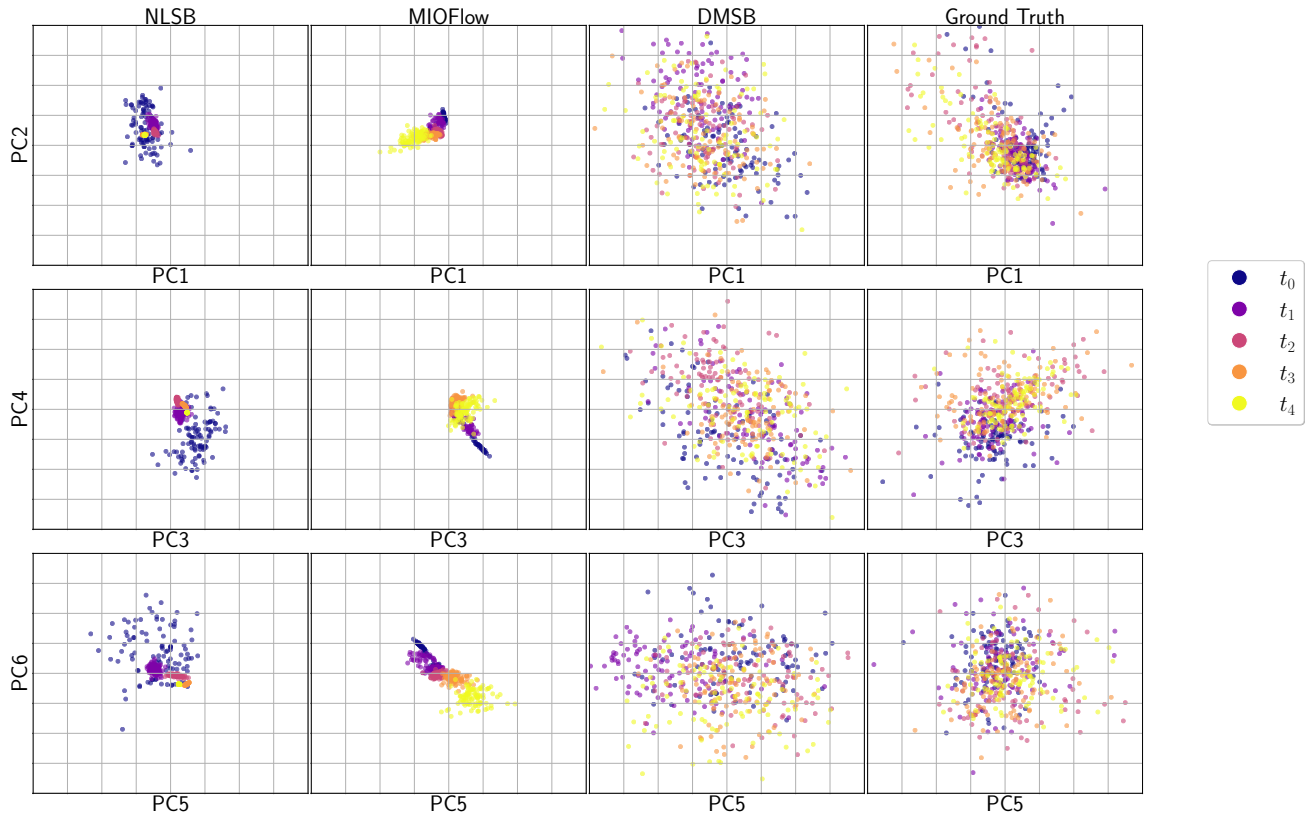
Figure 11. Comparison of estimated velocity on 100-dimensional PCA space at the moment of observation for scRNA-seq data using MIOFlow, NLSB, and DMSB. We display the plot of the first 6 principle components (PC).
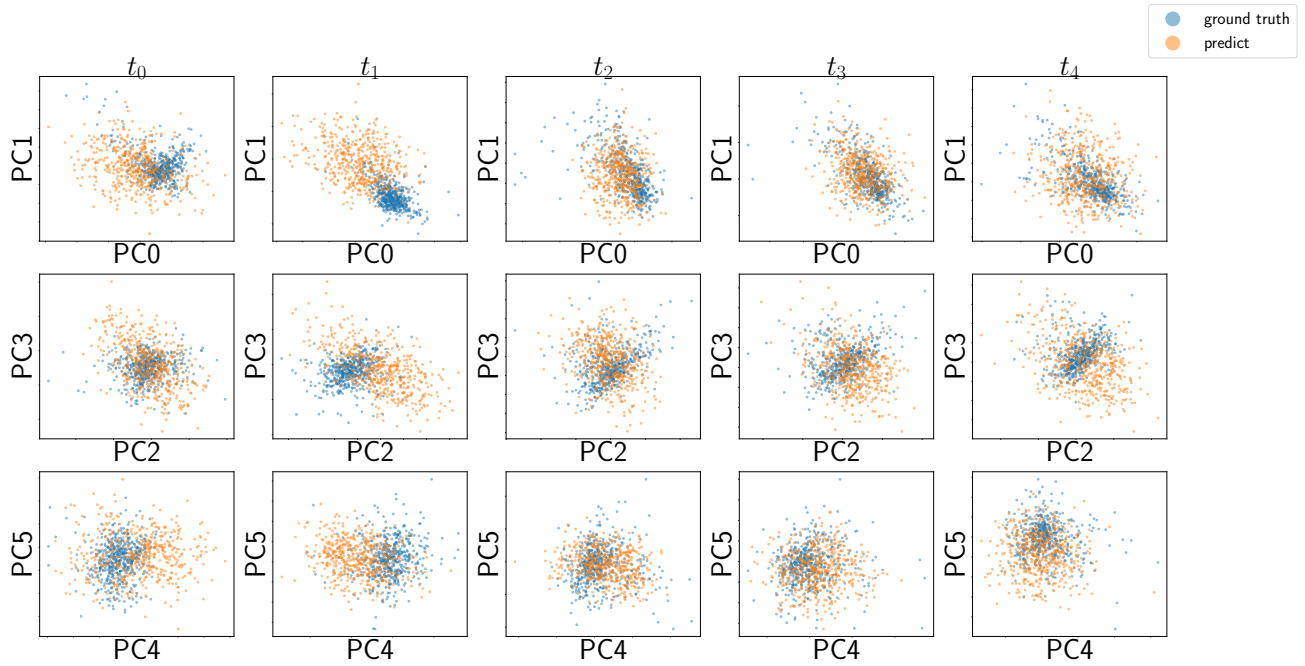
Figure 12. Comparison of estimated velocity on 100-dimensional PCA space at the moment of observation for scRNA-seq data using DMSB with ground truth. We display the plot of the first 6 principle components (PC).