Variational Online Mirror Descent for Robust Learning in Schrödinger Bridge

Anonymous authors
Paper under double-blind review

Abstract

Schrödinger bridge (SB) has evolved into a universal class of probabilistic generative models. In practice, however, estimated learning signals are often uncertain, and the reliability promised by existing methods is often based on speculative optimal-case scenarios. Recent studies regarding the Sinkhorn algorithm through mirror descent (MD) have gained attention, revealing geometric insights into solution acquisition of the SB problems. In this paper, we propose a variational online MD (OMD) framework for the SB problems, which provides further stability to SB solvers. We formally prove convergence and a regret bound for the novel OMD formulation of SB acquisition. As a result, we propose a simulation-free SB algorithm called Variational Mirrored Schrödinger Bridge (VMSB) by utilizing the Wasserstein-Fisher-Rao geometry of the Gaussian mixture parameterization for Schrödinger potentials. Based on the Wasserstein gradient flow theory, the algorithm offers tractable learning dynamics that precisely approximate each OMD step. In experiments, we validate the performance of the proposed VMSB algorithm across an extensive suite of benchmarks. VMSB consistently outperforms contemporary SB solvers on a range of SB problems, demonstrating the robustness predicted by our theory.

1 Introduction

Schrödinger bridge (SB; Schrödinger, 1932) has emerged as a universal class of probabilistic generative models. Nevertheless, learning methods of SB remain somewhat *atypical*, each requiring a sophisticated approach to derive a solution. Among various possible research directions, addressing the Schrödinger bridge problem (SBP) from a practical standpoint highlights the necessity of emphasizing *robustness*, a critical property ensuring solution reliability and stability in the presence of perturbations and distributional shifts (Xu et al., 2008; Duchi & Namkoong, 2021). Moreover, since SBP is fundamentally an infinite-dimensional distributional problem—where optimal probabilistic models inherently exhibit uncertainty when estimated from finite samples—establishing robust theoretical guarantees is both essential and challenging. The collective perspective of considering the SBP as an ordinary instance of optimization has broadly opened new avenues for algorithmic advancements of probabilistic generative models in a learning theoretical direction, particularly within the context of the learning theory and stability improvements of SB models.

Recently, learning an SB model with Sinkhorn (Peyré et al., 2019) has been generalized into mirror descent (MD; Nemirovsky & Yudin, 1983; Léger, 2021; Aubin-Frankowski et al., 2022). For parameters $\{w_t\}_{t=1}^T$ and a convex function Ω, an update of MD for a cost function F is derived as

$$\nabla\Omega(w_{t+1}) = \nabla\Omega(w_t) - \eta_t \nabla F(w_t), \tag{1}$$

where a gradient operation denoted as $\nabla\Omega(\cdot)$ creates a transformation that links a parametric space to a dual space. Aubin-Frankowski et al. (2022) formally expanded the notion of MD using directional derivatives and first variations. However, prior analyses have predominantly focused on optimal-case scenarios with a fixed cost function F, whereas theoretical improvements are possible under worst-case scenarios through the use of online convex programming methods (Zinkevich, 2003), which address unknown sequences of convex cost functions $\{F_t\}_{t=1}^\infty$. Although online mirror descent (OMD; Srebro et al., 2011; Lei & Zhou, 2020) has

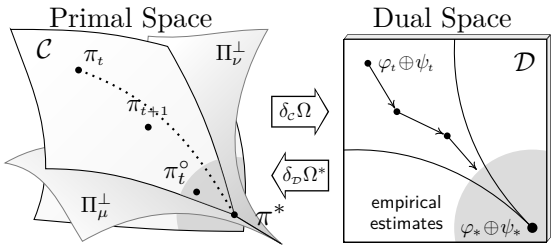

Figure 1: A schematic illustration. The primal and dual spaces $(\mathcal{C}, \mathcal{D})$ retain bidirectional maps $(\delta_{\mathcal{C}}\Omega, \delta_{\mathcal{D}}\Omega^*)$. $\Pi_\nu^\perp$ and $\Pi_\mu^\perp$ indicate projection spaces of $\gamma_1\pi = \mu$ and $\gamma_2\pi = \nu$, respectively. The current π_t performs an online learning update following a "unreliable" leader π_t° in a region shaded in gray.

Table 1: A technical overview. Combining chracteristics of existing methods, VMSB offers a simulation-free SB solver that produces iterative OMD solutions. Our VMSB additionally provides a strong theoretical guarantee of convergence based on a regert analysis.

	Iterative	Simulation-free	Regret analysis
DSB (De Bortoli et al.)	✓	✗	✗
DSBM (Shi et al.)	✓	✗	✗
LightSB (Korotin et al.)	✗	✓	✗
LightSB-M (Gushchi et al.)	✗	✓	✗
VMSB (ours)	✓	✓	✓

demonstrated effectiveness in classical settings, adapting OMD to the SBP context requires developing a novel theory and computational tools to guarantee robust algorithmic performance.

In general, one can consider constrained distributional optimization problems with generalized gradient dynamics on the space of distributions endowed with the Wasserstein metric. Leveraging the Wasserstein gradient flow (WGF) discovered by Jordan, Kinderlehrer, and Otto (JKO; Jordan et al., 1998), the desired dynamics of minimizing a cost functional $F : \mathcal{P}_2(\mathcal{X}) \to \mathbb{R}$ can be modeled, where $\mathcal{P}_2(\mathcal{X})$ denotes the set of probability distributions with finite second-order moments. Despite the extensive theoretical findings of WGF regarding OT problems (Ambrosio et al., 2005a; Santambrogio, 2015; Villani, 2021), the computational challenges remain. The established methods are commonly based on numerical methods of solving partial differential equations (PDEs) (Carlier et al., 2017; Carrillo et al., 2023), whose exhaustive numerical computations make them unsuitable for systems with high dimensional probability densities.

A favored strategy to mitigate the issue is to narrow down the solution space into a subset of tractable distributions, often referred to as taking a *variational* form (Paisley et al., 2012; Blei et al., 2017). For example, mean-field formulations of SB (Liu et al., 2022; Claisse et al., 2023) are variational approximations. Unfortunately, it does not faithfully yield an analytical submanifold and is obligated to physically simulate among particles. Recently, a Gaussian mixture parameterization of the Schrödinger potentials has been proposed by Korotin et al. (2024). The simulation-free LightSB solver is simple yet general, with the guarantee of universal approximation for SB. The expressiveness of the solver coincides with geometric properties of Gaussian variational inference and mixture models (Chen et al., 2018; Daudel et al., 2021; Lambert et al., 2022; Diao et al., 2023). However, its shortcoming—as well as other *simulation-free* solvers (Tong et al., 2024b; Gushchin et al., 2024a)—is the uncertainty of data-driven learning signals of non-convex objectives. This reveals room for improvement with the rich geometric properties of SB in a variational form.

In this paper, we explore a novel formulation of SB acquisition through the lens of dualistic geometry for ensuring further stability of learning. As illustrated in Fig. 1, our OMD method utilizes primal and dual spaces $(\mathcal{C}, \mathcal{D})$ whose geometry naturally arises in the SBP, and transformations between coupling π_t and dual potential $\varphi_t \oplus \psi_t$ are uniquely defined, and first variation operators $(\delta_{\mathcal{C}}, \delta_{\mathcal{D}})$ transform such elements (Aubin-Frankowski et al., 2022). For online learning, we postulate optimization errors of an SB solver, and propose a Variational OMD (VOMD) framework to formally reduce the errors in terms of regrets. To this end, we propose a robust simulation-free SB algorithm called Variational Mirrored Schrödinger Bridge (VMSB). The proposed VMSB offers a tractable approximation of OMD that solves iterative subproblems by Wasserstein gradient flows. To solve SB in a robust manner, the variational method is based on gradient flows with respect to the Wasserstein-Fisher-Rao (WFR) geometry. Our variational framework allows us to perform OMD which is tolerant of unreliable empirical estimates of arbitrary data-driven SB solvers. The experiments indicate that VMSB outperforms existing solvers in various benchmarks.

Our contributions. We aim to build a novel learning theoretical algorithm derived from a geometric perspective for SBP. To the best of our knowledge, VMSB is the first VOMD-based SB algorithm and

inherits the theoretical essence of OMD. Table 1 shows that VMSB is a simulation-free solver equipped with a rigorous regret bound for general situations. Our main contributions are summarized below:

- We develop a robust SB learning algorithm built upon an OMD formulation, whose rules follow Wasserstein-2 dynamics derived from local MD objectives. Under mild assumptions, we formally prove convergence of the proposed VOMD formulation in general online learning scenarios (§ 4).
- We introduce a simulation-free SB method leveraging the Wasserstein-Fisher-Rao geometry, ensuring asymptotic stability within Wasserstein gradient flows. The resulting VMSB algorithm admits closed-form dynamics, enabling accurate and computationally efficient implementation using LightSB (§ 5).
- We validate our algorithm across diverse SB problem settings, highlighting the effectiveness of our VOMD-based framework in contexts including online learning, EOT benchmarks, and image-to-image translation tasks. Empirical results consistently demonstrate that our proposed methods outperform existing SB solvers, strongly supporting our theoretical claims (§ 6).

2 Related Work

Simulation-free SB. The SBP is originated from a physical formulation for evolution of a dynamical system between measures (Léonard, 2012; Pavon & Wakolbinger, 1991). The study of SB has gained popularity due to its connection to entropy-regularized optimal transport (EOT; Peyré et al., 2019; Nutz, 2021). Its association with EOT suggests various applications across various fields related to machine learning, such as image processing, natural language processing, and control systems (Caron et al., 2020; Liu et al., 2023; Alvarez-Melis & Jaakkola, 2018; Chen et al., 2022). Historically, the most representative algorithm for SBP is Sinkhorn (Kullback, 1968), there has been progress in training SB with nonlinear networks (Vargas et al., 2021; De Bortoli et al., 2021) by "simulating" a half-bridge of forward and backward diffusion at each time. An SB solver is called as simulation-free (Tong et al., 2024a;b) if the solver is trained without samples from the simulation of SB diffusion processes. LightSB (Korotin et al., 2024) is a special type of simulation-free solver using the maximum likelihood method of Gaussian mixture models (GMMs). Building upon these advancements, our approach focuses on enhancing simulation-free SB solvers by leveraging geometric insights derived from the generalized dual geometry inherent to the SBP.

MD and Sinkhorn. The Bregman divergence (Bregman, 1967) is a family of statistical divergence that is particularly useful when analyzing constrained convex problems (Beck & Teboulle, 2003; Boyd & Vandenberghe, 2004; Hiriart-Urruty & Lemaréchal, 2004). Notably, Léger (2021) and Aubin-Frankowski et al. (2022) adopted the Bregman divergence into EOT and SB problems with probability measures, and the studies revealed that Sinkhorn can be considered to be an MD with a constant step size $\eta \equiv 1$. In statistical geometries, the Bregman divergence is a first-order approximation of a Hessian structure (Shima & Yagi, 1997; Butnariu & Resmerita, 2006), which makes MD as natural discretization on a gradient flow. Deb et al. (2023) introduced Wasserstein mirror flow, and the results include a geometric interpretation of Sinkhorn for unconstrained OT, *i.e.*, when $\varepsilon \to 0$ from our noisy setup. Karimi et al. (2024) formulated a *half-iteration* of the Sinkhorn algorithm for SB into a mirror flow, *i.e.*, $\eta_t \to 0$.

Wasserstein gradient flows have drawn significant attention whose geometry is formally described by the Wasserstein-2 metric (Ambrosio et al., 2005a; Villani, 2009; Santambrogio, 2017). Otto (2001) introduced a formal Riemannian structure to interpret various evolutionary equations as gradient flows with the Wasserstein space, which is closely related to our variational approach. The mirror Langevin dynamics is an early work describing the evolution of the Langevin diffusion (Hsieh et al., 2018), and was later incorporated in the geometry of the Bregman Wasserstein divergence (Rankin & Wong, 2023). We relate our methodology with recent approaches of variational inference on the Bures–Wasserstein space (Lambert et al., 2022; Diao et al., 2023). Utilizing Bures–Wasserstein geometry, the Wasserstein-Fisher-Rao geometry (Liero et al., 2016; Chizat et al., 2018; Liero et al., 2018; Lambert et al., 2022) additionally provides "liftings," which yield an interaction among measures.

Learning theory. Suppose we have time-varying costs $\{F_t\}_{t=1}^{\infty}$. We generally referred to learning through these signals as *online learning* (Fiat & Woeginger, 1998). Our interest lies in temporal costs defined in a probability space, where following the ordinary gradient may not be the best choice due to the geometric constraints (Amari, 2016; Amari & Nagaoka, 2000). In this sense, we primarily relate our work to the online

form of MD (Srebro et al., 2011; Raskutti & Mukherjee, 2015; Lei & Zhou, 2020). The OMD algorithm provides a generalization of robust learning by seeking solutions that are optimal in a worst-case sense, ensuring performance guarantees under adversarial or uncertain conditions (Xu et al., 2008; Zinkevich, 2003; Madry et al., 2017). Another relevant design of the online algorithm is the follow-the-regularized-leader (FTRL; McMahan, 2011). OMD focuses on scheduling proximity of updates through $\{\eta_t\}_{t=1}^T$, whereas FTRL minimizes historical losses with a fixed proximity term.

3 Generalization of Schrödinger Bridge Problems

Notation. Let $\mathcal{P}(\mathcal{S})$ ($\mathcal{P}_2(\mathcal{S})$) denote the set of (absolutely continuous) Borel probability measures on $\mathcal{S} \subseteq \mathbb{R}^d$ (with a finite second moment). For marginals $\mu, \nu \in \mathcal{P}_2(\mathcal{S})$, $\Pi(\mu, \nu)$ denotes the set of couplings (Peyré et al., 2019). For an transportation plan π, we often use a shorthand notation $\vec{\pi}^x$ ($\overleftarrow{\pi}^y$) to denotes a conditional distribution for a sample data $\vec{\pi}(\cdot|x)$ (or $\overleftarrow{\pi}(\cdot|y)$; see Fig. 2). We use $\mathrm{KL}(\cdot\|\cdot)$ to denote the KL functional and assume $+\infty$ if an argument is not absolutely continuous. We employ $\mathbb{P}([0,1], \mathcal{S})$ for a set of path measures from the from the time interval $[0, 1]$.

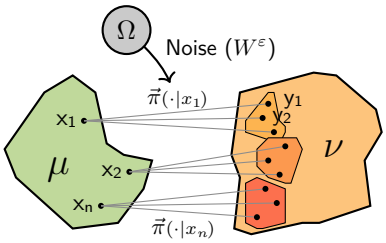

Figure 2: The SB problem.

For a positive regularization coefficient $\varepsilon \in \mathbb{R}^+$, the EOT problem, or the *static* SB problem with a quadratic cost function $c(x, y) = \frac{1}{2}\|x - y\|^2$, is defined as finding the unique minimizer π^* for the following problem:

$$\mathrm{OT}_\varepsilon(\mu, \nu) \coloneqq \inf_{\pi \in \Pi(\mu,\nu)} \iint_{\mathcal{S}\times\mathcal{S}} c(x, y)\, \mathrm{d}\pi(x, y) + \varepsilon \mathrm{KL}(\pi \| \mu \otimes \nu), \tag{2}$$

where $\mu \otimes \nu$ denotes the product of measures. For the dual problem, we consider the *log-Schrödinger potentials* (Nutz, 2021) $(\varphi^*, \psi^*) \in L^1(\mu) \times L^1(\nu)$, which represent the primal solution with $\mathrm{d}\pi^* = e^{\varphi^* \oplus \psi^* - c_\varepsilon} \mathrm{d}(\mu \otimes \nu)$, $(\mu \otimes \nu)$-almost surely, with the quadratic cost $c_\varepsilon(x, y) \coloneqq \frac{1}{2\varepsilon}\|x - y\|^2$. The Sinkhorn algorithm is given as following alternating updates (Cuturi, 2013):

$$\psi_{2t+1}(y) = -\log \int_{\mathcal{S}} e^{\varphi_{2t}(x) - c_\varepsilon(x,y)} \mu(\mathrm{d}x), \qquad \varphi_{2t+2}(x) = -\log \int_{\mathcal{S}} e^{\psi_{2t+1}(y) - c_\varepsilon(x,y)} \nu(\mathrm{d}y), \tag{3}$$

where each update is called iterative proportional fitting (IPF; Kullback, 1968). Meanwhile, let us consider $W^\varepsilon \in \mathbb{P}(\mathcal{S}, [0, 1])$ be the Wiener process with volatility ε. The *dynamic* SBP aims to find an optimal process $\mathcal{T}^*$ such that

$$\mathcal{T}^* \coloneqq \operatorname*{arg\,min}_{\mathcal{T} \in \mathcal{Q}(\mu,\nu)} \mathrm{KL}(\mathcal{T} \| W^\varepsilon), \tag{4}$$

where $\mathcal{Q}(\mu, \nu) \subset \mathbb{P}(\mathcal{S}, [0, 1])$ is the set of processes with marginals μ and ν. The SB process $\mathcal{T}^*$ is uniquely described by a stochastic differential equation (SDE): $\mathrm{d}X_t = g^*(t, X_t) + \mathrm{d}W_t^\varepsilon$ in $t \in [0, 1)$, governed by a drift function g^* along with some noise. The fundamental equivalence between static and dynamic SBPs (Pavon & Wakolbinger, 1991; Léonard, 2012) allows us to consider the optimal coupling π^* when finding the SB process $\mathcal{T}^*$, vice versa.

Since the SB objective does not ensure the Gâteaux differentiability (Aubin-Frankowski et al., 2022; see Definition 4 in the appendix), one needs an alternative notion of gradients, in order to generalize SBPs from the view of classical MD. Consequently, we provide the definitions of *directional derivatives* (Aliprantis & Border, 2006) and *first variations* (Aubin-Frankowski et al., 2022).

Definition 1 (Directional derivative)**.** Given a locally convex topological vector space $\mathcal{M}$, the directional derivative of F in the direction ξ is defined as $d^+F(x; \xi) = \lim_{h\to 0^+} \frac{F(x+h\xi)-F(x)}{h}$.

Definition 2 (First variation)**.** Given a topological vector space $\mathcal{M}$ and a convex constraint $\mathcal{C} \subseteq \mathcal{M}$, for a function F and $x \in \mathcal{C} \cup \mathrm{dom}(F)$, define the first variation of F over $\mathcal{C}$ to be an element $\delta_c F(x) \in \mathcal{M}^*$, where $\mathcal{M}^*$ is the topological dual of $\mathcal{M}$, such that it holds for all $y \in \mathcal{C} \cup \mathrm{dom}(F)$ and $v = y - x \in \mathcal{M}$: $\langle \delta_c F(x), v\rangle = d^+F(x; v)$. $\langle \cdot, \cdot \rangle$ denotes the duality product of $\mathcal{M}$ and $\mathcal{M}^*$.

Following Karimi et al. (2024), this work considers the generalized Bregman divergence defined with a weak notion of the directional derivative. We explicitly set the Bregman potential $\Omega(\cdot) = \mathrm{KL}(\cdot \| e^{-c_\varepsilon} \mu \otimes \nu)$ in the

SB problems, which enforces the Gibbs parameterization for the couplings, such that SB models follow the specific form of $\mathrm{d}\pi = e^{\varphi \oplus \psi - c_\varepsilon}\mathrm{d}(\mu \otimes \nu)$.

Definition 3 (Bregman divergence)**.** Let a convex functional $\Omega : \mathcal{M} \to \mathbb{R} \cup \{+\infty\}$ be a Bregman potential. Define the Bregman divergence associated with Ω as

$$D_\Omega(x \| y) \coloneqq \Omega(x) - \Omega(y) - d^+\Omega(y; x - y) \tag{5}$$

for every point $x, y \in \mathcal{M}$.

Lastly, our theoretical analysis works with a certain form of *measure concentration* property, and we formally address asymptotically strong log-concave (alc) distributions to address the desired properties of OMD. Let us consider the following informal definition of asymptotically strong log-concave distributions

$$\mathcal{P}_{\mathrm{alc}}(\mathbb{R}^d) \coloneqq \Big\{ \zeta(\mathrm{d}x) = \exp\big(-U(x)\big)\mathrm{d}x \,:\, U \in C_2\big(\mathbb{R}^d\big),\ U \text{ is asymptotically strongly convex} \Big\}, \tag{6}$$

where Appendix A contains a formal version on asymptotical convexity. Note that asymptotically log-concave functions satisfy a certain form of log Sobolev inequality (LSI; Gross, 1975). The condition can be an extension of Sobolev space (Adams & Fournier, 2003) for informational geometric problems. The simplest case of LSI for the Gaussian measure is represented as follows.

Remark 1 (LSI for the standard Gaussian)**.** Suppose that f is a nonnegative function, integrable with respect to a measure γ, and that the entropy is defined as $\mathrm{Ent}_\gamma(f) = \int_{\mathbb{R}^d} f \log f \mathrm{d}\gamma - \big(\int_{\mathbb{R}^d} f \mathrm{d}\gamma\big) \log\big(\int_{\mathbb{R}^d} f \mathrm{d}\gamma\big)$. the log Sobolev inequality when γ is the standard Gaussian measure reads $\mathrm{Ent}_\gamma(f) \leq \frac{1}{2}\int_{\mathbb{R}^d} \frac{|f|^2}{f}\mathrm{d}\gamma$.

Historically, LSI arises from the implication of satisfying the Talagrand's inequality for bounding the Wasserstein-2 distance, and is closely related to measure concentration (Otto & Villani, 2000). The important extension of asymptotically strong log-concave distributions for Schrödinger bridge $\mathrm{d}\pi = e^{\varphi \oplus \psi - c_\varepsilon}\mathrm{d}(\mu \otimes \nu)$, $(\mu \otimes \nu)$-a.s. is that induced SB model also satisfies asymptotically strongly log-concaveness and the LSI condition (Conforti, 2024). For a representative model related to our work, the Gaussian mixture parameterization (Korotin et al., 2024) is a representative model that our theoretical analysis holds, because GMM weights does not alter the characteristic of Fisher information as an upper bound of KL in the asymptotical sense.

Remark 2 (Conforti, 2024)**.** Let $\mu, \nu \in \mathcal{P}_{\mathrm{alc}}(\mathbb{R}^d)$ with finite entropy on Lebesgue measures and $\pi \in \mathcal{C}$ be a coupling in the static Schrödinger bridge problem. Then, for a quadratic cost function, the coupling distribution is also asymptotically log-concave and satisfies a form of logarithmic Sobolev inequality.

Let us suppose that a parameterized SB model $d\pi_t = e^{\varphi_t \oplus \psi_t - c_\varepsilon} d(\mu \otimes \nu)$ obeys the following constraints for marginals and potentials:

$$\mathcal{C} \coloneqq \big\{ \pi : (\mu, \nu) \in \mathcal{P}_2(\mathbb{R}^d) \cap \mathcal{P}_{\mathrm{alc}}(\mathbb{R}^d),\ \ (\varphi, \psi) \in L^1(\mu) \times L^1(\nu),\ \text{ and } \quad \varphi, \psi \in C^2(\mathbb{R}^d) \cap \mathrm{Lip}(\mathcal{K}) \big\}, \tag{7}$$

where $\mathrm{Lip}(\mathcal{K})$ denotes a set of functions with $\mathcal{K}$-Lipschitz continuity. Using the disintegration theorem for probability measures (Léonard, 2014), we assume the boundedness of Bregman divergence between two transport plans using derivatives of first variations with some positive constraint $\omega > 0$ by the following assumption.

Assumption 1 (LSI for couplings)**.** Let us suppose $\Omega = \mathrm{KL}(\pi \| \mathcal{R})$ for a reference measure $\mathcal{R}$. Suppose that arbitrary $\pi, \bar{\pi} \in \mathcal{C}$ satisfy a type of logarithmic Sobolev inequality for relative entropy (KL divergence) is upper bounded by (relative) Fisher information (Gross, 1975), namely $\mathrm{LSI}(\omega)$ for some $\bar{\omega} \in \mathbb{R}_+$ as follows.

$$D_\Omega(\pi \| \mathcal{R}) = \mathrm{KL}(\pi \| \mathcal{R}) \leq \frac{1}{2\bar{\omega}} \iint_{\mathbb{R}^d \times \mathbb{R}^d} \left| \nabla \log \frac{d\pi(x,y)}{d\mathcal{R}(x,y)} \right|^2 \pi(\mathrm{d}x, \mathrm{d}y)$$

where $\Omega = \mathrm{KL}(\cdot \| \mathcal{R})$. By the first variation of KL, equivalence in the first variation of Bregman divergences (explained later in Lemma 5) and an application of the Hölder's inequality, assume that we can find a constant $\omega > 0$ such that

$$D_\Omega(\pi \| \bar{\pi}) \leq \frac{1}{2\omega} \big\| \nabla(\delta_c \Omega(\pi) - \delta_c \Omega(\bar{\pi})) \big\|^2_{L^2(\pi)} \tag{8}$$

for the Bregman potential $\Omega = \mathrm{KL}(\cdot \| e^{-c_\varepsilon} \mu \otimes \nu)$ and the first variation δ_c.

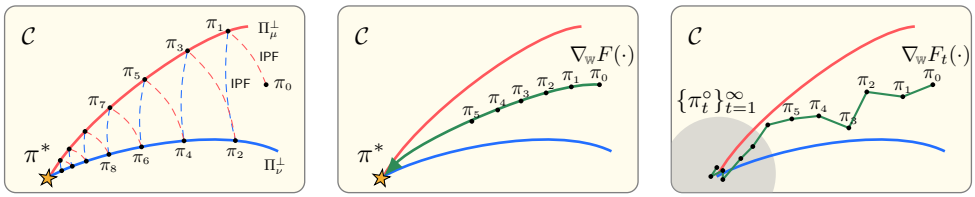

Figure 3: Learning for an SB model $\{\pi_t\}_{t=1}^\infty$ in $\mathcal{C}$ (see Fig. 1 for the detailed illustration). Left: Sinkhorn (Lemma 1). Middle: Wasserstein gradient descent in the distributional space $\mathcal{C}$ for fixed F (Lemma 2). Right: Variational online mirror descent with sequence of convex costs using uncertain estimates $\{\pi_t^\circ\}_{t=1}^\infty$.

In general, the LSI condition also has often been used to analyze the convergence of partial differential equations (Malrieu, 2001). In the same vein, to make an analysis on improvement (Lemma 13) and a solid regret bound of OMD (Lemma 15), this work finds that Assumption 1 is necessary to ensure a certain asymptotical type of measure concentration in the following section.

4 Learning Schrödinger Bridge via Online Mirror Descent

The goal in this section is to derive an OMD update rule for SB, and analyze its convergence. To accomplish this, we postulate on the existence of temporal estimates and an online learning problem. Our analysis suggests that applying an MD approach can reduce the uncertainty of these estimates.

4.1 Sinkhorn and Wasserstein descent as mirror descent algorithms

We start with a novel characterization of Sinkhorn and a static MD variant illustrated in the left side of Fig. 3, which will lead to a better understanding of the OMD framework. OMD updates are determined by the first order approximation of costs F_t and proximity of previous iterate with respect to a Bregman divergence (Beck & Teboulle, 2003). Using the first variation δ_c in Definition 2 instead of standard gradient ∇, the proximal form of OMD is derived as (Karimi et al., 2024)

$$\pi_{t+1} = \arg\min_{\pi \in \mathcal{C}} \Big\{ \big\langle \delta_c F_t(\pi_t), \pi - \pi_t \big\rangle + \tfrac{1}{\eta_t} D_\Omega(\pi \| \pi_t) \Big\}, \tag{9}$$

where F_t denotes a temporal cost function for SB models in $\mathcal{C}$. In Eq. (9), the updates are determined by the first order approximation of F_t and proximity of previous iterate π_t with respect to the Bregman divergence (Beck & Teboulle, 2003). In contrast to the 'half-bridge' interpretation provided by Karimi et al. (2024), the online MD iteration (9) involves a temporal cost F_t, which offers more general reinterpretation of the Sinkhorn algorithm. Using the feasible model space $\mathcal{C}$ in (7), IPF projections (3) are reformulated as following subproblems of alternating Bregman projections:

$$\arg\min_{\pi \in \Pi_\mu^\perp} \big\{ \mathrm{KL}(\pi \| \pi_{2t}) : \pi \in \mathcal{C}, \gamma_2 \pi = \nu \big\}, \qquad \arg\min_{\pi \in \Pi_\nu^\perp} \big\{ \mathrm{KL}(\pi \| \pi_{2t+1}) : \pi \in \mathcal{C}, \gamma_1 \pi = \mu \big\}, \tag{10}$$

where $\gamma_1 \pi(x) := \int \pi(x,y)\,\mathrm{d}y$ and $\gamma_2 \pi(y) := \int \pi(x,y)\,\mathrm{d}x$ denotes the marginalization operations, and the symbols $(\Pi_\mu^\perp, \Pi_\nu^\perp)$ denote the Sinkhorn projection spaces that preserve the property of marginals. As a generalized optimization problem in $\mathcal{C}$, one can consider a temporal cost $\widetilde{F}_t(\pi) := a_t \mathrm{KL}(\gamma_1 \pi \| \mu) + (1 - a_t)\mathrm{KL}(\gamma_2 \pi \| \nu)$ with sequence $\{a_t\}_{t=1}^\infty = \{0, 1, 0, 1, \dots\}$. By construction, an online form of MD for $\widetilde{F}_t$ with a constant step size $\eta_t \equiv 1$ matches the Sinkhorn.

Lemma 1 (Sinkhorn)**.** *For $\Omega(\cdot) = \mathrm{KL}(\pi \| e^{-c_\varepsilon} \mu \otimes \nu)$, iterates from $\pi_{t+1} = \arg\min_{\pi \in \mathcal{C}} \big\{ \langle \delta_c \widetilde{F}_t(\pi_t), \pi - \pi_t \rangle + D_\Omega(\pi \| \pi_t) \big\}$ is equivalent to estimates from (φ_t, ψ_t) of (3), for every update step $t \in \mathbb{N}_0$.*

The proof is in Appendix A. Consequently, we established that the Sinkhorn algorithm corresponds to an instance OMD; however, its inherent structure limits flexibility on step sizes and other underlying assumptions, making it challenging to analyze directly using standard OMD theoretical arguments. Instead, one

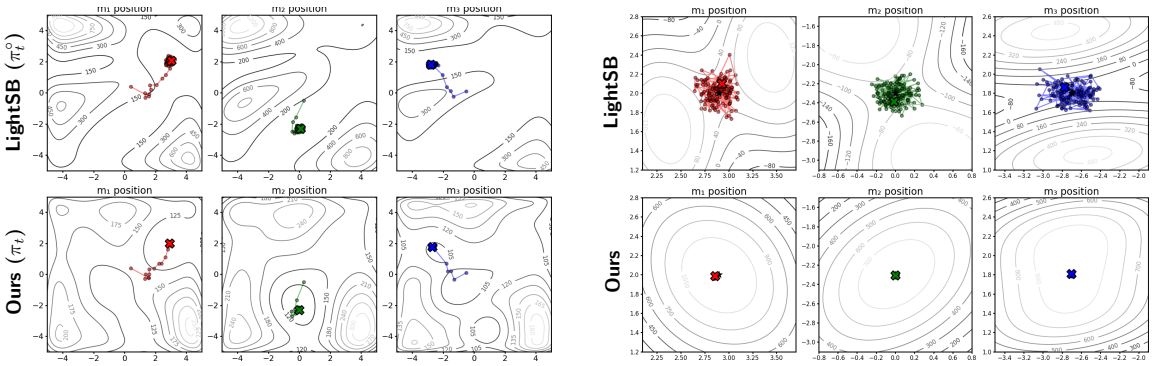

Figure 4: Loss landscapes and gradient dynamics in a 2D problem. Left: In an early stage, parameters of three modalities $\{m_k\}_{k=1}^3$ (mean estimations) for both LightSB (top) and VMSB (bottom) methods approach the optimality with different costs. Right: When magnified the landscapes in the late stages (10 times), while LightSB is vibrant, whereas our method emits strictly convex landscape and stable dynamics.

can alternatively consider OMD by recovering a "static" objective, namely $F(\cdot) \coloneqq \mathrm{KL}(\cdot\|\pi^*)$, where the KL functional is originated from the formal definition of SBP (Vargas et al., 2021; Chen et al., 2022). The following lemma shows that the MD updates for the static cost correspond to discretization of a Wasserstein gradient flow for SB models, a Riemannian steepest descent in the SB model space $\mathcal{C}$.

Lemma 2 (MD in the Wasserstein space)**.** *Suppose that $F(\pi) \coloneqq \mathrm{KL}(\pi\|\pi^*)$ for $\pi \in \mathcal{C}$. The MD formulation of F corresponds to a discretization of a geodesic flow such that $\lim_{\eta_t \to 0^+} \frac{\pi_{t+1}-\pi_t}{\eta_t} = -\nabla_{\mathbb{W}} F(\pi_t)$, where $\nabla_{\mathbb{W}}$ denotes the Wasserstein-2 gradient operator.*

According to Lemma 2, gradients of F are tangential to the geodesic curve from π_0 to π^* (green line in the middle of Fig. 3) in terms of the Wasserstein-2 metric W_2. Hence, the geometric interpretation allows us to consider the static cost F as the ground-truth cost for optimization in our variational OMD framework. However, the ideal case falls short in practice since π^* is inherently unknown. Therefore, we postulate on an online learning problem that nonstationary estimates $\{\pi_t^\circ\}_{t=1}^\infty$ are offered instead of π^* as learning signals, making an optimization process with $F_t(\cdot) \coloneqq \mathrm{KL}(\cdot\|\pi_t^\circ)$. As shown in in the right side of Fig. 3, we focus on the online learning setting where $\{\pi_t^\circ\}_{t=1}^\infty$ are fundamentally uncertain with perturbation, since the optimal plan π^* is not accessible during the training time.

4.2 Online mirror descent for Schrödinger bridges: theoretical analysis

We postulate on an online learning problem that nonstationary ergodic estimates $\{\pi_t^\circ\}_{t=1}^\infty$ are offered instead of π^* (gray region in Fig. 1). Let Ω^* be the Fenchel conjugate of $\Omega + i_{\mathcal{C}}$ with the convex indicator[1] $i_{\mathcal{C}}$. For the space $\mathcal{D} \coloneqq \delta_{\mathcal{C}}\Omega(\mathcal{C})$, a directional derivative $\delta_{\mathcal{D}}$ of Ω^* exists by the Danskin's theorem (Danskin, 1967; Bernhard & Rapaport, 1995), such that

$$\delta_{\mathcal{D}}\Omega^*(\varphi \oplus \psi) = \arg\max_{\pi \in \mathcal{C}} \big\{ \langle \varphi \oplus \psi, \pi \rangle - \Omega(\pi) \big\}. \tag{11}$$

Note that $(\delta_{\mathcal{C}}\Omega, \delta_{\mathcal{D}}\Omega^*)$ form bidirectional maps; a direct sum of potentials $\varphi \oplus \psi \in \mathcal{D}$ represent an element of the generalized dual space. The key assumption is that the learning target π_t° is asymptotically mean stationary (Gray & Kieffer, 1980) for the dual space, which have been used to analyze stochastic dynamics. Since iterates are updated through dual parameters in MD, we refer to the process as being dually stationary.

Assumption 2 (Dually stationary process)**.** Suppose that $\pi_{\mathcal{D}}^\circ \in \mathcal{C}$ exists, which is a primal representation of an asymptotic mean $\pi_{\mathcal{D}}^\circ \coloneqq \delta_{\mathcal{D}}(\lim_{t\to\infty} \mathbb{E}_t[\delta_{\mathcal{C}}\Omega(\pi_t^\circ)])$. Here $\mathbb{E}_t$ denotes the time-average.

Plugging $F_t(\cdot) = \mathrm{KL}(\cdot\|\pi_t^\circ)$ to (9) from § 4.1, we achieve a distinct OMD setup. Fig. 4 demonstrates toy experiment regarding our online learning hypothesis. OMD decomposes the global problem into local convex

[1] Defined as $i_{\mathcal{C}}(x) = 0$ if $x \in \mathcal{C}$ and $+\infty$ otherwise.

problems, and prevented iterates from being vibrant by stopping at a single point $\pi_{\mathcal{D}}^{\circ}$. This verifies that OMD stabilizes learning of π_t, even when the reference π_t° tends to inherently have some perturbation.

For stability, we state two conditions for OMD step sizes $\{\eta_t\}_{t=0}^{\infty}$, which will be justified in Theorem 1 and Proposition 1. Fig. 5 shows a plot of a well-known example: harmonic progression $\frac{1}{a+td}$ for $a \in \mathbb{R}$ and $d \in \mathbb{R}^+$ with respect to t. The sequence satisfies both conditions of the following assumption.

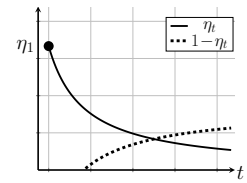

Figure 5: A sequence example of η_t and $1-\eta_t$.

Assumption 3 (Step sizes)**.** Assume two conditions for step sizes $\{\eta_t\}_{t=0}^{\infty}$. (a) *Convergent sequence & divergent series:* $\lim_{t\to\infty}\eta_t = 0$ and $\sum_{t=1}^{\infty}\eta_t = \infty$. (b) *Convergent series for squares:* $\sum_{t=1}^{\infty}\eta_t^2 < \infty$.

Using the conditions above, we firstly argue that OMD for the temporal cost $\mathrm{KL}(\cdot\|\pi_t^{\circ})$ with respect to the Bregman potential $\Omega = \mathrm{KL}(\cdot\|e^{-c_\varepsilon}\mu\otimes\nu)$ requires step size scheduling for the sake of convergence. The following theorem states that convergence in the case of $\pi^* = \pi_{\mathcal{D}}^{\circ}$ is assured when η_t follows Assumption (3a). In contrast to the well-known linear convergence guarantees for the Sinkhorn algorithm under bounded costs with fixed marginals (Carlier, 2022), our OMD-based analysis establishes sublinear convergence rates, accommodating scenarios involving unbounded and non-stationary costs.

Theorem 1 (Step size considerations)**.** *Suppose the idealized case of $\pi^* = \pi_{\mathcal{D}}^{\circ}$. Then, for $\{\pi_t\}_{t_1}^{T} \subset \mathcal{C}$ we get $\lim_{T\to\infty}\mathbb{E}_{1:T}[D_\Omega(\pi_{\mathcal{D}}^{\circ}\|\pi_T)] = 0$ if and only if Assumption (3a) is satisfied. Furthermore, if the step size is in the form of $\eta_t = \frac{2}{t+1}$, then $\mathbb{E}_{1:T}[D_\Omega(\pi^*\|\pi_t)] = \mathcal{O}(1/T)$.*

Therefore, we can guarantee the ideal convergence in the SB learning when the scheduling of η_t follows the step size assumptions. Next, we argue that general convergence toward $\pi_{\mathcal{D}}^{\circ}$ is guaranteed under Assumption (3b). Given the convex nature of SB cost functionals, we argue that this convergence toward $\pi_{\mathcal{D}}^{\circ}$ is beneficial as long as π_t° is trained to approximate π^* and remain bounded. Therefore, we argue that the convergence of SB is beneficial and address the following statement.

Proposition 1 (Convergence)**.** *Suppose that $\pi^* \neq \pi_{\mathcal{D}}^{\circ}$, hence $\inf_{\pi\in\mathcal{C}}\mathbb{E}[F_t(\pi)] > 0$. If the step sizes $\{\eta_t\}_{t=0}^{\infty}$ satisfies Assumption 3, then $\lim_{t\to\infty}\mathbb{E}_{1:t}[D_\Omega(\pi_{\mathcal{D}}^{\circ}\|\pi_t)]$ converges to 0 almost surely.*

Lastly, assume that a type of log Sobolev inequality holds (see Assumption 1) with continuity of potentials. We establish an online learning regret bound of $\mathcal{O}(\sqrt{T})$ for certain instance of step sizes, demonstrating that imposing specific measure-theoretic properties in SBPs generalizes classical OMD results (Nesterov, 2009; Srebro et al., 2011; Orabona & Pál, 2018; Lei & Zhou, 2020). We show that analysis on OMD is compatible using the dual norm $\|\hat{g}_t\|$ induced by the first variation $\delta_{\mathcal{C}}$.

Proposition 2 (Regret bound)**.** *Assume $\varphi, \psi \in C^2(\mathbb{R}^d)\cap\mathrm{Lip}(\mathcal{K})$, Assumption 1 holds, and the given costs $\{F_t\}_{t=1}^{T}$ are bounded. For arbitrary $u \in \mathcal{C}$ and a total step T, define $D^2 = \max_{1\le t\le T} D_\Omega(u\|\pi_t)$. (a) When the number of time step is known a priori, the regret is bounded to $2D\sqrt{2\omega^{-1}\mathcal{K}T}$ for a constant step size $\eta \equiv \frac{D\sqrt{\omega}}{\sqrt{2\mathcal{K}T}}$. (b) For an adaptive scheduling $\eta_t = D\sqrt{\omega}/\sqrt{2\sum_{i=1}^{t}\|\hat{g}_i\|^2}$ the regret is bounded to $2D\sqrt{2\omega^{-1}\sum_{t=1}^{T}\|\hat{g}_t\|^2}$ where $\hat{g}_t = \delta_{\mathcal{C}}\Omega(\pi_t) - \delta_{\mathcal{C}}\Omega(\pi_t^{\circ})$.*

Note that although our analysis establishes a rigorous connection between SB and OMD, it inherits certain limitations from classical OMD regret studies. For instance, sublinear regrets in Proposition 2 relies on an additional boundedness assumption on costs, and there exist some cases of Assumption 3 that may yield asymptotically linear regret (Orabona & Pál, 2018). Addressing these limitations may involve advanced hybrid OMD methodologies which are actively being studied, such as dual averaging (Fang et al., 2022) or FTRL (Chen & Orabona, 2023). As exploring (as well as computing) such extensions for SBPs falls beyond our scope, from now on we focus on practical implementations of OMD that demonstrate theoretical convergence (Proposition 1) over a relatively long horizon of several hundred steps. Hence, we adopt Assumption 3 and provide corresponding experimental evidence to substantiate its validity in practical scenarios.

4.3 Online mirror descent updates using Wasserstein gradient flows

The remaining issue is challenge is practicality of OMD, where the computation of critical points for (9). To resolve this issue, we present a new approximation method using Wasserstein gradient flows (Jordan et al.,

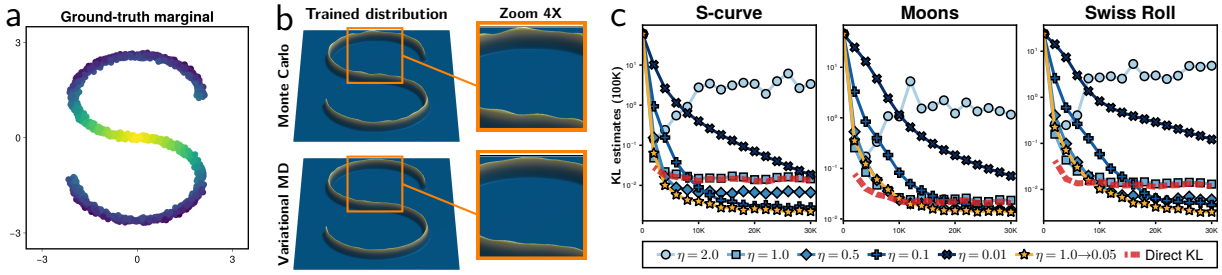

Figure 6: Variational MD with synthetic datasets. (a) A distribution is accessible by finite batch data. (b) 3D surfaces of $(\vec{\pi}_T^\circ, \vec{\pi}_T)$ trained by Monte Carlo method for KL (top) and variational MD (bottom) show that the MD results in more stable outcomes. (c) The plots show the estimated $\mathrm{KL}(\vec{\pi}_t \| \vec{\pi}^*)$ with different step size scheduling (5 runs), with red dashed baselines $\mathrm{KL}(\vec{\pi}_t^\circ \| \vec{\pi}^*)$.

1998). Suppose we expand a subinterval $[t, t+1)$ for each OMD step into continuous dynamics of $\rho(\tau) \in \mathcal{C}$ for a $\tau \in [0, \infty)$. By Otto's calculus on the Wasserstein space (Otto, 2001), it is well known that one can describe the dynamics of minimizing a convex functional $\mathcal{E}_t : \mathcal{C} \to \mathbb{R}$ as the following PDE

$$\partial_\tau \rho_\tau = -\nabla_{\mathbb{W}} \mathcal{E}_t(\rho), \tag{12}$$

where $\nabla_{\mathbb{W}}$ denotes the Wasserstein-2 gradient operator $\nabla_{\mathbb{W}} := \nabla \cdot \big(\rho \nabla \frac{\delta}{\delta \rho}\big)$. In this work, we adopt the Wasserstein gradient flow theory (Jordan et al., 1998) to efficiently perform OMD where the equilibrium indicates the subsequent iterate π_{t+1}. Note that Wasserstein gradient flows are asymptotically stable by the LaSalle's invariance principle (Carrillo et al., 2023). We present a simple and exact closed-form expression for the VOMD update. Note that the cost $F_t(\cdot) = \mathrm{KL}(\cdot \| \pi_t^\circ)$ satisfies the 1-relative-smoothness and 1-strong-convexity relative to Ω (see Definition 6; Aubin-Frankowski et al., 2022). Then, a first variation of the OMD problem can be decomposed into multiple variations of another problem with similar characteristics (*e.g.*, equilibrium, smoothness, and convexity). We present the following theorem for the computation of OMD.

Theorem 2 (Dynamics equivalence in first variation)**.** *Consider the Wasserstein gradient dynamics of* (12) *which solves a local update of (9). The gradient dynamics of updates are equivalent to that of a linear combination of KL functionals such that for any* $\rho_\tau \in \mathcal{C}$

$$\eta_t \delta_c \mathcal{E}_t(\rho_\tau) = \delta_c \big\{ \eta_t \mathrm{KL}(\rho_\tau \| \pi_t^\circ) + (1 - \eta_t) \mathrm{KL}(\rho_\tau \| \pi_t) \big\} \quad \forall \rho_\tau \in \mathcal{C}, \tag{13}$$

and the PDE converges to a unique critical point of subsequent OMD iterate (9) *as* $\tau \to \infty$.

Sketch of Proof. We identify $\delta \mathcal{E}_t$ as a dynamics that reaches an equilibrium solution for

$$\begin{aligned} &\underset{\pi \in \mathcal{C}}{\text{minimize}} \, \big\langle \delta_c F_t(\pi_t), \pi - \pi_t \big\rangle + \tfrac{1}{\eta_t} D_\Omega(\pi \| \pi_t) \\ \iff \quad &\underset{\pi \in \mathcal{C}}{\text{minimize}} \; \eta_t \underbrace{D_\Omega(\pi \| \pi_t^\circ)}_{\text{empirical estimates}} + (1 - \eta_t) \underbrace{D_\Omega(\pi \| \pi_t)}_{\text{proximity}}, \end{aligned} \tag{14}$$

and then the equivalence of first variation for recursively defined Bregman divergences is applied (Lemma 5). At a glance, Eq. (14) appears analogous to the interpolation search between two points, where the influence of π_t° is controlled by η_t. We leave the entire proof in Appendix A.5. □

We argue that Theorem 2 holds practical importance for OMD computation, since following the argument allows us to perform gradient-based updates without directly constructing a desired Bregman divergence. That is, updates can be drawn based on a linear combination of gradient flows $\eta_t \nabla_{\mathbb{W}} \mathrm{KL}(\rho_\tau \| \pi_t^\circ) + (1 - \eta_t) \nabla_{\mathbb{W}} \mathrm{KL}(\rho_\tau \| \pi_t)$, where such expression has been extensively studied both theoretically and computationally (Carrillo et al., 2023; Lambert et al., 2022). Therefore, we can utilize interpolation of Wasserstein gradient flows for performing updates and utilize a certain variational class for reducing the computational cost. Fig. 6 shows our actual experiments using Gaussian mixture models (GMMs). Let a reference estimation

be fitted using a Monte Carlo method, and our model be trained through a variational OMD method which is explained in the subsequent section. We initially observed that the VOMD method provides stability improvement when $\eta < 1$. In contrast, the condition of $\eta > 1$ performed worse than the Monte Carlo method and $\eta = 1$ showed almost equivalent performance. Furthermore, the performance of VOMD was greatly improved by choosing a harmonic step size scheduling in the interval $[1.0, 0.05]$. All of these results on variational approximation precisely matches our analysis.

5 Algorithm: Variational Mirrored Schrödinger Bridge

In this section, we propose variational mirrored Schrödinger bridge, a simulation-free method that offers iterative MD updates for parameterized SB models with mixture models, using the Wasserstein-Fisher-Rao geometry. We provide a variational interpretation for LightSB models and draw a practical VOMD updates algorithm that closely resembles ordinary machine learning methods.

5.1 Gaussian mixture parameterization for the Schrödinger bridge problem

Recently, Korotin et al. (2024) proposed the GMM parameterization, which provides theoretically and computationally desirable models for our variational OMD approach. The parameterization considers the *adjusted* Schrödinger potential $u^*(x) \coloneqq \exp(\varphi^*(x) - \|x\|^2/2\varepsilon)$ and $v^*(y) \coloneqq \exp(\psi^*(y) - \|y\|^2/2\varepsilon)$ such that we have a proportional property $\pi^*(y|x) \propto \exp(\langle x,t\rangle/\varepsilon) v^*(y)$. With a finite set of parameters $\theta \triangleq \{\alpha_k, m_k, \Sigma_k\}_{k=1}^K$ for weights $\alpha_k > 0$, means $m_k \in \mathbb{R}^d$ and covariances $\Sigma_k \in \mathbf{S}_{++}^d$, Korotin et al. (2024) proposed to approximate the adjusted Schrödinger potential v_θ and conditional probability density $\vec{\pi}_\theta$

$$v_\theta(y) \coloneqq \sum_{k=1}^K \alpha_k \mathcal{N}(y|m_k, \varepsilon\Sigma_k), \qquad \vec{\pi}_\theta^x(y) \coloneqq \frac{1}{z_\theta^x} \sum_{k=1}^K \alpha_k^x \mathcal{N}(y|m_k^x, \varepsilon\Sigma_k), \tag{15}$$

where GMM component for $\vec{\pi}_\theta^x$ is conditioned by an input x: $m_k^x \coloneqq m_k + \Sigma_k x$, $\alpha_k^x \coloneqq \alpha_k \exp\big(\frac{x^\top \Sigma_k x + \langle m_k, x\rangle}{2\varepsilon}\big)$, $z_\theta^x \coloneqq \sum_{k=1}^K \alpha_k^x$ (see Proposition 3.2 of Korotin et al.). For this parameterization, the closed-from expression of SB process $\mathcal{T}_\theta$ is given as the following SDE for $t \in [0, 1)$:

$$\begin{gathered} \mathcal{T}_\theta : \mathrm{d}X_t = g_\theta(t, X_t)\,\mathrm{d}t + \sqrt{\varepsilon}\,\mathrm{d}W_t, \\ g_\theta(t,x) \coloneqq \varepsilon \nabla \log \mathcal{N}(x|0, \varepsilon(1-t)I_d) \sum_{k=1}^K \alpha_k \mathcal{N}(m_k|0, \varepsilon\Sigma_k) \mathcal{N}\big(m_k(t,x)\big|0, A_k(t)\big), \end{gathered} \tag{16}$$

where $m_k(t,x) \triangleq \frac{x}{\varepsilon(1-t)} + \frac{1}{\varepsilon}\Sigma_k^{-1} m_k$ and $A_k(t) \triangleq \frac{t}{\varepsilon(1-t)} I_d + \frac{1}{\varepsilon}\Sigma_k^{-1}$. Therefore, the LightSB parameterization represent both static and dynamic SB models and arbitrary SB solvers can be applied without restrictions. We utilize the GMM parameterization for our computational algorithm for three key reasons. Firstly, the parameterization induce the universal approximation property for both $\vec{\pi}_\theta$ and $\mathcal{T}_\theta$ (Korotin et al., 2024). Secondly, GMMs are asymptotically log-concave (see Lemma 3), which is a fundamental assumption for our theory. Lastly, the parameterization makes the computation of WGFs with respect to the KL divergence tractable, which is delineated in the subsequent section.

5.2 Computation of VOMD in the WFR geometry

The space of Gaussian parameters $\mathbb{R}^d \times \mathbf{S}_{++}^d$, endowed with the Wasserstein-2 metric W_2, is formally recognized as the Bures–Wasserstein (BW) geometry (Bures, 1969; Bhatia et al., 2019; Lambert et al., 2022) $\mathtt{BW}(\mathbb{R}^d) \subseteq \mathcal{P}_2(\mathbb{R}^d)$. Wasserstein-Fisher-Rao geometry, equivalently characterized by the spherical Hellinger–Kantorovich distance, extends this setting by considering liftings of positive, complete, and separable measures while preserving total mass (Liero et al., 2018; Chizat et al., 2018; Lu et al., 2019). Building upon the BW space, the Wasserstein-Fisher-Rao geometry of GMMs, namely $\mathcal{P}_2(\mathtt{BW}(\mathbb{R}^d))$, naturally provides liftings of Gaussian particles satisfying distributional consistency. In this work, we introduce the following proposition, which refines and extends the results from Lambert et al. (2022, § 6) specifically enhancing their framework through the introduction of freely trainable GMM weights α_k.

Algorithm 1 Variational Mirrored SB (VMSB).

Input: SB models $(\vec{\pi}_\theta, \vec{\pi}_\phi)$ parameterized by Gaussian mixtures, step sizes (η_1, η_T), $n_y, B \in \mathbb{N}$.

1: **for** $t \leftarrow 1$ **to** T **do**
2: Acquire ϕ_t with an external data-driven SB solver.
3: $\theta_t \leftarrow \theta$, $\eta_t \leftarrow 1/\big(\eta_1^{-1} + (\eta_T^{-1} - \eta_1^{-1})(t-1/T-1)\big)$
4: **for** $n \leftarrow 1$ **to** N **do**
5: $\{x_i\}_{i=1}^B \leftarrow$ sample mini batch data from μ.
6: $\frac{\partial \mathcal{L}}{\partial \theta} \leftarrow \frac{1}{B}\sum_{i=1}^B \eta_t \texttt{WFRgrad}(\theta; \phi_t, x_i, n_y) + (1-\eta_t)\texttt{WFRgrad}(\theta; \theta_t, x_i, n_y)$
7: Update θ with the gradient $\frac{\partial \mathcal{L}}{\partial \theta}$.
8: **end for**
9: **end for**

Output: Trained SB model $\vec{\pi}_\theta$.

Proposition 3 (WFR gradient dynamics)**.** *Suppose a GMM ρ_{θ_τ} with $\theta_\tau = \{\alpha_{k,\tau}, m_{k,\tau}, \Sigma_{k,\tau}\}_{k=1}^K$. Let $y_{k,\tau} \sim \mathcal{N}(m_{k,\tau}, \Sigma_{k,\tau})$ denote a sample from the k-th Gaussian particle of ρ_{θ_τ}. Then, the WFR dynamics $\nabla_{\texttt{WFR}}\mathrm{KL}(\rho_{\theta_\tau}\|\rho^*)$ wrt $\dot{\theta}_\tau = \{\dot{\alpha}_{k,\tau}, \dot{m}_{k,\tau}, \dot{\Sigma}_{k,\tau}\}_{k=1}^K$ are given as*

$$\begin{aligned} \dot{\alpha}_{k,\tau} &= -\left(\mathbb{E}\left[\log \frac{\rho_{\theta_\tau}}{\rho^*}(y_{k,\tau})\right] - \frac{1}{z_\tau}\sum_{\ell=1}^K \alpha_\ell \mathbb{E}\left[\log \frac{\rho_{\theta_\tau}}{\rho^*}(y_{\ell,\tau})\right]\right)\alpha_{k,\tau}, \\ \dot{m}_{k,\tau} &= -\mathbb{E}\left[\nabla \log \frac{\rho_{\theta_\tau}}{\rho^*}(y_{k,\tau})\right], \quad \dot{\Sigma}_{k,\tau} = -\mathbb{E}\left[\nabla^2 \log \frac{\rho_{\theta_\tau}}{\rho^*}(y_{k,\tau})\right]\Sigma_{k,\tau} - \Sigma_{k,\tau}\mathbb{E}\left[\nabla^2 \log \frac{\rho_{\theta_\tau}}{\rho^*}(y_{k,\tau})\right], \end{aligned} \tag{17}$$

for $\tau \in [0, \infty)$, where $z_\tau \coloneqq \sum_{k=1}^K \alpha_k$; ∇ and ∇^2 denote gradient and Hessian with respect to $y_{k,\tau}$.

Appendices A.6 and B contain the complete theory. Proposition 3 argues that the one parameter family θ_τ predicts a gradient-based algorithm of $\nabla_{\texttt{WFR}}\mathrm{KL}(\rho_{\theta_\tau}\|\rho^*)$, and thus Eq. (17) can be directly used for training GMM models. Recall that GMMs have a closed form expression of likelihoods, which means each log likelihood difference can be calculated without errors. Given that the target has the identical number of Gaussian particles, both Eq. (17) and its approximation using finite samples strictly induce zero gradients at the equilibrium. Hence, we argue that the simulation-free algorithm VMSB will result in more robust and stable outcomes than standard data-driven SB learning.

5.3 Algorithmic considerations

Algorithm 1 outlines the overall procedure. VMSB requires SB parameters θ and ϕ, which represents $\vec{\pi}_t$ and $\vec{\pi}_t^\circ$ from the theoretical framework in § 4.2. The target model $\vec{\pi}_\phi$ is independently fitted using an arbitrary SB solver. By the results of analysis, one can schedule of the step size η_t with a harmonic progression satisfying Assumption 3; thus, we propose to schedule by the series for $1 \geq \eta_1 \geq \eta_T > 0$ as in Line 3 of the algorithm. In our settings, the hyperparameters are set $\eta_1 = 1$ and $\eta_T \in \{0.05, 0.01\}$ which varies depending on each length of training. The algorithm can also put "warm up" steps leveraged by a existing solver, and start from $\theta = \phi_t$ enforcing $\eta_t \equiv 1$ for a certain period of the early stage.

The VMSB algorithm is essentially designed to perform the following approximation of the WFR gradient operation (17) in Proposition 3, approximated with finite data samples $\{x_i\}_{i=1}^B \sim \mu$

$$\frac{1}{B}\sum_{i=1}^B \texttt{WFRgrad}(\theta; \phi, x_i, n_y) \approx \nabla_{\texttt{WFR}}\,\mathrm{KL}(\vec{\pi}_\theta \| \vec{\pi}_\phi),$$

where each expectation is estimated using n_y samples from each Gaussian particle. Following Theorem 2, we propose to update the SB model $\vec{\pi}_\theta$ with $\eta_t\texttt{WFRgrad}(\theta; \phi_t, x_i, n_y) + (1-\eta_t)\texttt{WFRgrad}(\theta; \theta_{t-1}, x_i, n_y)$ at each VOMD iteration t (see Line 6). When μ is a zero-centered distribution, we set $B = 1$ and $x = 0$ for the fast training time. This trick is equivalent to training the adjusted Schrödinger potential (Korotin et al., 2024) $v_\theta \coloneqq \sum_{k=1}^K \alpha_k \mathcal{N}(y|m_k, \varepsilon\Sigma_k) \propto \pi_\theta(\cdot|x=0)$ directly, which makes the VMSB algorithm run efficiently for certain tasks.

6 Experimental Results

Experiment goals. We aimed to test our online learning hypothesis and verify that the VMSB effectively induces OMD updates. Since our theoretical claims are intended to be highly versatile, consistent performance improvements for each setting coincides with the generality of the proposed VOMD method. We delineate our objectives as follows: ① We aimed to affirm our online learning hypothesis by demonstrating consistent improvements. ② We sought to corroborate our theoretical results, aiming for stable performance that consistently exceeds that of benchmarks. ③ We aimed to verify that our algorithm effectively induces OMD by the Wasserstein gradient flow.

Baselines and VMSB variants. Korotin et al. (2024) introduced a streamlined, simulation-free solver referred to as **LightSB** that optimizes ϕ through Monte Carlo approximation of $\mathrm{KL}(\vec{\pi}^* \| \vec{\pi}_\phi)$. As an alternative, **LightSB-M** (Gushchin et al., 2024a) reformulated the reciprocal projection from DSBM (Shi et al., 2023) to a projection method termed *optimal projection*, establishing approximated bridge matching for the path measure $\mathcal{T}_\phi$. Applying Algorithm 1, we derived two distinct methods called **VMSB** and **VMSB-M** ($\vec{\pi}_\theta$), trained upon LightSB and LightSB-M solvers ($\vec{\pi}_\phi$), respectively. Since the theoretical arguments imply that the algorithm is agnostic to targets, the performance benefits of VMSB variants from their references support the generality of our claims. Additionally, we adopted VMSB on *hybrid* settings, leveraging networks or embeddings for complex problems. We considered such technique for as a part of the problem and applied VMSB on top of the settings.

6.1 Online SB learning for synthetic data streams

To validate our online learning hypothesis, we considered 2D SBPs for data streams depicted in Fig. 7 (a). We applied an angle-based rotating filter, making the marginal as a data stream where only 12.5% (or 45-degree angle) of the total data is accessible for each step t. We trained conditional models $\vec{\pi}_\theta$ for ordinary SB for the 2D coordinates. Fig. 7 (b) shows the plots of squared energy distance (ED), which is a special instance of squared maximum mean discrepancy (MMD), approximating the L^2 distance between distributions: $\mathrm{ED}(P, Q) = \int (P(x) - Q(x))^2 \mathrm{d}x$ (Rizzo & Székely, 2016). In our ED evaluation, the VMSB algorithm achieved a strictly lower divergence than the LightSB and LightSB-M solvers for various numbers of Gaussian particles K. Therefore, we concluded that these results aligned with our hypothesis and theory of online mirror descent.

6.2 Quantitative Evaluation

EOT benchmark. Next, we considered the EOT benchmark proposed by Gushchin et al. (2024b), which contains 12 entropic OT problems with different volatility and dimensionality settings. Table 2 shows that LightSB and VMSB methods outperforms other method in terms of the cBW_2^2-UVP metric as previously reported by Korotin et al. (2024) and Gushchin et al. (2024a). We also observed that a hybrid approach combining LightSB and the exponential moving average (EMA; Morales-Brotons et al., 2024) named as

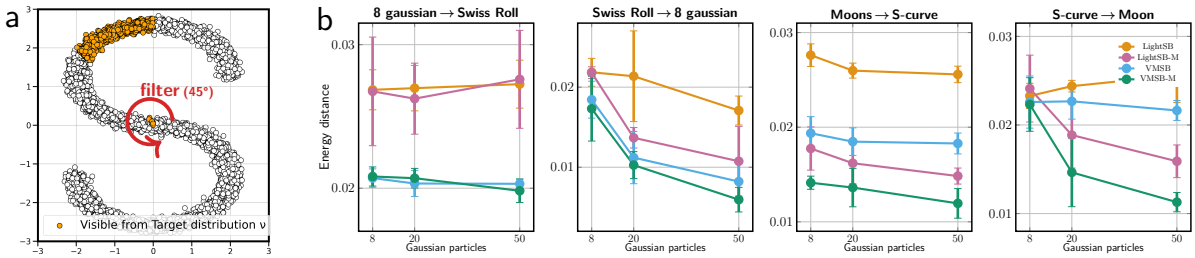

Figure 7: Online SBPs for synthetic dataset streams. (a) We designed an online learning problem with a rotating filter where an algorithm is allowed to observe the data in $y \sim \nu$ only 12.5% at a time. (b) The plots show that our VMSB and VMSB-M show consistent improvements from their references regarding the ED metric with 95% confidence intervals for 5 runs with different seeds.

Table 2: EOT benchmark scores with cB$\mathbb{W}_2^2$-UVP $\downarrow$ (%) between the optimal plan π^* and the learned plan π_θ (five runs). Results of classical EOT solvers marked with † are taken from (Korotin et al., 2024), and ‡ from (Gushchin et al., 2024a). Additionally, LightSB-EMA indicates a hybrid approach using the exponential moving average techniques (EMA; Morales-Brotons et al., 2024) for LightSB parameters ($decay = 0.99$). Our VMSB and VMSB-M results are highlighted in bold when VMSB methods exceed their reference algorithm.

Type	**Solver**	$\varepsilon = 0.1$				$\varepsilon = 1$				$\varepsilon = 10$			
		$d = 2$	$d = 16$	$d = 64$	$d = 128$	$d = 2$	$d = 16$	$d = 64$	$d = 128$	$d = 2$	$d = 16$	$d = 64$	$d = 128$
Classical solvers (best; Korotin et al.)†		1.94	13.67	11.74	11.4	1.04	9.08	18.05	15.23	1.40	1.27	2.36	1.31
Bridge-M	DSBM (Shi et al.)‡	5.2	10.8	37.3	35	0.3	1.1	9.7	31	3.7	105	3557	15000
Bridge-M	SF^{2}M-Sink (Tong et al.)‡	0.54	3.7	9.5	10.9	0.2	1.1	9	23	0.31	4.9	319	819
rev. KL	LightSB (Korotin et al.)	0.007	0.040	0.100	0.140	0.014	0.026	0.060	0.140	0.019	0.027	0.052	0.092
Bridge-M	LightSB-M (Gushchin et al.)	0.017	0.088	0.204	0.346	0.020	0.069	0.134	0.294	0.014	0.029	0.207	0.747
EMA	LightSB-EMA	0.005	0.040	0.078	0.149	0.012	0.022	0.051	0.127	0.017	0.021	0.025	0.042
Var-MD	VMSB (ours)	**0.004**	**0.012**	**0.038**	**0.101**	**0.010**	**0.018**	**0.044**	**0.114**	**0.013**	**0.019**	**0.021**	**0.040**
Var-MD	VMSB-M (ours)	**0.015**	**0.067**	**0.108**	**0.253**	**0.010**	**0.019**	**0.094**	**0.222**	**0.013**	**0.029**	**0.193**	0.748

LightSB-EMA was affective for improving stability. Among 24 different settings, our MD approach exceeded the reference model and the EMA method in 23 settings in terms of the cB$\mathbb{W}_2^2$-UVP metric (Gushchin et al., 2024b). Our replication of LightSB/LightSB-M achieved better performance than originally reported results, and our method accordingly reached the state-of-the-art performance in this benchmark with stability, which represents strong evidence of Proposition 1. Among all cases, the only exception was LightSB-M, which had the highest dimension and volatility. We suspected that the drift form Eq. (16), which is proportional to ε, may have violated our assumptions Assumption 2 and the boundedness assumption during the training. Thus, we conclude that our variational MD training is effective in various EOT setups.

SB on single cell dynamics. We evaluated VMSB on unpaired single-cell data problems in the high-dimensional single cell dynamics experiment (Tong et al., 2024a). The dataset provided single cell data from four donors on days 2, 3, 4, and 7, describing the gene expression levels of distinct cells. Given samples collected on two different dates, the task involves performing inference on temporal evolution, such as interpolation and extrapolation of PCA projections with $\{50, 100, 1000\}$ dimensions. Table 3 shows that our VMSB method achieved the best results, verifying that its effectiveness for the real-world biological problems.

Table 3: Energy distance on the MSCI dataset (95% confidence interval, ten trials with different instances). Results marked with ‡ are from (Gushchin et al., 2024a).

Type	**Solver**	$d = 50$	$d = 100$	$d = 1000$
Sinkhorn	Vargas et al. (2021)†	2.34	2.24	1.864
Bridge-M	DSBM (Shi et al.)‡	2.46 ± 0.1	2.35 ± 0.1	1.36 ± 0.04
Bridge-M	SF^{2}M-Sink (Tong et al.)‡	2.66 ± 0.18	2.52 ± 0.17	1.38 ± 0.05
rev. KL	LightSB	2.31 ± 0.08	2.15 ± 0.09	1.264 ± 0.06
Bridge-M	LightSB-M	2.30 ± 0.08	2.15 ± 0.08	1.267 ± 0.06
Var-MD	VMSB (ours)	$\mathbf{2.28 \pm 0.09}$	$\mathbf{2.13 \pm 0.09}$	$\mathbf{1.260 \pm 0.06}$
Var-MD	VMSB-M (ours)	$\mathbf{2.26 \pm 0.10}$	$\mathbf{2.12 \pm 0.09}$	$\mathbf{1.265 \pm 0.05}$

6.3 Unpaired image-to-image transfer

MNIST-EMNIST. We applied VMSB to unpaired image translation tasks for MNIST and EMNIST datasets. In these tasks, LightSB methods struggled to generate raw pixels due to the limited scalability of the loss function. To solve this issue, we opted to find a viable alternative to LightSB the raw pixel space, and we discovered that the capabilities of GMM parameterization can be extended by incorporating the adversarial learning technique (Goodfellow et al., 2014; see Appendix C.5) was effective in providing rich learning signals for π_ϕ. Therefore, we named the adversarial method and the VMSB adaptation **LightSB-adv** and **VMSB-adv**. Also, we pretrained encoder networks using the *Adversarial Latent AutoEncoder* (ALAE; Pidhorskyi et al., 2020) technique, and applied the LightSB and VMSB algorithms on the 128-dimensional latent space that represent the both of data. Fig. 8 shows that VMSB/VMSB-adv outperformed Light/LightSB-adv (with identical architecture) in the fidelity of samples and semantics of letters for latent and pixel spaces. In Table 4, the VMSB method on the ALAE embedding space was able to surpass deep SB models with a fewer number of parameters of $K = 256$. Even for raw pixels, our algorithm also achieved

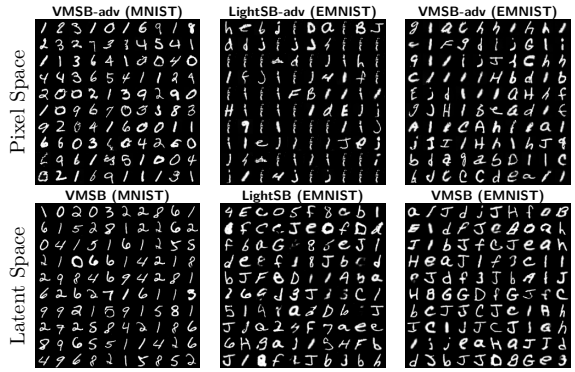

Figure 8: Generated MNIST/EMNIST samples. Top: Raw pixel SB results. Bottom: Latent SB results.

Table 4: FID and MSD scores in EMNIST-to-MNIST translation tasks. Hyperparameters between LightSB and VMSB are shared. We examined the scores with five runs for the ALAE case.

	Method	FID	MSD
U-net	SF^{2}M-Sink	23.215	0.456
	DSBM-IPF	15.211	0.352
	DSBM-IMF	11.429	0.373
Pixel	LightSB-adv	20.017	0.362
	VMSB-adv (ours)	15.471	0.356
ALAE	LightSB	$9.183_{\pm 0.569}$	$0.371_{\pm 0.018}$
	VMSB (ours)	$\mathbf{8.774}_{\pm \mathbf{0.065}}$	$0.365_{\pm 0.002}$

Figure 9: Image-to-Image translation on a latent space. Left: Generation results for the FFHQ dataset (1024×1024) using our two SB variants. Right: Quantitative results using MMD metrics.

competitive FID and input/output MSD similarity scores for $K = 4096$. The consistent performance gains from the LightSB and LightSB-adv algorithms strongly supports our theoretical claims on online learning.

FFHQ. Following the latent SB setting of Korotin et al. (2024), we assessed our method by utilizing a pretrained ALAE model for generating 1024×1024 images of the FFHQ dataset (Karras et al., 2019). With the predefined 512-dimensional embedding space, we trained our SB models on the latent space to solve four distinct tasks: *Adult* $\rightarrow$ *Child*, *Child* $\rightarrow$ *Adult*, *Female* $\rightarrow$ *Male*, and *Male* $\rightarrow$ *Female*. Fig. 9 illustrates that our method delivered high-quality translation results. We also conducted a quantitative analysis using the ED on the ALAE embedding as a metric for evaluation. The result also verifies that our VMSB and VMSB-M algorithms consistently achieved lower ED scores than other baselines, demonstrating its applicability for the high dimensional embedding space. Consequently, the image-to-image transfer results showed that the generality of our online learning hypothesis and that the proposed algorithm is highly capable of interacting with neural networks of complex learning dynamics. Considering the significantly higher dimensionality of image domains relative to the batch sizes used in VOMD, the consistent and stable performance improvements demonstrated in our experiments strongly validate our theoretical claims regarding the robustness of our approach in online learning scenarios.

7 Conclusion

In this paper, we introduced VMSB, a practical simulation-free algorithm designed for effectively addressing SB problems encountered in real-world scenarios. We proposed a robust theoretical learning framework applicable to general SB solvers, leveraging a dual geometric interpretation of convex optimization to construct a robust OMD algorithm with rigorous guarantees on convergence and regret bounds. Furthermore, we proposed the computational algorithm for our OMD framework by employing the Wasserstein-Fisher-Rao geometry. Through extensive empirical evaluation, we validated the effectiveness of VMSB across diverse settings, including high-dimensional spaces, limited-sample regimes, and online learning environments. The

experimental results consistently demonstrated stable and superior benchmark performance, highlighting the enhanced robustness of our approach. Consequently, we argue that the proposed VMSB algorithm provides a promising and robust methodology for probabilistic generative modeling within learning-theoretic contexts.

Limitations. In this work, we significantly reduced the computational complexity inherent in the MD framework by adopting the Wasserstein-Fisher-Rao geometry. GMM-based models, due to the lack of deep structural processing, tend to focus on *instance-level* associations of images in EOT couplings rather than the *subinstance-* or *feature-level* associations that are intrinsic to deep generative models. As a result, while VMSB produces statistically valid representations of optimal transportation within the given architectural constraints, these outcomes may be perceived as somewhat synthetic compared to large neural networks. Nevertheless, GMM-based models still hold an irreplaceable role in numerous problems such as latent diffusion and variational methods, due to their simplicity and distinctive properties Korotin et al. (2024). As we successfully demonstrated in two distinct ways of interacting with neural networks for solving unpaired image transfer, we hope our theoretical and empirical findings help novel neural architecture studies.

Future research. One line of future studies in SB is a general understanding of learning in diffusion models with various regularizations. This includes diffusion models in various problem-specific constraints, and geometric constraints from manifolds. Another direction is the extension of the theoretical results into network architecture design. From § 4.2, a pair of Schrödinger potentials represent a dual representation of SB in a statistical manifold. In Gigli & Tamanini (2020), such potentials satisfy the Hamilton-Jacobi-Bellman (HJB) equations and, this can be trained with forward-backward SDE (SB-FBSDE) as presented by Liu et al. (2022). However, this requires many simulation samples from SDEs, and the requirements for applying VMSB contain a tractable way of estimating gradient flows, and a guarantee of measure concentration. Therefore, we expect there will be a new studies of energy-based neural architecture for efficiently representing SB, which will advance various subfields of machine learning. Lastly, a theoretical generalization of our work can be done by considering the Orlicz space for EOT studied by Lorenz & Mahler (2022). Since we essentially devised our theoretical framework to be compatible with arbitrary Bregman potentials, we believe controlling regularity of Young functionals can find more generalized learning algorithms for a wider range of OT problems.

References

Robert A Adams and John JF Fournier. *Sobolev spaces*, volume 140. Elsevier, 2003.

Charalambos D. Aliprantis and Kim C. Border. *Infinite Dimensional Analysis: a Hitchhiker's Guide*. Springer, 2006.

Jason Altschuler, Sinho Chewi, Patrik R Gerber, and Austin Stromme. Averaging on the Bures-Wasserstein manifold: dimension-free convergence of gradient descent. *Advances in Neural Information Processing Systems*, 34:22132–22145, 2021.

David Alvarez-Melis and Tommi S Jaakkola. Gromov–Wasserstein alignment of word embedding spaces. *arXiv preprint arXiv:1809.00013*, 2018.

Shun-ichi Amari. *Information geometry and its applications*, volume 194. Springer, 2016.

Shun-ichi Amari and Hiroshi Nagaoka. *Methods of information geometry*, volume 191. American Mathematical Soc., 2000.

Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2005a.

Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2005b.

Pierre-Cyril Aubin-Frankowski, Anna Korba, and Flavien Léger. Mirror descent with relative smoothness in measure spaces, with application to Sinkhorn and EM. *Advances in Neural Information Processing Systems*, 35:17263–17275, 2022.

Peter Auer, Nicolo Cesa-Bianchi, and Claudio Gentile. Adaptive and self-confident on-line learning algorithms. *Journal of Computer and System Sciences*, 64(1):48–75, 2002.

Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operations Research Letters*, 31(3):167–175, 2003.

Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numerische Mathematik*, 84(3):375–393, 2000.

Pierre Bernhard and Alain Rapaport. On a theorem of Danskin with an application to a theorem of von neumann-sion. *Nonlinear Analysis: Theory, Methods & Applications*, 24(8):1163–1181, 1995.

Rajendra Bhatia, Tanvi Jain, and Yongdo Lim. On the Bures–Wasserstein distance between positive definite matrices. *Expositiones Mathematicae*, 37(2):165–191, 2019.

Benjamin Birnbaum, Nikhil R Devanur, and Lin Xiao. Distributed algorithms via gradient descent for fisher markets. In *Proceedings of the 12th ACM conference on Electronic commerce*, pp. 127–136, 2011.

David M Blei, Alp Kucukelbir, and Jon D McAuliffe. Variational inference: A review for statisticians. *Journal of the American statistical Association*, 112(518):859–877, 2017.

Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

James Bradbury, Roy Frostig, Peter Hawkins, Matthew James Johnson, Chris Leary, Dougal Maclaurin, George Necula, Adam Paszke, Jake VanderPlas, Skye Wanderman-Milne, and Qiao Zhang. JAX: composable transformations of Python+NumPy programs. 2018. URL `http://github.com/google/jax`.

Lev M. Bregman. 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.

Donald Bures. An extension of Kakutani's theorem on infinite product measures to the tensor product of semifinite ω^*-algebras. *Transactions of the American Mathematical Society*, 135:199–212, 1969.

Dan Butnariu and Elena Resmerita. Bregman distances, totally convex functions, and a method for solving operator equations in Banach spaces. In *Abstract and Applied Analysis*, volume 2006, 2006.

Guillaume Carlier. On the linear convergence of the multimarginal sinkhorn algorithm. *SIAM Journal on Optimization*, 32(2):786–794, 2022.

Guillaume Carlier, Vincent Duval, Gabriel Peyré, and Bernhard Schmitzer. Convergence of entropic schemes for optimal transport and gradient flows. *SIAM Journal on Mathematical Analysis*, 49(2):1385–1418, 2017.

Mathilde Caron, Ishan Misra, Julien Mairal, Priya Goyal, Piotr Bojanowski, and Armand Joulin. Unsupervised learning of visual features by contrasting cluster assignments. *Advances in neural information processing systems*, 33:9912–9924, 2020.

José A Carrillo, Rishabh S Gvalani, and Jeremy S-H Wu. An invariance principle for gradient flows in the space of probability measures. *Journal of Differential Equations*, 345:233–284, 2023.

Gong Chen and Marc Teboulle. Convergence analysis of a proximal-like minimization algorithm using bregman functions. *SIAM Journal on Optimization*, 3(3):538–543, 1993.

Keyi Chen and Francesco Orabona. Generalized implicit follow-the-regularized-leader. In *International Conference on Machine Learning*, pp. 4826–4838. PMLR, 2023.

Tianrong Chen, Guan-Horng Liu, and Evangelos A. Theodorou. Likelihood training of Schrödinger bridge using forward-backward sdes theory. In *10th International Conference on Learning Representations*, 2022.

Yongxin Chen, Tryphon T Georgiou, and Allen Tannenbaum. Optimal transport for Gaussian mixture models. *IEEE Access*, 7:6269–6278, 2018.

Lenaic Chizat, Gabriel Peyré, Bernhard Schmitzer, and François-Xavier Vialard. An interpolating distance between optimal transport and Fisher–Rao metrics. *Foundations of Computational Mathematics*, 18:1–44, 2018.

Julien Claisse, Giovanni Conforti, Zhenjie Ren, and Songbo Wang. Mean field optimization problem regularized by Fisher information. *arXiv preprint arXiv:2302.05938*, 2023.

Giovanni Conforti. Weak semiconvexity estimates for schrödinger potentials and logarithmic sobolev inequality for schrödinger bridges. *Probability Theory and Related Fields*, pp. 1–27, 2024.

Giovanni Conforti, Alain Durmus, and Giacomo Greco. Quantitative contraction rates for sinkhorn algorithm: beyond bounded costs and compact marginals. *arXiv preprint arXiv:2304.04451*, 2023.

Isaac P Cornfeld, Sergej V Fomin, and Yakov Grigorevich Sinai. *Ergodic theory*, volume 245. Springer Science & Business Media, 2012.

Marco Cuturi. Sinkhorn distances: lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems*, volume 26, 2013.

John M Danskin. *The theory of max-min and its application to weapons allocation problems*, volume 5. Springer, 1967.

Kamélia Daudel et al. Mixture weights optimisation for alpha-divergence variational inference. *Advances in Neural Information Processing Systems*, 34:4397–4408, 2021.

Valentin De Bortoli, James Thornton, Jeremy Heng, and Arnaud Doucet. Diffusion Schrödinger bridge with applications to score-based generative modeling. *Advances in Neural Information Processing Systems*, 34:17695–17709, 2021.

Nabarun Deb, Young-Heon Kim, Soumik Pal, and Geoffrey Schiebinger. Wasserstein mirror gradient flow as the limit of the Sinkhorn algorithm. *arXiv preprint arXiv:2307.16421*, 2023.

Michael Ziyang Diao, Krishna Balasubramanian, Sinho Chewi, and Adil Salim. Forward-backward Gaussian variational inference via JKO in the Bures-Wasserstein space. In *International Conference on Machine Learning*, pp. 7960–7991. PMLR, 2023.

John C Duchi and Hongseok Namkoong. Learning models with uniform performance via distributionally robust optimization. *The Annals of Statistics*, 49(3):1378–1406, 2021.

Huang Fang, Nicholas JA Harvey, Victor S Portella, and Michael P Friedlander. Online mirror descent and dual averaging: keeping pace in the dynamic case. *Journal of Machine Learning Research*, 23(121):1–38, 2022.

Amos Fiat and Gerhard J Woeginger. *Online algorithms: The state of the art*, volume 1442. Springer, 1998.

Nicola Gigli and Luca Tamanini. Benamou-Brenier and duality formulas for the entropic cost on $RCD^*(K, N)$ spaces. *Probability Theory and Related Fields*, 176(1-2):1–34, 2020.

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in Neural Information Processing Systems*, pp. 2672–2680, 2014.

Robert M Gray and John C Kieffer. Asymptotically mean stationary measures. *The Annals of Probability*, pp. 962–973, 1980.

Leonard Gross. Logarithmic Sobolev inequalities. *American Journal of Mathematics*, 97(4):1061–1083, 1975.

Suriya Gunasekar, Blake E. Woodworth, and Nathan Srebro. Mirrorless mirror descent: A more natural discretization of riemannian gradient flow. *CoRR*, abs/2004.01025, 2020.

Nikita Gushchin, Sergei Kholkin, Evgeny Burnaev, and Alexander Korotin. Light and optimal schrödinger bridge matching. In *International Conference on Machine Learnin*. PMLR, 2024a.

Nikita Gushchin, Alexander Kolesov, Petr Mokrov, Polina Karpikova, Andrei Spiridonov, Evgeny Burnaev, and Alexander Korotin. Building the bridge of Schrödinger: A continuous entropic optimal transport benchmark. *Advances in Neural Information Processing Systems*, 36, 2024b.

Dong-Sig Han, Hyunseo Kim, Hyundo Lee, JeHwan Ryu, and Byoung-Tak Zhang. Robust imitation via mirror descent inverse reinforcement learning. *Advances in Neural Information Processing Systems*, 35:30031–30043, 2022.

Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Fundamentals of convex analysis*. Springer Science & Business Media, 2004.

Ya-Ping Hsieh, Ali Kavis, Paul Rolland, and Volkan Cevher. Mirrored Langevin dynamics. *Advances in Neural Information Processing Systems*, 31, 2018.

Eric Jang, Shixiang Gu, and Ben Poole. Categorical reparameterization with gumbel-softmax. In *5th International Conference on Learning Representations*, 2016.

Sadeep Jayasumana, Srikumar Ramalingam, Andreas Veit, Daniel Glasner, Ayan Chakrabarti, and Sanjiv Kumar. Rethinking fid: Towards a better evaluation metric for image generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9307–9315, 2024.

Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker–Planck equation. *SIAM journal on mathematical analysis*, 29(1):1–17, 1998.

Mohammad Reza Karimi, Ya-Ping Hsieh, and Andreas Krause. Sinkhorn flow as mirror flow: a continuous-time framework for generalizing the sinkhorn algorithm. In *International Conference on Artificial Intelligence and Statistics*, pp. 4186–4194. PMLR, 2024.

Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 4401–4410, 2019.

Alexander Korotin, Nikita Gushchin, and Evgeny Burnaev. Light Schrödinger bridge. In *13th International Conference on Learning Representations*, 2024.

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

Marc Lambert, Sinho Chewi, Francis Bach, Silv ère Bonnabel, and Philippe Rigollet. Variational inference via Wasserstein gradient flows. *Advances in Neural Information Processing Systems*, 35:14434–14447, 2022.

Flavien Léger. A gradient descent perspective on Sinkhorn. *Applied Mathematics & Optimization*, 84(2):1843–1855, 2021.

Yunwen Lei and Ding-Xuan Zhou. Convergence of online mirror descent. *Applied and Computational Harmonic Analysis*, 48(1):343–373, 2020.

Christian Léonard. From the Schrödinger problem to the Monge–Kantorovich problem. *Journal of Functional Analysis*, 262(4):1879–1920, 2012.

Christian Léonard. Some properties of path measures. *Séminaire de Probabilités XLVI*, pp. 207–230, 2014.

Matthias Liero, Alexander Mielke, and Giuseppe Savaré. Optimal transport in competition with reaction: The Hellinger–Kantorovich distance and geodesic curves. *SIAM Journal on Mathematical Analysis*, 48(4):2869–2911, 2016.

Matthias Liero, Alexander Mielke, and Giuseppe Savaré. Optimal entropy-transport problems and a new Hellinger–Kantorovich distance between positive measures. *Inventiones mathematicae*, 211(3):969–1117, 2018.

Guan-Horng Liu, Tianrong Chen, Oswin So, and Evangelos Theodorou. Deep generalized Schrödinger bridge. In *Advances in Neural Information Processing Systems*, pp. 9374–9388, 2022.

Guan-Horng Liu, Arash Vahdat, De-An Huang, Evangelos A Theodorou, Weili Nie, and Anima Anandkumar. I^2sb: Image-to-image Schrödinger bridge. In *International Conference on Machine Learning*, pp. 22042–22062. PMLR, 2023.

Dirk Lorenz and Hinrich Mahler. Orlicz space regularization of continuous optimal transport problems. *Applied Mathematics & Optimization*, 85(2):14, 2022.

John Lott. Some geometric calculations on Wasserstein space. *Communications in Mathematical Physics*, 277(2): 423–437, 2008.

Haihao Lu, Robert M Freund, and Yurii Nesterov. Relatively smooth convex optimization by first-order methods, and applications. *SIAM Journal on Optimization*, 28(1):333–354, 2018.

Yulong Lu, Jianfeng Lu, and James Nolen. Accelerating langevin sampling with birth-death. *arXiv preprint arXiv:1905.09863*, 2019.

Aleksander Madry, Aleksandar Makelov, Ludwig Schmidt, Dimitris Tsipras, and Adrian Vladu. Towards deep learning models resistant to adversarial attacks. *arXiv preprint arXiv:1706.06083*, 2017.

Florient Malrieu. Logarithmic sobolev inequalities for some nonlinear pde's. *Stochastic processes and their applications*, 95(1):109–132, 2001.

Brendan McMahan. Follow-the-regularized-leader and mirror descent: Equivalence theorems and l1 regularization. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, pp. 525–533. JMLR Workshop and Conference Proceedings, 2011.

John Milnor. Microbundles: Part i. *Topology*, 3:53–80, 1964.

Daniel Morales-Brotons, Thijs Vogels, and Hadrien Hendrikx. Exponential moving average of weights in deep learning: Dynamics and benefits. *Transactions on Machine Learning Research*, 2024.

Arkadiĭ Semenovich Nemirovsky and David Borisovich Yudin. *Problem complexity and method efficiency in optimization*. A Wiley-Interscience publication. Wiley, 1983.

Yurii Nesterov. Primal-dual subgradient methods for convex problems. *Mathematical programming*, 120(1):221–259, 2009.

Marcel Nutz. Introduction to entropic optimal transport. *Lecture notes, Columbia University*, 2021.

Marcel Nutz and Johannes Wiesel. Stability of Schrödinger potentials and convergence of sinkhorn's algorithm. *The Annals of Probability*, 51(2):699–722, 2023.

Francesco Orabona and Dávid Pál. Scale-free online learning. *Theoretical Computer Science*, 716:50–69, 2018.

Felix Otto. The geometry of dissipative evolution equations: the porous medium equation. *Communications in Partial Differential Equations*, 26(1-2):101–174, 2001.

Felix Otto and Cédric Villani. Generalization of an inequality by talagrand and links with the logarithmic sobolev inequality. *Journal of Functional Analysis*, 173(2):361–400, 2000.

John Paisley, David Blei, and Michael Jordan. Variational Bayesian inference with stochastic search. *arXiv preprint arXiv:1206.6430*, 2012.

Michele Pavon and Anton Wakolbinger. On free energy, stochastic control, and Schrödinger processes. In *Modeling, Estimation and Control of Systems with Uncertainty: Proceedings of a Conference held in Sopron, Hungary, September 1990*, pp. 334–348. Springer, 1991.

Gabriel Peyré, Marco Cuturi, et al. Computational optimal transport: With applications to data science. *Foundations and Trends® in Machine Learning*, 11(5-6):355–607, 2019.

Stanislav Pidhorskyi, Donald A Adjeroh, and Gianfranco Doretto. Adversarial latent autoencoders. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 14104–14113, 2020.

Alec Radford. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*, 2015.

Cale Rankin and Ting-Kam Leonard Wong. Bregman-Wasserstein divergence: geometry and applications. *arXiv preprint arXiv:2302.05833*, 2023.

G. Raskutti and S. Mukherjee. The information geometry of mirror descent. *IEEE Transactions on Information Theory*, 61(3):1451–1457, 2015.

Maria L Rizzo and Gábor J Székely. Energy distance. *wiley interdisciplinary reviews: Computational statistics*, 8(1): 27–38, 2016.

Filippo Santambrogio. Optimal transport for applied mathematicians. *Birkäuser, NY*, 55(58-63):94, 2015.

Filippo Santambrogio. {Euclidean, metric, and Wasserstein} gradient flows: an overview. *Bulletin of Mathematical Sciences*, 7:87–154, 2017.

Erwin Schrödinger. Sur la théorie relativiste de l'électron et l'interpré tation de la mécanique quantique. In *Annales de l'institut Henri Poincaré*, volume 2, pp. 269–310, 1932.

Yuyang Shi, Valentin De Bortoli, Andrew Campbell, and Arnaud Doucet. Diffusion Schrödinger bridge matching. In *Advances in Neural Information Processing Systems*, volume 36, pp. 62183–62223, 2023.

Hirohiko Shima and Katsumi Yagi. Geometry of Hessian manifolds. *Differential Geometry and its Applications*, 7 (3):277–290, 1997. ISSN 0926-2245.

Nati Srebro, Karthik Sridharan, and Ambuj Tewari. On the universality of online mirror descent. In *Advances in Neural Information Processing Systems*, pp. 2645–2653, 2011.

Alexander Tong, Nikolay Malkin, Kilian Fatras, Lazar Atanackovic, Yanlei Zhang, Guillaume Huguet, Guy Wolf, and Yoshua Bengio. Simulation-free Schrödinger bridges via score and flow matching. *arXiv preprint arXiv:2307.03672*, 2023.

Alexander Tong, Kilian FATRAS, Nikolay Malkin, Guillaume Huguet, Yanlei Zhang, Jarrid Rector-Brooks, Guy Wolf, and Yoshua Bengio. Improving and generalizing flow-based generative models with minibatch optimal transport. *Transactions on Machine Learning Research*, 2024a. ISSN 2835-8856. Expert Certification.

Alexander Y. Tong, Nikolay Malkin, Kilian Fatras, Lazar Atanackovic, Yanlei Zhang, Guillaume Huguet, Guy Wolf, and Yoshua Bengio. Simulation-free Schrödinger bridges via score and flow matching. In *Proceedings of The 27th International Conference on Artificial Intelligence and Statistics*, volume 238 of *Proceedings of Machine Learning Research*, pp. 1279–1287. PMLR, 2024b.

Francisco Vargas, Pierre Thodoroff, Austen Lamacraft, and Neil Lawrence. Solving Schrödinger bridges via maximum likelihood. *Entropy*, 23(9):1134, 2021.

Cédric Villani. *Optimal transport: old and new*, volume 338. Springer, 2009.

Cédric Villani. *Topics in optimal transportation*, volume 58. American Mathematical Soc., 2021.

Huan Xu, Constantine Caramanis, and Shie Mannor. Robust regression and lasso. *Advances in neural information processing systems*, 21, 2008.

Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th international conference on machine learning (icml-03)*, pp. 928–936, 2003.

Appendices for

Variational Online Mirror Descent for Robust Learning in Schrödinger Bridge

Abbreviation and Notation

Abbreviation	Expansion
SB	Schrödinger Bridge
SBP	Schrödinger Bridge Problem
EOT	Entropy-regularized Optimal Transport
MD	Mirror Descent
OMD	Online Mirror Descent
KL	Kullback-Leibler
IPF	Iterative Proportional Fitting
BW	Bures–Wasserstein
WFR	Wasserstein-Fisher-Rao
SDE	Stochastic Differential Equation
PDE	Partial Differential Equation
FP	Fokker–Planck
GMM	Gaussian mixture model

Notation	Usage
μ, ν	marginal distributions
ε	volatility of reference measure
c_ε	cost $c_\varepsilon(x,y) \coloneqq \frac{1}{2\varepsilon}\|x-y\|^2$
π	a coupling of μ and ν
$\vec{\pi}, \overleftarrow{\pi}$	conditional distributions
γ_n	n-th marginal
φ, ψ	log-Schrödinger potential
Ω, D_Ω	Bregman potential/divergence
d^+	directional derivative
$\delta_{\mathcal{C}}, \delta_{\mathcal{D}}$	First variations
$\nabla_{\mathbb{W}}$	Wasserstein-2 gradient operator
$\mathcal{T}$	dynamic stochastic process in SB
g	drift function
$i_{\mathcal{C}}$	indicator function

A Theoretical Details and Proofs

Background on first variation operators. In this paper, we utilize the notation of first variation operators $\delta_{\mathcal{C}}$ and $\delta_{\mathcal{D}}$ to identify the generalized primal and dual spaces in Schrödinger bridge. Since the problems are classified as an infinite-dimensional optimization (Aliprantis & Border, 2006), we introduce the essential background supporting the necessity of these operators. We introduce Gâteaux and Fréchet differentiablility (Aubin-Frankowski et al., 2022; Karimi et al., 2024).

Definition 4 (Gâteaux & Fréchet differentiablility)**.** Let $\mathcal{M}$ be a topological vector space of measures on a space $\mathcal{X}$. Define the Gâteaux differentiablity of a functional $F : \mathcal{M} \to \mathbb{R}$, if there exists a gradient operator $\nabla_{\text{Gât}}$ such that for an arbitrary direction $v \in \mathcal{M}$, defined as the limit

$$\nabla_{\text{Gât}} F(x)[v] = \lim_{h \to 0} \frac{F(x+hv) - F(x)}{h}, \quad x \in \mathcal{M}$$

If the limit exists in the unit ball in $\mathcal{M}$, the function F is called Fréchet differentiable with $\nabla_{\text{Fré}} F(x)$.

The problem of the Gâteaux and Fréchet differentiability in the context of SB is that the limit must be given in *all* directions, implying that every neighboring point must be within the domain of the topological space $\mathcal{M}$. For the case of functionals such as the KL divergence functional $F(\cdot) = \text{KL}(\cdot|\pi^*)$, the domain of F and has an empty interior (Aubin-Frankowski et al., 2022). To resolve this issue, we use *directional derivatives* and *first variations*, defined in Definitions 1 and 2.

First variations of KL. Suppose that we have two distributions $\rho, \rho' \in \mathcal{P}_2(\mathcal{X}), \mathcal{X} \subseteq \mathbb{R}^d$. Let us consider the log likelihood of ρ': $\ell'(x) \coloneqq \log \rho'(x)$, and an element of a (topological) tangent space $v \in T_\rho \mathcal{P}_2(\mathcal{X})$ (Milnor, 1964). Then, we can achieve the followings:

$$\text{KL}(\rho \| \rho') = \int_{\mathcal{X}} \log \rho(x) \, \mathrm{d}\rho(x) - \int_{\mathcal{X}} \ell'(x) \, \mathrm{d}\rho(x) \tag{18}$$

$$\int \ell'(x) \big[\rho(x) + hv(x)\big] \, \mathrm{d}x = \int \ell'(x)\rho(x) \, \mathrm{d}x + h \int \ell'(x) v(x) \, \mathrm{d}x \tag{19}$$

Given that $\log(z+\varepsilon)(z+\varepsilon) = \log(z)z + [\log(z)+1]\varepsilon + o(\varepsilon)$, and $\int_{\mathcal{X}} v(x)\,\mathrm{d}x = 0$, we achieve

$$\begin{aligned}\int_{\mathcal{X}} \log\big(\rho(x)+hv(x)\big)\big(\rho(x)+hv(x)\big)\,\mathrm{d}x &= \int_{\mathcal{X}} \log\rho(x)\rho(x) + [\log\rho(x)+1]hv(x) + o(h)\ \mathrm{d}x \\ &= \int_{\mathcal{X}} \log\rho(x)\rho(x)\,\mathrm{d}x + h\int_{\mathcal{S}} \log\rho(x)v(x)\mathrm{d}x + h\int v(x)\mathrm{d}x + o(h)\end{aligned} \tag{20}$$

Combining Eqs. (18-20), we achieve

$$F(\rho+hv) = F(\rho) + h\left\langle \log\left(\frac{\rho}{\rho'}\right),\, v\right\rangle + o(h). \tag{21}$$

By Eq. (21) and Definition 2, the first variation $\delta F_2 \in T^*\mathcal{P}(\mathcal{X})$ exists for infinitesimal $h > 0$. Therefore, the first variation of KL is derived as $\delta\mathrm{KL}(\rho\|\rho') = \log\frac{\rho}{\rho'}$. In machine learning, log likelihoods of probabilistic models are often given in a closed-form expression, incentivizing development of computational continuous EOT/SB methods. Generally, identical arguments generally apply to all KL functionals with respect to distributions (π, $\vec{\pi}$, and marginals) in our setup.

Asymptotically log-concave distributions. For convergence analysis, we assume each marginal distribution is in log-concave distribution, particularly satisfying the log Sobolev inequality of measures, motivated by relevant literature (Otto & Villani, 2000; Conforti, 2024). This assumption works a wider range of costs and marginals beyond popular choices with boundedness and compactness (Nutz & Wiesel, 2023; Conforti et al., 2023). Suppose that marginals admit densities of the form

$$\mu(\mathrm{d}x) = \exp\big(-U_\mu(x)\big)\mathrm{d}x \qquad \text{and} \qquad \nu(\mathrm{d}y) = \exp\big(-U_\nu(y)\big)\mathrm{d}y. \tag{22}$$

We exploit the following definition from (Conforti et al., 2023) in order to describe asymptotically log-concaveness.

Definition 5 (Asymptotically strongly log-concavity; Conforti et al., 2023)**.** Suppose that marginals μ and ν admit a positive density against the Lebesgue measure, which can be written in the form (22). In particular, consider a collection of functions $\mathcal{G} := \{g \in C^2((0,+\infty),\mathbb{R}_+) | r \mapsto r^{1/2}g(r^{1/2})$is non-increasing and concave, $\lim_{r\to 0} rg(r) = 0\}$. Accordingly, define a set

$$\tilde{\mathcal{G}} := \{g \in \mathcal{G} \text{ bounded and s.t. } \lim_{r\to 0^+} g(r) = 0,\ \ g' \geq 0 \quad \text{and} \quad 2g'' + gg' \leq 0\} \subset \mathcal{G}.$$

and *convexity profile* $\kappa_U : \mathbb{R}_+ \to \mathbb{R}$ of a differentiable function U as the following

$$\kappa_U(r) := \inf\left\{\frac{\langle \nabla U(x) - \nabla U(y), x-y\rangle}{|x-y|^2}\ :\ |x-y| = r\right\}.$$

We say a potential is asymptotically strongly convex if there exists $\alpha_U \in \mathbb{R}_+$ and $\tilde{g}_U \in \tilde{\mathcal{G}}$ such that

$$\kappa_U(r) \geq \alpha_U - r^{-1}\tilde{g}_U(r) \tag{23}$$

holds for all $r > 0$. We consider the set of asymptotically strongly log-concave probability measures

$$\mathcal{P}_{\mathrm{alc}}(\mathbb{R}^d) := \{\zeta(\mathrm{d}x) = \exp(-U(x))\mathrm{d}x : U \in C_2(\mathbb{R}^d),\ \ U \text{ is asymptotically strongly convex}\}.$$

It is essential to note that a mixture of asymptotically log concave is also asymptotically log concave.

Lemma 3. *For positive weights $\boldsymbol{\beta} = \{\beta_k\}_{k=1}^K$ with $\sum_{k=1}^K \beta_k = 1$ and asymptotically log concave distributions $\{\rho_k\}_{k=1}^K$, $\pi = \sum_{k=1}^K \beta_k\rho_k$.*

Proof. Let us reformulate the mixture as $\log\pi(x) = \log\sum_k \beta_k \exp(-U_k(x))$ for asymptotically strongly convex functions $\mathbf{U} = \{U_k\}_{k=1}^K$. The gradient is

$$\nabla\log\pi = \mathfrak{J}^T\mathbf{p}, \quad \mathfrak{J} = -\begin{bmatrix}\nabla U_1 \\ \vdots \\ \nabla U_K\end{bmatrix}, \quad \mathbf{p} = \mathrm{softmax}(\log\boldsymbol{\beta} - \mathbf{U})$$

If each U_k of mixture satisfy Eq. (23) with α_{U_k} and $\tilde{g}_{U_k}$, there exist $\alpha_U = \min_{1\le k\le K} \alpha_{U_k}$ and $\tilde{g}_U(r) = -r\log\sum_{k=1}^K \exp(-r^{-1}\tilde{g}_{U_k})$ that satisfies the condition (23) for $U = -\log\pi$. By direct calculation, one can easily see that soft min-like property of $r^{-1}\tilde{g}_U$ from $\{r^{-1}\tilde{g}_{U_k}\}_{k=1}^K$ does not change the conditions of $\tilde{\mathcal{G}}$. □

General assumptions and justifications. We additionally need the following general assumptions for our OMD framework. ① (Existence) The sequence of MD from Eq. (9) exists $\{\pi_t\}_{t\in\mathbb{N}} \subset \mathcal{C}$, and are unique, ② (Relative smoothness/convexity) For some $l, L \ge 0$, the functional F_t is L-smooth and l-strongly-convex relative to Ω. ③ (Existence of first variations) For each $t \ge 0$, the first variation $\delta_c\Omega(\pi_t)$ exists. ④ (Boundedness of estimations) The asymptotic dual mean $\pi_{\mathcal{D}}^\circ$ is almost surely bounded $\Pr(D_\Omega(\pi_t\|\pi_{\mathcal{D}}^\circ) \le R) = 1$ for some $R > 0$. ⑤ (Ergodicity) The estimation process of $\{\pi_t^\circ\}_{t=1}^\infty$ is governed by a measure-preserving transformation on a measure space $(\mathcal{Y}, \Sigma, \varsigma)$ with $\varsigma(\mathcal{Y}) = 1$; for every event $E \in \Sigma$, $\varsigma(T^{-1}(E)\triangle E) = 0$ (that is, E is invariant), either $\varsigma(E) = 0$ or $\varsigma(E) = 1$ (Cornfeld et al., 2012).[2] For ①, the temporal cost $F_t(\cdot) = \mathrm{KL}(\cdot|\pi_t^\circ)$ is well defined since KL is a strong Bregman divergence with lower semicontinuity, where the existence of a primal solution in guaranteed as discussed in Aubin-Frankowski et al. (2022). For ②-③, we can identify $l = L = 1$ and close-form expression of the first variation that is shown in Definition 6 and Proposition 3. For the assumptions ④-⑤, we postulate the existence of estimates produced from a Monte-Carlo method, using a fixed amount of updates on topological vector space. Hence, it is natural to consider that these estimates will be bounded in a probabilistic sense and yield Markovian transitions, which are aperiodic and irreducible.

A.1 Proofs of Lemmas 1 and 2

The EOT in Eq. (2) can be reformulated as a divergence minimization problem with respective to a reference measure. If a Gibbs parameterization is enforced with the quadratic cost functional $c_\varepsilon(x, y) = \frac{1}{2\varepsilon}\|x-y\|^2$ for $\varepsilon > 0$, it is well known that the problem has the equivalence with the entropy regularized optimal transport problem (Nutz, 2021)

$$\mathrm{OT}_\varepsilon(\mu,\nu) = \inf_{\pi\in\Pi(\mu,\nu)} \mathrm{KL}\big(\pi\|e^{-c_\varepsilon}\mu\otimes\nu\big). \tag{24}$$

Note that the above equation corresponds to the constrained minimization of $\mathrm{KL}(\mathcal{T}\|W^\varepsilon)$ in Eq. (4) by the disintegration theorem of Schrödinger bridge (Appendix A of Vargas et al., 2021). While the Bregman projection formulation of Sinkhorn Eq. (10) are described by the spaces $(\Pi_\mu^\perp, \Pi_\nu^\perp)$, it is (equally) natural to think that considering the problem as convex problem with the distributional constraint $\mathcal{C}$ (see the primal space in illustrated in Fig. 1). As a problem in the constraint $\mathcal{C}$, one can consider a temporal cost functional $\widetilde{F}_t(\pi) \coloneqq a_t\mathrm{KL}(\gamma_1\pi\|\mu) + (1-a_t)\mathrm{KL}(\gamma_2\pi\|\nu)$ with sequences $\{a_t\}_{t=1}^\infty = \{0,1,0,1,\dots\}$ for $\gamma_1\pi(x) \coloneqq \int \pi(x,y)\mathrm{d}y$ and $\gamma_2\pi(y) \coloneqq \int \pi(x,y)\mathrm{d}x$. By construction, we have the following MD update:

$$\underset{\pi\in\mathcal{C}}{\mathrm{minimize}}\big\langle\delta_c\widetilde{F}_t(\pi_t), \pi-\pi_t\big\rangle + D_\Omega(\pi\|\pi_t). \tag{25}$$

The optimization problem (25) is equivalent to having the property for subsequent π_{t+1}:

$$\begin{aligned} &d^+\widetilde{F}_t(\pi_t;\pi-\pi_t) + D_\Omega(\pi\|\pi_t) \ge d^+\widetilde{F}_t(\pi_t;\pi_{t+1}-\pi_t) + D_\Omega(\pi_{t+1}|\pi_t) \\ &\qquad\iff \big\langle\delta_c\widetilde{F}_t(\pi_t) - \delta_c\Omega(\pi_t),\, \pi-\pi_{t+1}\big\rangle + \big(\Omega(\pi)-\Omega(\pi_{t+1})\big) \ge 0, \quad \forall\pi\in\mathcal{C}. \end{aligned} \tag{26}$$

Setting the free parameter $\pi = \pi_{t+1} + h(\pi-\pi_{t+1})$ and taking the limit $h \to 0^+$ yields described the time evolution of the log-Schrödinger potentials for $\pi_t = e^{\varphi_t\oplus\psi_t - c_\varepsilon}\mathrm{d}(\mu\otimes\nu)$:

$$\dot{\varphi}_t = -\log\frac{\mathrm{d}(\gamma_1\pi_t)}{\mathrm{d}\nu_*} = -\alpha\bigg(\varphi_t - \varphi^* + \log\int_{\mathbb{R}^d} e^{\psi_t-\psi^*}\nu(\mathrm{d}y)\bigg), \tag{27a}$$

$$\dot{\psi}_t = -\log\frac{\mathrm{d}(\gamma_2\pi_t)}{\mathrm{d}\mu_*} = -\beta\bigg(\psi_t - \psi^* + \log\int_{\mathbb{R}^d} e^{\varphi_t-\varphi^*}\mu(\mathrm{d}x)\bigg), \tag{27b}$$

[2]Here, $\triangle$ denotes the symmetric difference, equivalent to the exclusive-or with respect to set membership.

for $\alpha = a_t$ and $\beta = 1 - a_t$.[3] Setting a discrete approximation of dynamics Eq. (27): $\varphi_{t+1} = \varphi_t + \dot{\varphi}_t$ and $\psi_{t+1} = \psi_t + \dot{\psi}_t$ yields the following alternating updates:

$$\psi_{2t+1}(y) = -\log \int_{\mathbb{R}^d} e^{\varphi_{2t}(x) - c_\varepsilon(x,y)} \mu(\mathrm{d}x), \quad \varphi_{2t+2}(x) = -\log \int_{\mathbb{R}^d} e^{\psi_{2t+1}(x) - c_\varepsilon(x,y)} \nu(\mathrm{d}y).$$

Therefore, the proof of Lemma 1 is complete.

From the dual iteration of KL stated in Eq. (40) (Aubin-Frankowski et al., 2022), the static, idealized MD cost $F(\cdot) = \mathrm{KL}(\cdot \| \pi^*)$ yield the following closed-form expression for the first variation:

$$\delta_c \Omega(\pi_t) - \delta_c \Omega(\pi_{t+1}) = \eta_t \big(\delta_c \Omega(\pi_t) - \delta_c \Omega(\pi^*) \big),$$

where the equation implies that setting $\eta_t \equiv 1$ for MD yields one-step optimality π^* in this idealized condition. Utilizing the equivalence of first variation stated in Lemma 5 and the disintegration theorem for the Radon-Nikodym derivatives, we get the first variation of F with respect to π for all x as

$$\delta F(\pi_t) = \log \frac{d\pi^*}{d\pi}. \tag{28}$$

And by the disintegration theorem (Léonard, 2014), we also achieve the first variation of f with respect to $\vec{\pi}$ for all x as

$$\delta f(\vec{\pi}_t^x) = \log \frac{d(\vec{\pi}_t^*)^x}{d\vec{\pi}^x}, \tag{29}$$

where $f(\vec{\pi}^x) = \mathrm{KL}(\vec{\pi}^x \| (\vec{\pi}^*)^x)$. Since this disintegration theorem always hold for every directional derivative, we can use expression for $\vec{\pi}^x$ and π interchangeably. It is well-known that MD is a discretization of natural gradient descent (Gunasekar et al., 2020), and our setting for Ω generates the geometry governed by the (generalized) Fisher information. In this particular case, one can use Otto's formalization of Riemannian calculus (Otto, 2001; § 3.2), and the probability space equipped with the Wasserstein-2 metric $(\mathcal{P}_2(\mathbb{R}^d), W_2)$, is generally represented as a Wasserstein gradient flow

$$\partial_t \pi_t = -\nabla_{\mathbb{W}} F(\pi_t), \qquad \forall \pi_t \in \mathcal{C}, \tag{30}$$

where $\nabla_{\mathbb{W}}$ denotes the Wasserstein-2 gradient operator $\nabla_{\mathbb{W}} \coloneqq \nabla \cdot \big(\rho \, \nabla \frac{\delta}{\delta \rho} \big)$. In particular, plugging Eq. (28) yields

$$\partial_t \pi_t = -\nabla \cdot (\pi \nabla \log \pi^*) + \Delta \pi, \tag{31}$$

where Δ denotes the Laplace operator. The foundational results concerning Wasserstein gradients were initially established by JKO (Jordan et al., 1998), who demonstrated that the formulation in Eq. (30) corresponds precisely to the Fokker–Planck equation (31). Consequently, it follows that Wasserstein gradients characterize the tangential direction of flows on a manifold constrained by distributional properties and endowed with the W_2 metric. □

A.2 Proof of Theorem 1

We start with introducing basic properties of the Bregman divergence in Definition 3. First, the *idempotence* property states that a Bregman divergence associated with another Bregman divergence $D_\Omega(\cdot|y)$ remains as the identical divergence with the original. Note that the (global or universal) idempotence initially stated by Aubin-Frankowski et al. (2022), but we apply some changes to the statement and only work with localized version of idempotence for the purpose of this paper.

Lemma 4 (Idempotence)**.** *Suppose a convex potential $\Omega : \mathcal{M}(\mathcal{X}) \to \mathbb{R} \cup \{+\infty\}$, where $\mathcal{M}(\mathcal{X})$ denotes a topological vector space for $\mathcal{X}$. Assume that for all $z \in \mathrm{dom}(\Omega)$, $\delta_c \Omega(z)$ exists. Then, $\forall x, y \in \mathcal{C} \cap \mathrm{dom}(\Omega)$: $D_{D_\Omega(\cdot|y)}(x|y) = D_\Omega(x|y)$.*

[3] More precisely, one needs to apply Lemma 5 for KL, and the disintegration theorem to get Eq. (27).

Proof of Lemma 4. Both Bregman divergences and Bregman potentials are convex functionals. By definition, we have $D_{D_\Omega(\cdot|z)}(x|y) = D_\Omega(x\|z) - D_\Omega(y\|z) - \langle \delta_c\Omega(y) - \delta_c\Omega(z),\, x - y\rangle$ for arbitrary z, and setting $z = y$ completes the proof. Another (informal) point of view is considering the Bregman divergence as a first-order approximation of a Hessian structure, and $D_{D_\Omega(\cdot|z)}$ converges to $D_\Omega(\cdot|z)$ by taking a limit, knowing that $D_\Omega(y|y) = 0$. □

We then proceed to the equivalence property of the family of recursive Bregman divergences. The property is important for proving the theorem and representing the dual representation of MD. Moreover, it is also used in Theorem 2 as a key ingredient which constructs our VOMD framework.

Lemma 5 (Equivalence of first variations)**.** *Suppose $\Omega : \mathcal{M}(\mathcal{X}) \to \mathbb{R} \cup \{+\infty\}$ Assume that for all $z \in \mathrm{dom}(\Omega)$, the first variation $\delta_c\Omega(z)$ exists, then, for all $x, y, y_1, y_2 \in \mathrm{dom}(\Omega)$, the first variation taken for the first argument x of the following Bregman divergences are equivalent: $\delta_c D_\Omega(x|y) = \delta_c D_{D_\Omega(\cdot|y_1)}(x|y) = \delta_c D_{D_\Omega(\cdot|y_2)}(x|y)$.*

Proof of Lemma 5. First, it can be analytically driven $\delta_c D_\Omega(x|y) = \delta_c\Omega(x) - \delta_c\Omega(y)$. Next, by definition, taking the first variation of $D_{D_\Omega(\cdot|z)}(x|y)$ with respect to x for arbitrary $z \in \mathrm{dom}(\Omega)$ yields $\delta_c D_\Omega(x\|z) - \delta_c\langle \Omega(y) - \Omega(z), x - y\rangle$. Knowing that the second term $\delta_c\langle \Omega(y) - \Omega(z), x - y\rangle$ is linear, we achieve $\delta D_{D_\Omega(\cdot|z)}(x|y) = \delta_c\Omega(x) - \delta_c\Omega(z) - (\delta_c\Omega(y) - \delta_c\Omega(z)) = \delta_c\Omega(x) - \delta_c\Omega(y)$, which completes the proof. □

By an inductive reasoning, we arrive at the basic property of family of Bregman divergences, that all divergence recursively defined by the Bregman potential Ω, has the (local) idempotence and the (global) equivalence of first variation. To address characteristics for particular Bregman potential Ω, we apply the notions of relative smoothness and convexity with respect to Ω, which was first introduced by Birnbaum et al. (2011).

Definition 6 (Relative smoothness and convexity)**.** Let $G : \mathcal{M}(\mathcal{X}) \to \mathbb{R} \cup \{+\infty\}$ be a proper convex functional. Given scalar $l, L \geq 0$, we define that G is L-smooth and l-strongly-convex relative to Ω over $\mathcal{C}$ if for every $x, y \in \mathrm{dom}(G) \cap \mathrm{dom}(\Omega) \cap \mathcal{C}$, we have

$$D_G(x\|y) \leq L D_\Omega(x\|y), \qquad D_G(x\|y) \geq l D_\Omega(x\|y),$$

respectively, where D_G and D_G are Bregman divergences associated with G defined in Definition 3.

Applying the idempotence lemma Lemma 4, we immediately recognize that the Bregman divergence D_Ω is relatively 1-smooth and 1-strongly-convex for Ω. To start our analysis, we reintroduce the well-known three-point identity for a Bregman divergence.

Lemma 6 (Three-point identity)**.** *For all $\pi_a, \pi_b, \pi_c \in \mathcal{C} \cap \mathrm{dom}(\Omega)$, we have the following identity*

$$\big\langle \delta_c\Omega(\pi_a) - \delta_c\Omega(\pi_b), \pi_c - \pi_b \big\rangle = D_\Omega(\pi_c\|\pi_b) - D_\Omega(\pi_c\|\pi_a) + D_\Omega(\pi_b\|\pi_a)$$

when D_Ω is the Bregman divergence defined in Definition 3.

Proof of Lemma 6. By the definition of Bregman divergence, we have

$$\begin{aligned} D_\Omega(\pi_c\|\pi_b) - D_\Omega(\pi_c\|\pi_a) + D_\Omega(\pi_b\|\pi_a) &= \Omega(\pi_c) - \Omega(\pi_b) - \big\langle \delta_c\Omega(\pi_b), \pi_c - \pi_b \big\rangle \\ &\quad - \Omega(\pi_c) + \Omega(\pi_a) + \big\langle \delta_c\Omega(\pi_a), \pi_c - \pi_a \big\rangle \\ &\quad + \Omega(\pi_b) - \Omega(\pi_a) - \big\langle \delta_c\Omega(\pi_a), \pi_b - \pi_a \big\rangle \\ &= \big\langle \delta_c\Omega(\pi_a) - \delta_c\Omega(\pi_b), \pi_c - \pi_b \big\rangle. \end{aligned}$$

Therefore, the proof is complete. □

Utilizing the three-point identity, we present the following useful lemmas for dealing inequalities regarding improvements by Han et al. (2022), which we call left and right Bregman differences.

Lemma 7 (Left Bregman difference)**.** *For all $\pi_a, \pi_b, \pi_c \in \mathcal{C} \cap \mathrm{dom}(\Omega)$, the following identity holds.*

$$D_\Omega(\pi_b \| \pi_a) - D_\Omega(\pi_c \| \pi_a) = -\big\langle \delta_c\Omega(\pi_c) - \delta_c\Omega(\pi_a), \pi_c - \pi_b \big\rangle + D_\Omega(\pi_b \| \pi_c). \tag{32}$$

Proof of Lemma 7. Using Lemma 6, we have

$$D_\Omega(\pi_b \| \pi_a) - D_\Omega(\pi_c \| \pi_a) = -D_\Omega(\pi_c \| \pi_b) + \big\langle \delta_c\Omega(\pi_a) - \delta_c\Omega(\pi_b), \pi_c - \pi_b \big\rangle.$$

Utilizing an identity of two Bregman divergences for arbitrary $(\rho, \bar{\rho})$:

$$D_\Omega(\rho \| \bar{\rho}) + D_\Omega(\bar{\rho} \| \rho) = \big\langle \delta_c\Omega(\rho) - \delta_c\Omega(\bar{\rho}), \rho - \bar{\rho} \big\rangle. \tag{33}$$

We separate $\delta_c\Omega(\pi_a) - \delta_c\Omega(\pi_b)$ into $\delta_c\Omega(\pi_a) - \delta_c\Omega(\pi_c)$ and $\delta_c\Omega(\pi_c) - \delta_c\Omega(\pi_b)$ and write the rest of the derivation as follows.

$$\begin{aligned}
&D_\Omega(\pi_b \| \pi_a) - D_\Omega(\pi_c \| \pi_a) \\
&\quad = \underbrace{-D_\Omega(\pi_c \| \pi_b) + \big\langle \delta_c\Omega(\pi_c) - \delta_c\Omega(\pi_b), \pi_c - \pi_b \big\rangle}_{\text{Eq. (33)}} + \big\langle \delta_c\Omega(\pi_a) - \delta_c\Omega(\pi_c), \pi_c - \pi_b \big\rangle \\
&\quad = D_\Omega(\pi_b \| \pi_c) + \big\langle \delta_c\Omega(\pi_a) - \delta_c\Omega(\pi_c), \pi_c - \pi_b \big\rangle
\end{aligned}$$

Therefore, we achieve the desired identity. □

Lemma 8 (Right Bregman difference)**.** *For all π_a, π_b, π_c, the following identity holds.*

$$D_\Omega(\pi_c \| \pi_b) - D_\Omega(\pi_c \| \pi_a) = D_\Omega(\pi_a \| \pi_b) + \big\langle \delta_c\Omega(\pi_a) - \delta_c\Omega(\pi_b), \pi_c - \pi_a \big\rangle \tag{34}$$

Proof of Lemma 8. By Lemma 6, we have

$$D_\Omega(\pi_c \| \pi_b) - D_\Omega(\pi_c \| \pi_a) = -D_\Omega(\pi_b \| \pi_a) + \big\langle \delta_c\Omega(\pi_a) - \delta_c\Omega(\pi_b), \pi_c - \pi_b \big\rangle.$$

We separate $\pi_c - \pi_b$ into $\pi_c - \pi_a$ and $\pi_a - \pi_b$ and write the rest of the derivation as follows.

$$\begin{aligned}
&D_\Omega(\pi_c \| \pi_b) - D_\Omega(\pi_c \| \pi_a) \\
&\quad = \underbrace{-D_\Omega(\pi_b \| \pi_a) + \big\langle \delta_c\Omega(\pi_a) - \delta_c\Omega(\pi_b),\, \pi_a - \pi_b \big\rangle}_{\text{Eq. (33)}} + \big\langle \delta_c\Omega(\pi_a) - \delta_c\Omega(\pi_b), \pi_c - \pi_a \big\rangle \\
&\quad = D_\Omega(\pi_a \| \pi_b) + \big\langle \delta_c\Omega(\pi_a) - \delta_c\Omega(\pi_b), \pi_c - \pi_a \big\rangle
\end{aligned}$$

Therefore, we achieve the desired identity. □

Additionally, we introduce the three-point inequality (Chen & Teboulle, 1993), which has been a key statement for proving MD convergence for a static cost functional (Aubin-Frankowski et al., 2022), and OMD improvement for temporal costs. The proof mostly follows Aubin-Frankowski et al. (2022) with a slight change of notation.

Lemma 9 (Three-point inequality)**.** *Given $\pi \in \mathcal{M}(\mathcal{X})$ and some proper convex functional $\Psi : \mathcal{M}(\mathcal{X}) \to \mathbb{R} \cup \{+\infty\}$, if $\delta_c\Omega$ exists, as well as $\bar{\rho} = \arg\min_{\rho \in \mathcal{C}} \{\Psi(\rho) + D_\Omega(\rho \| \pi)\}$, then for all $\rho \in \mathcal{C} \cap \mathrm{dom}(\Omega) \cap \mathrm{dom}(\Psi)$: $\Psi(\rho) + D_\Omega(\rho \| \pi) \geq \Psi(\bar{\rho}) + D_\Omega(\bar{\rho} \| \pi) + D_\Omega(\rho \| \bar{\rho})$.*

Proof of Lemma 9. The existence of $\delta_c\Omega$ implies $\mathcal{C} \cap \mathrm{dom}(D_\Omega(\cdot|y)) = \mathcal{C} \cap \mathrm{dom}(\Omega) \cap \mathrm{dom}(\Psi)$. Set $G(\cdot) = \Psi(\cdot) + D_\Omega(\cdot \| y)$. By linearity and idempotence, we have for any $\rho \in \mathcal{C} \cap \mathrm{dom}(\Omega) \cap \mathrm{dom}(\Psi)$

$$D_G(\rho \| \bar{\rho}) = D_\Psi(\rho \| \bar{\rho}) + D_\Omega(\rho \| \bar{\rho}) \geq D_\Omega(\rho \| \bar{\rho}). \tag{35}$$

By $\bar{\rho}$ being the optimality for G, for all $x \in \mathcal{C}$,

$$d^+G(\bar{\rho}; \rho - \bar{\rho}) = \lim_{h \to 0^+} \frac{G((1-h)\bar{\rho} + h\rho) - G(\bar{\rho})}{h} \geq 0,$$

which suggests $G(\rho) \geq G(\bar{\rho}) + D_G(\rho \| \bar{\rho})$. Applying (35) to this inequality complete the proof. □

The following argument is from the convergence rate of mirror descent for relatively smooth and convex pairs of functionals, and extend to infinite dimensional convergence results of Lu et al. (2018) and Aubin-Frankowski et al. (2022). We aim to reformulate the statements in online learning, addressing one-step improvement of OMD.

Lemma 10 (OMD improvement)**.** *Suppose a temporal cost $F_t : \mathcal{M}(\mathcal{X}) \to \mathbb{R}$ which is L-smooth and l-strongly-convex relative to Ω and $\eta_t \leq \frac{1}{L}$. Then, OMD improves for current cost $F_t(\pi_{t+1}) \leq F_t(\pi_t)$.*

Proof of Lemma 10. Since F is L relatively smooth, we initially have the inequality

$$F_t(\pi_{t+1}) \leq F_t(\pi_t) + d^+F(\pi_t; \pi_{t+1} - \pi_t) + LD_\Omega(\pi_{t+1}|\pi_t) \tag{36}$$

Applying the three-point inequality (Lemma 9) to Eq. (36), and setting a linear functional $\Psi(\rho) = \eta_t d^+F_t(\pi_t; \pi - \pi_t)$, $\rho = \pi_t$ and $\bar{\rho} = \pi_{t+1}$ yields

$$d^+F_t(\pi_t; \pi_{t+1} - \pi_t) + \tfrac{1}{\eta_t} D_\Omega(\pi_{t+1}|\pi_t) \leq d^+F_t(\pi_t; \rho - \pi_t) + \tfrac{1}{\eta_t} D_\Omega(\rho|\pi_t) - \tfrac{1}{\eta_t} D_\Omega(\rho\|\pi_{t+1}).$$

Since F_t is assumed to be l-strongly convex relative to Ω, we also have

$$d^+F(\pi_t; \rho - \pi_t) \leq F_t(\rho) - F_t(\pi_t) - lD_\Omega(\rho|\pi_t), \tag{37}$$

Then, by using (37), Eq. (36) becomes

$$F_t(\pi_{t+1}) \leq F_t(\rho) + (\tfrac{1}{\eta_t} - l)D_\Omega(\rho|\pi_t) - \tfrac{1}{\eta_t} D_\Omega(\rho|\pi_{t+1}) + (L - \tfrac{1}{\eta_t})D_\Omega(\pi_{t+1}\|\pi_t). \tag{38}$$

By substituting $\rho = \pi_t$, since $D_\Omega(\rho|\pi_{t+1}) \geq 0$ and $L - \frac{1}{\eta_t} \leq 0$, this shows $F_t(\pi_{t+1}) \leq F_t(\pi_t)$, *i.e.*, F_t is decreasing at each iteration. This completes the proof. □

A fundamental property with the dual space $\mathcal{D}$ induced by the first variation δ_c holds in our OMD setting. The existence of such learning sequence–particularly in Sinkhorn–is well discussed by Nutz (2021) and Aubin-Frankowski et al. (2022). Focusing on the dual geometry, we explicitly call this relationship with arbitrary step size η_t as "dual iteration."

Lemma 11 (Dual iteration)**.** *Suppose that first variations $\delta_c F_t(\pi_t)$ and $\delta_c \Omega(\pi_t)$ exists for $t \geq 0$. Then, online mirror descent updates Eq. (9) is equivalent to $\delta_c \Omega(\pi_{t+1}) - \delta_c \Omega(\pi_t) = -\eta_t \delta_c F_t(\pi_t)$, for all $\pi_t \in \mathcal{C}, t \in \mathbb{N}$.*

Proof of Lemma 11. The optimization (9) is equivalent to having the property for subsequent π_{t+1}:

$$\begin{aligned} &d^+F_t(\pi_t; \pi - \pi_t) + \tfrac{1}{\eta_t} D_\Omega(\pi\|\pi_t) \geq d^+F_t(\pi_t; \pi_{t+1} - \pi_t) + \tfrac{1}{\eta_t} D_\Omega(\pi_{t+1}|\pi_t) \\ &\iff \big\langle \delta_c F_t(\pi_t) - \tfrac{1}{\eta_t}\delta_c \Omega(\pi_t),\, \pi - \pi_{t+1} \big\rangle + \tfrac{1}{\eta_t}\big(\Omega(\pi) - \Omega(\pi_{t+1})\big) \geq 0, \quad \forall \pi \in \mathcal{C}. \end{aligned} \tag{39}$$

Setting the free parameter $\pi = \pi_{t+1} + h(\pi - \pi_{t+1})$ and taking the limit $h \to 0^+$ yields the result. □

Remark 3. With applications of Lemma 11 and Lemma 5, we can achieve a concise form of iteration in the dual using our temporal cost as:

$$\begin{aligned} \delta_c \Omega(\pi_t) - \delta_c \Omega(\pi_{t+1}) &= \eta_t \big(\delta_c(-H)(\pi_t) - \delta_c(-H)(\pi_t^\circ)\big) \\ &= \eta_t \big(\delta_c \Omega(\pi_t) - \delta_c \Omega(\pi_t^\circ)\big), \end{aligned} \tag{40}$$

where H denotes the entropy, *i.e.*, the minus KL divergence with the Lebesgue measure.

Leveraging the aforementioned lemmas, we have systematically introduced and rigorously formalized the essential concepts necessary to progress with our analysis within the OMD framework. Finally, we are ready to describe a suitable step size scheduling by the following arguments.

Lemma 12 (Step size I)**.** *Suppose that $F_t = \mathrm{KL}(\pi\|\pi_t^\circ)$ and $\Omega = \mathrm{KL}(\pi\|e^{-c_\varepsilon}\mu \otimes \nu)$. If ① $\lim_{t\to\infty} \eta_t = 0^+$ and ② $\sum_{t=1}^\infty \eta_t = +\infty$ ③ $\eta \leq \frac{1}{L}$, the OMD algorithm converges to a certain $\pi_{\mathcal{D}}^\circ$*

Proof of Lemma 12. From Lemma 10, we have

$$\eta_t(F_t(\pi_{t+1}) - F_t(\pi_t)) \leq -D_\Omega(\pi_t\|\pi_{t+1}) + (\eta_t L - 1)D_\Omega(\pi_{t+1}\|\pi_t). \tag{41}$$

Taking $\lim_{t\to\infty}\eta_t = 0$ ensures improvements; this means for any $\varepsilon > 0$ there exists some $0 < \delta \leq 1$ such that $D_\Omega(\pi_t\|\pi_{t+1}) + D_\Omega(\pi_{t+1}\|\pi_t) < \varepsilon$ whenever $\eta_t < \delta$. Since convexity and the lower semicontinuity of the Bregman divergence D_Ω induced by KL, we conclude that OMD to a certain point upon the assumed step size scheduling. □

Lemma 13 (Step size II)**.** *Assume that* $\min_{\pi\in\mathcal{C}}\mathbb{E}_t[D_\Omega(\pi_t,\pi_t^\circ)] > 0$ *for all* $t \in [1,\infty)$*. Suppose that* $\eta_t \to 0$ *and* $\lim_{T\to\infty}\mathbb{E}[\frac{1}{T}\sum_{t=1}^T D_\Omega(\pi_t\|\pi_t^\circ)] = 0$ *if and only if* $\sum_{t=1}^\infty \eta_t = +\infty$*.*

Proof of Lemma 13. We note that due to dual iteration equation Eq. (40), improvements on KL in Lemma 10 are also improvements in the Bregman divergence, *i.e.* $D_\Omega(\pi_{t+1}\|\pi_t^\circ) \leq D_\Omega(\pi_t\|\pi_t^\circ)$, and if $\eta_t \to 0$, then the process $\{\pi_t\}_{t=1}^\infty$ is convergent. By the dominated convergence theorem, assuming ergodicity of non-stationary $\{\pi_t^\circ\}_{t=1}^\infty$ (Cornfeld et al., 2012), there is a constant ε that satisfies $\mathbb{E}_{1:t+1}[D_\Omega(\pi_{t+1}\|\pi_{t+1}^\circ)] \geq \mathbb{E}_{1:t+1}[D_\Omega(\pi_{t+1}\|\pi_t^\circ)] + \varepsilon$ for $t > n$ for some n as $\eta_t \to 0$, where an expectation subscripted by "$1:t$" indicates the notation of time-averaging from 1 to t. Consequently, we achieve the following inequality

$$\begin{aligned}
&\mathbb{E}_{1:t+1}[D_\Omega(\pi_{t+1}\|\pi_{t+1}^\circ)] \\
&\quad\geq \mathbb{E}_{1:t+1}[D_\Omega(\pi_{t+1}\|\pi_t^\circ)] + \varepsilon \\
&\quad\geq \mathbb{E}_{1:t}\big[D_\Omega(\pi_t\|\pi_t^\circ) - \langle\delta_c\Omega(\pi_{t+1}) - \delta_c\Omega(\pi_t), \pi_t^\circ - \pi_t\rangle\big] + \mathbb{E}_{1:t+1}\big[D_\Omega(\pi_{t+1}\|\pi_t)\big] + \varepsilon && \text{Lemma 7} \\
&\quad= \mathbb{E}_{1:t}\big[D_\Omega(\pi_t\|\pi_t^\circ) - \eta_t D_\Omega(\pi_t\|\pi_t^\circ) + \eta_t D_\Omega(\pi_t^\circ\|\pi_t)\big] + \mathbb{E}_{1:t+1}\big[D_\Omega(\pi_{t+1}\|\pi_t)\big] + \varepsilon && \text{Eq. (40)} \\
&\quad= (1-\eta_t)\mathbb{E}_{1:t}\big[D_\Omega(\pi_t\|\pi_t^\circ)\big] + \mathbb{E}_{1:t+1}\big[D_\Omega(\pi_{t+1}\|\pi_t) + \eta_t D_\Omega(\pi_t^\circ\|\pi_t)\big] + \varepsilon \\
&\quad\geq (1-\eta_t)\mathbb{E}_{1:t}\big[D_\Omega(\pi_t\|\pi_t^\circ)\big] + \varepsilon'
\end{aligned} \tag{42}$$

for some t and $0 < \varepsilon < \varepsilon'$, where Lemma 7 and Eq. (40) are used.

Necessity. For big enough $t \geq n$ where $n \in \mathbb{N}$, we can achieve the inequality in Eq. (42) as

$$\mathbb{E}_{1:t+1}\big[D_\Omega(\pi_{t+1}\|\pi_{t+1}^\circ)\big] \geq (1-\eta_t)\mathbb{E}_{1:t}\big[D_\Omega(\pi_t\|\pi_t^\circ)\big], \tag{43}$$

Since we have assumed that η_t converges to 0, consider a step size sequence $0 < \eta_t \leq \frac{2}{2+k}$ for $k > 0$. Denote a constant $a = \frac{2+k}{2}\log\frac{2+k}{k}$ and apply the elementary inequality (Lei & Zhou, 2020)

$$1 - x \geq \exp(-ax), \quad \text{such that} \quad 0 < x \leq \frac{2}{2+k}.$$

From Eq. (43), we achieve

$$\mathbb{E}_{1:t+1}\big[D_\Omega(\pi_{t+1}\|\pi_{t+1}^\circ)\big] \geq \exp(-a\eta_t)\mathbb{E}_{1:t}\big[D_\Omega(\pi_t\|\pi_t^\circ)\big].$$

for all $t \geq n$. Iteratively applying this inequality iterative for $t = n, n+1, \cdots, T-1$ gives

$$\begin{aligned}
\mathbb{E}_{1:T}[D_\Omega(\pi_T\|\pi_T^\circ)] &\geq \mathbb{E}_{1:n}[D_\Omega(\pi_n\|\pi_n^\circ)]\prod_{t=n}^{T-1}\exp(-a\eta_t) \\
&= \exp\bigg\{-a\sum_{t=n}^{T-1}\eta_t\bigg\}\mathbb{E}_{1:n}[D_\Omega(\pi_n\|\pi_n^\circ)].
\end{aligned} \tag{44}$$

From the assumption $\pi^* \neq \pi_n$, $D_\Omega(\pi_n\|\pi_n^\circ) > 0$ by the property of divergence. Therefore, by Eq. (44), the convergence $\lim_{t\to\infty}\mathbb{E}_{1:t}[D_\Omega(\pi_t\|\pi_t^\circ)] = 0$ implies the series $\sum_{t=1}^\infty \eta_t$ diverges to $+\infty$ so that $\exp(-a\sum_{t=n}^{T-1}\eta_t)$ converges to 0.

Sufficiency. Consider a static Schrödinger bridge problem with couplings $\pi \in \Pi(\mu,\nu)$, which is in a constraint set

$$\mathcal{C} = \big\{\pi|(\mu,\nu) \in \mathcal{P}_2(\mathbb{R}^d) \cap \mathcal{P}_{\text{alc}}(\mathbb{R}^d), (\varphi,\psi) \in L^1(\mu)\times L^1(\nu), \text{and}\, \varphi,\psi \in C^2(\mathbb{R}^d) \cap \text{Lip}(\mathcal{K})\big\}.$$

For $\rho, \bar{\rho} \in \mathcal{C}$, we can see

$$D_\Omega(\bar{\rho}\|\rho) = \Omega(\bar{\rho}) - \Omega(\rho) - \langle \delta_c\Omega(\rho), \bar{\rho} - \rho\rangle \geq 0 \iff -\langle \delta_c\Omega(\rho), \bar{\rho} - \rho\rangle \geq \Omega(\rho) - \Omega(\bar{\rho}).$$

By adding $\langle \delta_c\Omega(\bar{\rho}), \bar{\rho} - \rho\rangle$, we achieve a property:

$$\langle \delta_c\Omega(\rho) - \delta_c\Omega(\bar{\rho}), \rho - \bar{\rho}\rangle \geq D_\Omega(\rho\|\bar{\rho}). \tag{45}$$

Then, suppose that we have the asymptotic dual mean $\pi_{\mathcal{D}}^\circ$. Using the right Bregman difference Lemma 8, the one-step progress from the perspective of dual mean writes as

$$\begin{aligned}
D_\Omega(\pi_{\mathcal{D}}^\circ\|\pi_{t+1}) - D_\Omega(\pi_{\mathcal{D}}^\circ\|\pi_t) &= \big\langle \delta_c\Omega(\pi_t) - \delta_c\Omega(\pi_{t+1}), \pi_{\mathcal{D}}^\circ - \pi_t\big\rangle + D_\Omega(\pi_t\|\pi_{t+1}). \\
&= \eta_t\big\langle \delta_c\Omega(\pi_t) - \delta_c\Omega(\pi_t^\circ), \pi_{\mathcal{D}}^\circ - \pi_t\big\rangle + D_\Omega(\pi_t\|\pi_{t+1}) \\
&= \eta_t\big\langle \delta_c\Omega(\pi_t) - \delta_c\Omega(\pi_{\mathcal{D}}^\circ), \pi_{\mathcal{D}}^\circ - \pi_t\big\rangle + \eta_t\big\langle \delta_c\Omega(\pi_{\mathcal{D}}^\circ) - \delta_c\Omega(\pi_t^\circ), \pi_t^\circ - \pi_t\big\rangle + D_\Omega(\pi_t\|\pi_{t+1}) \\
&\leq -\eta_t D(\pi_{\mathcal{D}}^\circ\|\pi_t) + \eta_t\big\langle \delta_c\Omega(\pi_{\mathcal{D}}^\circ) - \delta_c\Omega(\pi_t^\circ), \pi_t^\circ - \pi_t\big\rangle + D_\Omega(\pi_t\|\pi_{t+1})
\end{aligned}$$

where the inequality is from Eq. (45). By applying the definition of $\pi_{\mathcal{D}}^\circ$ and ergodicity of $\{\pi_t^\circ\}_{t=1}^\infty$, we can bound the expectation by finding some $t > n$ such that

$$\begin{aligned}
\mathbb{E}_{1:t+1}\big[D_\Omega(\pi_{\mathcal{D}}^\circ\|\pi_{t+1})\big] &\leq \mathbb{E}_{1:t}\big[(1-\eta_t)D_\Omega(\pi_{\mathcal{D}}^\circ\|\pi_t)\big] + \mathbb{E}_{1:t+1}\big[D_\Omega(\pi_t\|\pi_{t+1})\big] \\
&\leq \mathbb{E}_{1:t}[(1-\eta_t)D_\Omega(\pi_{\mathcal{D}}^\circ\|\pi_t)] + \frac{1}{2\omega}\mathbb{E}_{1:t+1}\big[\big\|\nabla(\delta_c\Omega(\pi_t) - \delta_c\Omega(\pi_{t+1}))\big\|_{L^2(\pi_t)}^2\big] \\
&\leq \mathbb{E}_{1:t}[(1-\eta_t)D_\Omega(\pi_{\mathcal{D}}^\circ\|\pi_t)] + \frac{\eta_t^2}{2\omega}\mathbb{E}_{1:t}\big[\|\nabla(\delta_c\Omega(\pi_t) - \delta_c\Omega(\pi_t^\circ))\|_{L^2(\pi_t)}\big] \\
&\leq \mathbb{E}_{1:t}[(1-\eta_t)D_\Omega(\pi_{\mathcal{D}}^\circ\|\pi_t)] + 2\eta_t^2\omega^{-1}\mathcal{K},
\end{aligned} \tag{46}$$

where $\mathcal{K}$ is the Lipschitz constant for each log Schrödinger potential in $\mathcal{C}$. For the first inequality, we use Assumption 2, and we use the log Sobolev inequality LSI(ω) from Assumption 1 in the second inequality. Let $\{A_t\}_{t=1}^\infty$, denote a sequence of $A_t = \mathbb{E}_{1:t}[D_\Omega(\pi_{\mathcal{D}}^\circ\|\pi_t)]$. As a result, we have

$$A_{t+1} \leq (1-\eta_t)A_t + z\eta_t^2, \quad \forall t > n, \tag{47}$$

where $z \coloneqq 2\omega^{-1}\mathcal{K}$. For a constant $h > 0$, we argue that $A_{t_1} < h$ for some $t_1 > n'$. Suppose that this statement is *not* true; we find some $t \geq t_1$ such that $A_t > h$, $\forall t \geq t_2$. Since $\lim_{t\to\infty}\eta_t = 0$, there are some $t > t_3 > t_2$ that $\eta_t \leq \frac{h}{4}$. However, Eq. (47) tells us that for $t \geq t_3$, for $t \geq t_3$,

$$A_{t+1} \leq (1-\eta_t)A_t + z\eta_t^2 \leq A_{t_3} - \frac{h}{4}\sum_{k=t_3}^{T}\eta_k \to -\infty \quad (\text{as } t \to \infty).$$

This results to a contradiction, which verifies $A_t < h$ for $t > n'$. Since $\lim_{t\to\infty}\eta_t = 0$, we can find some η_t which makes A_t monotonically decreasing. Therefore, we conclude the nonnegative sequence $\{A_t\}_{t=1}^\infty$ finds convergence by iteratively applying the upper bound in Eq. (47).

We now prove the theorem under consideration of the particular case of $\eta_t = \frac{2}{t+1}$. Then, Eq. (47) becomes

$$A_{t+1} \leq \left(1 - \frac{2}{t+1}\right)A_t + \frac{4z}{(t+1)^2}, \quad \forall t \geq n.$$

It follows that recursive relation writes as

$$t(t+1)A_{t+1} \leq (t-1)tA_t + 4z, \quad \forall t \geq n.$$

Iterative applying the relation, we achieve the following inequality:

$$(T-1)TA_T \leq (n-1)nA_n + 4z(T-n), \quad \forall T \geq n.$$

Therefore, we finally achieve inequality as follows:

$$\mathbb{E}_{1:T}[D_\Omega(\pi_{\mathcal{D}}^\circ\|\pi_T)] \leq \frac{(n-1)n\mathbb{E}_{1:n}[D_\Omega(\pi_{\mathcal{D}}^\circ\|\pi_n)]}{(T-1)T} + \frac{4z}{T}, \quad \forall T \geq n. \tag{48}$$

Since we assumed $\pi^* = \pi_{\mathcal{D}}^\circ$, $\mathbb{E}_{1:T}[D_\Omega(\pi^*\|\pi_T)] = \mathcal{O}(1/T)$, the proof of Theorem 1 is complete. □

A.3 Proof of Proposition 1

The proof is based on the Doob's forward convergence theorem.

Theorem 3 (Doob's forward convergence theorem)**.** *Let $\{X_t\}_{t\in\mathbb{N}}$ be a sequence of nonnegative random variables and let $\{\mathcal{F}_t\}_t$ be a random variable and let $\{\mathcal{F}_t\}_{t\in\mathbb{N}}$ be a filtration with $\mathcal{F}_t \subset \mathcal{F}_{t+1}$ for every $t \in \mathbb{N}$. Assume that $\mathbb{E}[X_{t+1}|\mathcal{F}_t] \leq X_t$ almost surely for every $t \in \mathbb{N}$. Then, the sequence $\{X_t\}$ converges to a nonnegative random variable X_∞ almost surely.*

We follow the derivation of Eq. (46): there exists $n \in \mathbb{N}$ which satisfies

$$\mathbb{E}_t[D_\Omega(\pi^\circ_{\mathcal{D}}\|\pi_{t+1})] \leq D_\Omega(\pi^\circ_{\mathcal{D}}\|\pi_t) + 2\eta_t^2\omega^{-1}\mathcal{K}, \quad \forall t \geq n$$

and since the step size is scheduled as $\lim_{t\to\infty}\eta_t = 0$, the condition $\sum_{t=1}^\infty \eta_t^2 < \infty$ enables us to define a stochastic process $\{X_t\}_{t\in\mathbb{N}}$:

$$X_t = D_\Omega(\pi^\circ_{\mathcal{D}}\|\pi_t) + 2\omega^{-1}\mathcal{K}\sum_{i=t}^\infty \eta_i^2. \tag{49}$$

It is straightforward that the defined random variable satisfies $\mathbb{E}_t[X_{t+1}] \leq X_t$ for $t \geq n$. Since $X_t \geq 0$, the process is a sub martingale. By Theorem 3, the sequence $\{X_t\}_{t\in\mathbb{N}}$ converges to a nonnegative random variable X_∞ almost surely. Therefore $D_\Omega(\pi^\circ_{\mathcal{D}}\|\pi_t)$ converges to 0 almost surely. □

A.4 Proof of Proposition 2

To achieve a meaningful regret bound for our problem setup, we first demonstrate the following.

Lemma 14. *For all $w = \arg\min_y\{\langle\hat{g}, y\rangle + \frac{1}{\eta}D_\Omega(y\|z)\}$ with $\eta > 0$, the following equation.*

$$\forall u. \langle\eta\hat{g}, w - u\rangle \leq D_\Omega(u\|z) - D_\Omega(u\|w) - D_\Omega(w\|z) \tag{50}$$

Proof of Lemma 14. By the first order optimality of $\{\langle g, y\rangle + D_\Omega(y\|z)\}$ as a function of w, we have

$$\begin{aligned}&\langle\hat{g} + \tfrac{1}{\eta}\delta_c D_\Omega(w\|z), u - w\rangle \geq 0\\ &\quad\implies \langle\hat{g}, w - u\rangle \leq \tfrac{1}{\eta}\langle -\delta_c D_\Omega(w\|z), w - u\rangle = \tfrac{1}{\eta}(D_\Omega(u\|z) - D_\Omega(u\|w) - D_\Omega(w\|z)).\end{aligned}$$

where used Lemma 7 in the derivation. This completes the proof. □

Next, we derive the one-step relationship for OMD. The result entails that the regret at each step is related to a quadratic expression of η_t, which is a key aspect of sublinear total regret. From a technical standpoint, we can see that the assumption for log Sobolev inequality generally works as a premise for Lipschitz continuity of gradient, *i.e.*, $\nabla\Omega$ in classical MD analyses.

Lemma 15 (Single step regret)**.** *Suppose a static Schrödinger bridge problem with the aforementioned constraint $\mathcal{C}$. Let D_Ω be the Bregman divergence wrt $\Omega : \mathcal{P}(\mathcal{X}) \to \mathbb{R} + \{+\infty\}$. Then,*

$$\eta_t(F_t(\pi_t) - F_t(u)) \leq D_\Omega(u\|\pi_t) - D_\Omega(u\|\pi_{t+1}) + \frac{\eta_t^2}{2\omega}\|\hat{g}_t\|^2_{L^2(\pi_t)}, \quad \forall u \in \mathcal{C} \tag{51}$$

holds, where $\hat{g}_t \coloneqq \delta_c F_t(\pi_t) = \frac{1}{\eta_t}(\delta_c\Omega(\pi_t) - \delta_c\Omega(\pi_{t+1}))$ in an MD iteration for the dual space for a step size η_t, and $\omega > 0$ is drawn from a type of log Sobolev inequality in Assumption 1.

Proof of Lemma 15. Consider single step regrets by the adversary plays of a linearization for $\hat{g}_t$:

$$F_t(\pi_t) - F_t(u) \leq \langle\hat{g}_t, \pi_t - u\rangle.$$

Therefore, we derive a inequality for $\langle\hat{g}_t, \pi_t - u\rangle$ as follows.

$$\begin{aligned}\langle\eta_t\hat{g}_t, \pi_t - u\rangle &= \langle\eta_t\hat{g}_t, \pi_{t+1} - u\rangle + \langle\eta_t\hat{g}_t, \pi_t - \pi_{t+1}\rangle\\ &\leq D_\Omega(u\|\pi_t) - D_\Omega(u\|\pi_{t+1}) - D_\Omega(\pi_{t+1}\|\pi_t) + \langle\eta_t\hat{g}_t, \pi_t - \pi_{t+1}\rangle\\ &= D_\Omega(u\|\pi_t) - D_\Omega(u\|\pi_{t+1}) - D_\Omega(\pi_{t+1}\|\pi_t) + \langle\delta_c\Omega(\pi_{t+1}) - \delta_c\Omega(\pi), \pi_t - \pi_{t+1}\rangle\\ &= D_\Omega(u\|\pi_t) - D_\Omega(u\|\pi_{t+1}) + D_\Omega(\pi_t\|\pi_{t+1}).\end{aligned}$$

Since we assumed that $\hat{g}_t = \frac{1}{\eta_t}(\delta_c\Omega(\pi_t) - \delta_c\Omega(\pi_{t+1}))$ by the dual iteration and that Assumption 1 holds, we can achieve the upperbound $D_\Omega(\pi_t\|\pi_{t+1}) \leq \frac{\eta_t^2}{2\omega}\|\hat{g}_t\|^2_{L^2(\pi_t)}$ by direct calculation. □

We now show our upper bound of total regret by utilizing Lemma 15.

Lemma 16. *Assume $\eta_{t+1} \leq \eta_t$. Then, $u \in \mathcal{C}$, the following regret bounds for fixed $u \in \mathcal{C}$ hold*

$$\sum_{t=1}^{T} F_t(\pi_t) - F_t(u) \leq \max_{1\leq t\leq T} \frac{D_\Omega(u\|\pi_t)}{\eta_T} + \frac{1}{2\omega}\sum_{t=1}^{T}\eta_t\|\tilde{g}_t\|^2_{L^2(\pi_t)} \tag{52}$$

where $\hat{g}_t = \frac{1}{\eta_t}(\delta_c\Omega(\pi_t) - \delta_c\Omega(\pi_{t+1}))$.

Proof of Lemma 16. Define $D^2 = \max_{1\leq t\leq T} D_\Omega(u\|\pi_t)$. We get

$$\begin{aligned}
\mathrm{Regret}(u) &= \sum_{t=1}^{T}(F_t(\pi_t) - F_t(u)) \\
&\leq \sum_{t=1}^{T}\left(\frac{1}{\eta_t}D_\Omega(u\|\pi_t) - \frac{1}{\eta_t}D_\Omega(u\|\pi_{t+1})\right) + \sum_{t=1}^{T}\frac{\eta_t}{2\omega}\|\hat{g}_t\|^2_{L^2(\pi_t)} \\
&= \frac{1}{\eta_1}D_\Omega(u\|\pi_1) - \frac{1}{\eta_T}D_\Omega(u\|\pi_{T+1}) + \sum_{t=1}^{T-1}\left(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\right)D_\Omega(u\|\pi_{t+1}) + \sum_{t=1}^{T}\frac{\eta_t}{2\omega}\|\hat{g}_t\|^2_{L^2(\pi_t)} \\
&\leq \frac{1}{\eta_1}D^2 + D^2\sum_{t=1}^{T-1}\left(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\right) + \sum_{t=1}^{T}\frac{\eta_t}{2\omega}\|\hat{g}_t\|^2_{L^2(\pi_t)} = \frac{D^2}{\eta_T} + \sum_{t=1}^{T}\frac{\eta_t}{2\omega}\|\hat{g}_t\|^2_{L^2(\pi_t)}.
\end{aligned}$$

Therefore, the proof is complete. □

Following Lemma 16 and Assumption 1, we can have the inequality

$$\sum_{t=1}^{T} F_t(\pi_t) - F_t(u) \leq \frac{D^2}{\eta_T} + \sum_{t=1}^{T}\frac{\eta_t}{2\omega}\|\hat{g}_t\|^2_{L^2(\pi_t)} \leq \frac{D^2}{\eta_T} + 2\eta_1\omega^{-1}\mathcal{K}T.$$

where $D^2 = \max_{1\leq t\leq T} D_\Omega(u\|\pi_t)$. Setting a constant step size $\eta_t \equiv \frac{D\sqrt{\omega}}{\sqrt{2\mathcal{K}T}}$ yields an upper bound of $2D\sqrt{2\omega^{-1}\mathcal{K}T}$ which proves the regret bound of $\mathcal{O}(\sqrt{T})$. Also, recall that the following lemma.

Lemma 17 (Lemma 3.5 of Auer et al., 2002)**.** *Let a sequence $a_1, a_2, \ldots, a_T$ be non-negative real numbers. If $a_1 > 0$, then*

$$\sum_{t=1}^{T}\frac{a_t}{\sqrt{\sum_{i=1}^{t} a_i}} \leq 2\sqrt{\sum_{t=1}^{T} a_t}. \tag{53}$$

Setting a adaptive scheduling $\eta_t = \frac{D\sqrt{\omega}}{\sqrt{2\sum_{i=1}^{t}\|\hat{g}_i\|^2}}$ yields $2D\sqrt{2\omega^{-1}\sum_{t=1}^{T}\|\hat{g}_t\|^2}$ which has a possibility to be lower than $\mathcal{O}(\sqrt{T})$ depending on $\{\pi_t^\circ\}_{t=1}^T$. Therefore, we have formally expanded the convergence results of OMD (Lei & Zhou, 2020; Orabona & Pál, 2018; Srebro et al., 2011) to SBPs. □

A.5 Proof of Theorem 2

Since $D_\Omega(\cdot\|\cdot) \coloneqq D_{\mathrm{KL}(\cdot\|\mathcal{R})}(\cdot\|\cdot)$ for a reference measure $\mathcal{R} \in \mathcal{C}$, we can apply Lemma 5 and achieve Eq. (13). We write the following equivalent convex problems, using the equivalence of first variation for recursively

defined Bregman divergences.

$$\begin{aligned}\left\langle \delta_c F_t(\pi_t), \pi - \pi_t \right\rangle + \tfrac{1}{\eta_t} D_\Omega(\pi \| \pi_t) &= \left\langle \delta_c D_\Omega(\pi_t \| \pi_t^\circ), \pi - \pi_t \right\rangle + \tfrac{1}{\eta_t} D_\Omega(\pi \| \pi_t) \\ &= \left\langle \delta_c \Omega(\pi_t) - \delta_c \Omega(\pi_t^\circ), \pi - \pi_t \right\rangle + \tfrac{1}{\eta_t} D_\Omega(\pi \| \pi_t) \\ &= D_\Omega(\pi \| \pi_t^\circ) - D_\Omega(\pi \| \pi_t) + \tfrac{1}{\eta_t} D_\Omega(\pi \| \pi_t) \\ &= \left(\frac{1}{\eta_t}\right) D_\Omega(\pi \| \pi_t^\circ) + \left(\frac{1-\eta_t}{\eta_t}\right) D_\Omega(\pi \| \pi_t)\end{aligned}$$

We refer to Appendix B for the stability of Wasserstein gradient flows according to the LaSalle's invariance principle. We can now interpret $\delta_c \mathcal{E}_t$ as a dynamics that reaches an equilibrium solution

$$\underset{\pi \in \mathcal{C}}{\text{minimize}} \left\langle \delta_c F_t(\pi_t), \pi - \pi_t \right\rangle + \tfrac{1}{\eta_t} D_\Omega(\pi \| \pi_t) \quad \Leftrightarrow \quad \underset{\pi \in \mathcal{C}}{\text{minimize}}\ \eta_t \underbrace{D_\Omega(\pi \| \pi_t^\circ)}_{\text{empirical estimates}} + (1-\eta_t) \underbrace{D_\Omega(\pi \| \pi_t)}_{\text{proximity}},$$

At a glance, the above equation appears analogous to the interpolation search between two points, where the influence of π_t° is controlled by η_t. □

A.6 Proof of Proposition 3

The proof is closely related to the work of Lambert et al. (2022) where the difference lies in we correct the Wasserstein gradient term $\dot{\alpha}_{k,\tau}$ for suitable for generally unbalanced weight. Suppose take parameterization $\theta \in (\mathcal{P}_2(\texttt{BW}(\mathbb{R}^d)), \texttt{WFR})$, the space of Gaussian mixtures equipped with the Wasserstein-Fisher-Rao metric, over the measure space of Gaussian particles. Following the arguments from Appendix B.2 and the studies for this particular GMM problem (Lu et al., 2019; Lambert et al., 2022) of the Wasserstein-Fisher-Rao of the KL functional is derived as

$$\nabla_{\texttt{WFR}} \mathrm{KL}(\rho_\theta \| \rho^*) = \left(\nabla_{\texttt{BW}} \delta \mathrm{KL}(\rho \| \rho^*), \frac{1}{2} \left(\delta \mathrm{KL}(\rho_\theta \| \rho^*) - \int \delta \mathrm{KL}(\rho \| \rho^*) \mathrm{d}\rho \right) \right), \tag{54}$$

where we can consider the WFR gradient is taken with respect to θ of its first argument. By Eq. (54), we separately consider Wasserstein gradient in the Bures–Wasserstein space and the space of lighting that controls the amount of each Gaussian particle.

Given a functional $F : \mathcal{P}_2(\mathcal{X}) \to \mathbb{R} \cup \{+\infty\}$, the Wasserstein gradient $\nabla_{\mathbb{W}} F \cap T_\rho \mathcal{P}_2(\mathcal{X})$ such that all $\{\rho_t\}_{t \in \mathbb{R}^+}$ satisfy the continuity eqatuion starting from ρ_0 (Jordan et al., 1998; Villani, 2021). If the functional is the KL divergence $\mathrm{KL}(\rho \| \pi)$ we can compute the Bures–Wasserstein gradient for the Gaussian distribution with respect to (m, Σ) using Eq. (71)

$$\begin{aligned}\nabla_{\texttt{BW}} F(m, \Sigma) &= (\nabla_m F(m, \Sigma), 2 \nabla_\Sigma F(m, \Sigma)) \\ &= \left(\int \nabla_m \rho_{m,\Sigma} \log \frac{\rho_{m,\Sigma}}{\pi}, 2 \int \nabla_\Sigma \rho_{m,\Sigma} \log \frac{\rho_{m,\Sigma}}{\pi} \right),\end{aligned}$$

with some abuse of notation for ρ. Using the following closed-form identities for the Gaussian distributions

$$\forall x. \quad \nabla_m \rho_{m,\Sigma}(x) = -\nabla_x \rho_{m,\Sigma}(x) \quad \text{and} \quad \nabla_\Sigma \rho_{m,\Sigma}(x) = \frac{1}{2} \nabla_x^2 \rho_{m,\Sigma}(x).$$

and the equivalence between the Hessian and Fisher information, we achieve the following form:

$$\nabla_{\texttt{BW}} F(m, \Sigma) = \left(\mathbb{E}_\rho \left[\nabla \frac{\rho}{\pi} \right], \mathbb{E}_\rho \left[\nabla^2 \log \frac{\rho}{\pi} \right] \right).$$

Define $r_{k,\tau} = \sqrt{\alpha_{k,\tau}}$. Since r_t follows the Fisher–Rao metric in Definition 7, by the Proposition A.1 from Lu et al. (2019) and specialization of Lambert et al. (2022), we can think of dynamics of K Gaussian particles $\{\alpha_{k,\tau}, m_{k,\tau}, \Sigma_{k,\tau}\}_{k=1}^K$ such that

$$\dot{r}_{k,\tau} = -\frac{1}{2} \left(\mathbb{E}\left[\log \frac{\rho_{\theta_\tau}}{\rho^*}(y_{k,\tau}) \right] - \frac{1}{z_\tau} \sum_{\ell=1}^K \alpha_\ell \mathbb{E}\left[\log \frac{\rho_{\theta_\tau}}{\rho^*}(y_{\ell,\tau}) \right] \right) r_{k,\tau},$$

$$\dot{m}_{k,\tau} = -\mathbb{E}\left[\nabla \log \frac{\rho_{\theta_\tau}}{\rho^*}(y_{k,\tau}) \right], \ \dot{\Sigma}_{k,\tau} = -\mathbb{E}\left[\nabla^2 \log \frac{\rho_{\theta_\tau}}{\rho^*}(y_{k,\tau}) \right] \Sigma_{k,\tau} - \Sigma_{k,\tau} \mathbb{E}\left[\nabla^2 \log \frac{\rho_{\theta_\tau}}{\rho^*}(y_{k,\tau}) \right],$$

Since $\alpha_{k,\tau} = \sqrt{r_{k,\tau}}$ by previous definition, it is straightforward that

$$\dot{\alpha}_{k,\tau} = -\left(\mathbb{E}\left[\log \frac{\rho_{\theta_\tau}}{\rho^*}(y_{k,\tau})\right] - \frac{1}{z_\tau}\sum_{\ell=1}^{K}\alpha_\ell \mathbb{E}\left[\log \frac{\rho_{\theta_\tau}}{\rho^*}(y_{\ell,\tau})\right]\right)\alpha_{k,\tau}.$$

For $\alpha_k > 0$. This completes the proof. □

B A Riemannian Perspective on Wasserstein Geometries

B.1 An introduction to Otto calculus and the LaSalle invariance principle

In this appendix, we introduce a basic notion of Wasserstein gradient flows in the space of continuous probability measures. We focus on describing the particular example, the KL cost, initially studied by JKO (Jordan et al., 1998) and formally generalized by Otto (2001) in the context of Riemannian geometry. For more details and mathematical rigor, we refer the reader to (Ambrosio et al., 2005b; Carrillo et al., 2023). For $\mathcal{X} \subset \mathbb{R}^d$, and functions $U : \mathbb{R}_{\geq 0} \to \mathbb{R}$; $V, W : \mathcal{X} \to \mathbb{R}$. We first consider an energy function $\mathcal{E} \colon \mathcal{P}_2(\mathcal{X}) \to \mathbb{R}$:

$$\mathcal{E}(\rho) = \underbrace{\int_{\mathcal{X}} U\big(\rho(x)\big)\,\mathrm{d}x}_{\text{internal potential } \mathcal{U}} + \underbrace{\int_{\mathcal{X}} V(x)\,\mathrm{d}\rho(x)}_{\text{external potential } \mathcal{E}_V} + \underbrace{\frac{1}{2}\int_{\mathcal{X}} (W * \rho)(x)\,\mathrm{d}\rho(x)}_{\text{interaction energy } \mathcal{W}}, \quad \rho \in \mathcal{P}_2(\mathcal{X}). \tag{55}$$

For this function, we refer to the solution of the following PDE:

$$\partial_t \rho_t = \nabla \cdot \big[\rho\, \nabla (U' + V + W * \rho)\big], \qquad t \geq 0 \tag{56}$$

as the Wasserstein gradient flow of $\mathcal{E}$. Following Otto's formalization of Riemannian calculus on the continuous probability space equipped with the Wasserstein metric $(\mathcal{P}_2(\mathcal{X}), W_2)$, the PDE (56) can be interpreted close to an ODE of Riemannian gradient flow:

$$\partial_t \rho_t = -\nabla_{\mathbb{W}} \mathcal{E}(\rho), \tag{57}$$

where $\nabla_{\mathbb{W}}$ denotes the Wasserstein-2 gradient operator $\nabla_{\mathbb{W}} := \nabla \cdot \big(\rho\, \nabla \frac{\delta}{\delta \rho}\big)$. Considering the Otto's Wasserstein-2 Riemannian metric $\mathfrak{g}$ (Otto, 2001; Lott, 2008), under the absolute continuity, we see that

$$\frac{\partial}{\partial t}\mathcal{E}(\rho_t) = -\mathfrak{g}_\rho\left(\frac{\partial \rho}{\partial t}, \frac{\partial \rho}{\partial t}\right) = -\int_{\mathcal{X}} \big|\nabla (U' + V + W * \rho)\big|^2 \mathrm{d}\rho(x) \leq 0, \tag{58}$$

which is closely related to the strict Lyapunov condition. As a result, dynamical systems following the PDE are guaranteed to reach an equilibrium solution, under the LaSalle invariance principle for probability measures (Carrillo et al., 2023).

For a representative example, we identify Eq. (55) for the relative entropy (the KL functional) for a target density $\rho^* \in \mathcal{P}_2(\mathcal{X})$ writes

$$\mathcal{E}(\rho) = \mathrm{KL}(\rho \| \rho^*) = \underbrace{\int_{\mathcal{X}} U\big(\rho(x)\big)\,\mathrm{d}x}_{\mathcal{U}} + \underbrace{\int_{\mathcal{X}} V(x)\,\mathrm{d}\rho(x)}_{\mathcal{E}_V} - C,$$

where $U(s) = s \log s$, $V(x) = -\log \rho^*(x)$, and $C = \mathcal{U}(\rho^*) + \mathcal{E}_V(\rho^*)$. Recall that $\delta \mathcal{E}(\rho) = \log \frac{\rho(x)}{\rho^*}$, then we have

$$\nabla_{\mathbb{W}} \mathcal{E}(\rho) = \mathfrak{G}_\rho^{-1} \delta E(\rho) = -\nabla \cdot [\rho \nabla \delta E(\rho)] = \nabla \cdot \left[\rho \nabla \log \frac{\rho}{\rho_*}\right] \tag{59}$$

where $\mathfrak{G}$ denotes the metric tensor in matrix form. We can derive the the Fokker–Planck equation

$$\partial_t \rho_t = -\nabla \cdot (\rho \nabla \log \rho^*) + \Delta \rho_t,$$

describing the time evolution of the probability density. Combining the convexity of KL and the LaSalle invariance principle Wasserstein gradient flows, the PDE reaches a unique stationary solution of $\frac{e^{-V(x)}}{\int_{\mathcal{X}} e^{-V(y)} \mathrm{d}y}$.

B.2 Background on Wasserstein-Fisher-Rao and other related geometries

The Wasserstein-Fisher-Rao geometry is also known as *Hellinger–Kantorovich* in some of papers (Liero et al., 2016; 2018). In this section, we provide an overview of the geometry tailored to meet our technical needs. Along the way, we also briefly describe various metrics and geometries related to the Wasserstein space.

The Wasserstein space. Let $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ be a probability densities with respect to the Lebesgue measure. We define the squared Wasserstein distance by a problem of couplings (Villani, 2009)

$$W_2^2(\mu, \nu) \coloneqq \min_{\pi \in \Pi(\mu,\nu)} \int_{\mathbb{R}^2 \times \mathbb{R}^2} \frac{1}{2} \|x - y\|^2 \mathrm{d}\pi(x, y). \tag{60}$$

Then, the Brenier theorem (Villani, 2021) states that there exists an optimal mapping function that pushes forward μ to ν, *i.e.* $\nu = \nabla\zeta_{\#}\mu$, where $\zeta : \mathbb{R}^d \to \mathbb{R}^d \cup \{+\infty\}$ is a convex and lower semicontinuous function. In optimal transport, the equation is generally referred to as the Monge–Ampere equation. In the fluid dynamics, the Brenier map yields a constant-speed of geodesic $\{\mu_t\}_{t\in[0,1]}$ formally described by the following differential equation

$$\rho_t = (\nabla\zeta_t)_{\#}\mu, \qquad \nabla\zeta_t \coloneqq (1 - t)\mathrm{id} + t\nabla\zeta. \tag{61}$$

Assuming the existence of such geodesic, we can understand finding optimality of Eq. (61) the Benamou–Brenier formulation (Benamou & Brenier, 2000), which finds a velocity v_t by minimizing the L^2 cost

$$W_2^2(\mu, \nu) = \min_{\rho,v} \left\{ \int_0^1 \int_{\mathbb{R}^d} \frac{1}{2} \|v_t(x)\|^2 \mathrm{d}\rho_t(x)\mathrm{d}t \;\middle|\; \rho_0 = \mu,\ \rho_1 = \nu,\ \partial_t\rho_t = -\nabla \cdot (v_t\rho_t) \right\}. \tag{62}$$

The equation dictates *how* the fluid should be transported (which shall be controlled by speed v_t) while satisfying the continuity equation of path measure on the right hand side. In the Otto calculus (Otto, 2001), we can understand the Benamou–Brenier formula (62) as a Riemannian formulation for W_2. In this interpretation, the tangent space at $\rho \in \mathcal{P}_2(\mathcal{X})$ are measures of the form $\delta\rho = -\nabla \cdot (v\rho)$ with a velocity field $v \in L^2(\rho, \mathbb{R}^d)$ and the metric is given by

$$\|\rho\|_\rho^2 = \inf_{v \in L^2(\rho,\mathbb{R}^d)} \left\{ \int \|v\|^2 \mathrm{d}\rho \;\middle|\; \delta\rho = -\nabla \cdot (v\rho) \right\}. \tag{63}$$

The Benamou–Brenier formula exhibits dynamics in the Wasserstein space of probability densities metric generally governed by the continuity equation, implying the mass of probability is preserved.

Fisher-Rao metric. The Fisher–Rao metric is a metric on the space of positive measures $\mathcal{P}_+$ with possibly different total masses. We are interested in the simple case where such measure are represented with a fininte number of parameters such as exponential families. We use the following definition throughout the paper.

Definition 7 (Fisher–Rao metric)**.** The Fisher–Rao distance between measures $\rho_0, \rho_1 \in \mathcal{M}_+$ is given by

$$d_{\mathtt{FR}}^2(\rho_0, \rho_1) \coloneqq \min_{\rho,v \in \mathcal{A}[\rho_0,\rho_1]} \int_0^1 \int_{\mathbb{R}^d} \frac{1}{2} \omega_t^2(x) \mathrm{d}\rho_t(x)\mathrm{d}t = 2 \int_{\mathbb{R}^d} \left| \sqrt{\frac{\mathrm{d}\rho_0}{\mathrm{d}\lambda}} - \sqrt{\frac{\mathrm{d}\rho_1}{\mathrm{d}\lambda}} \right|^2 \mathrm{d}\lambda$$

where $\mathcal{A}$ is an admissible set for a scalar field on positive measures; λ is any reference measure such that ρ and ρ' are both absolutely continuous with respect to λ, with Radon-Nikodym derivatives $\frac{\mathrm{d}\rho_i}{\mathrm{d}\lambda}$.

The equivalence between the square Fisher–Rao distance and squared Hellinger distance (Liero et al., 2016; 2018) quantifies the similarity between two probability distributions ranging from 0 to 1. The total variation bounds the squared form and is well-studied in the information geometry (Amari, 2016). The partial differential equations of the form $\partial_t\rho_t = \alpha_t\rho_t$ are called reaction equations of α_t, which describes dynamics regarding concentration.

Wasserstein-Fisher-Rao. The Wasserstein-Fisher-Rao geometry, or equivalently, spherical Hellinger–Kantorovich distance, considers liftings of positive, complete, and separable measures while preserving the total mass. This can be expresses as combining the Fisher–Rao and Wasserstein geometries characterized by PDE such as (Liero et al., 2016):

$$\partial_t\rho_t + \nabla \cdot (v_t\rho_t) = \frac{\omega_t}{2}\rho_t. \tag{64}$$

One problem, is that the PDE (64) In order to stay the dynamics on the space of probability measures, which is our interest, we adopt the definition from (Lu et al., 2019; Lambert et al., 2022) the equation becomes

$$\partial_t \rho_t + \nabla \cdot (\rho_t v_t) = \frac{1}{2}\left(\beta_t - \int \beta_t \mathrm{d}\rho_t\right)\rho_t, \tag{65}$$

which satisfies mass conservation. For the geometry, the norm on tangent space is given by

$$\|(\beta_t, \rho)\|_\rho^2 \coloneqq \int \left\{ \left(\omega - \int \beta_t \, \mathrm{d}\rho\right)^2 + \|v\|^2 \right\} \mathrm{d}\rho. \tag{66}$$

and we define the WFR distance as

$$d_{\mathtt{WFR}}^2(\rho_0, \rho_1) \coloneqq \inf_{\rho, \beta_t, v} \left\{ \int_0^1 \|(\beta_t, v_t)\|_{\rho_t}^2 \mathrm{d}t \;\middle|\; \{\rho_t, \beta_t, v_t\}_{t\in[0,1]} \text{ satisfies (65)} \right\}. \tag{67}$$

(Lu et al., 2019) demonstrated that WFR gradient dynamics over the Bures–Wasserstein space can be analytically derived. In this work, we were able to design a computational method for OMD iterates in the WFR geometry. Using Proposition 3, this geometry allowed the VMSB algorithm to perform tractable gradient computation within Wasserstein space.

B.3 The Bures–Wasserstein space and a mixture of Gaussians

The space of Gaussian distribution in the Wasserstein space is known as Bures–Wasserstein space, denoted as $\mathtt{BW}(\mathbb{R}^d)$. Given $\theta_0, \theta_1 \in \mathtt{BW}(\mathbb{R}^d)$, we can identify the space with the manifold $\mathbb{R}^d \times \mathbf{S}_{++}^d$, where $\mathbf{S}_{++}^d$ denotes the space of symmetric positive definite matrices. For $\theta_0 = (m_0, \Sigma_0)$ and $\theta_1 = (m_1, \Sigma_1)$ an affine map from p_{θ_0} to p_{θ_1} is given as a closed-form expression:

$$\nabla\zeta(x) = m_1 + \Sigma_0^{-1/2}\big(\Sigma_0^{1/2}\Sigma_1\Sigma_0^{1/2}\big)^{1/2}\Sigma^{-1/2}(x - m_0).$$

Note that the constant-speed geodesic also lies in $\mathtt{BW}(\mathbb{R}^d)$, as pushforward of a Gaussian with an affine map is also a Gaussian. Therefore, it can be said that $\mathtt{BW}(\mathbb{R}^d)$ is a geodesically convex subset of $\mathcal{P}_2(\mathbb{R}^d)$. For the Brenier map, a constant-speed geodesic in $\mathtt{BW}(\mathbb{R}^d)$, for the tangent vector to the geodesic (r, S)

$$p_{\theta_t} = \exp_{p_{\theta_0}}\big(t \cdot (r, S)\big) = \mathcal{N}\big(m_0 + tr, (tS + I_d)\Sigma_0(tS + I_d)\big), \tag{68}$$

and the dynamics at its current position at time $t = 0$ is represented as

$$\dot{m}_0 = r, \tag{69}$$

$$\dot{\Sigma}_0 = S\Sigma_0 + \Sigma_0 S. \tag{70}$$

Generalizing this geodesic dynamics, the Bures–Wasserstein gradient $\nabla_{\mathtt{BW}} f$ of a function $f : \mathbb{R}^d \times \mathbf{S}_{++}^d \to \mathbb{R}$ for a tangent vector (r, S) at time 0 Altschuler et al. (2021)

$$\big\langle \nabla_{\mathtt{BW}} f(m_0, \Sigma_0), (r, S) \big\rangle_{\mathtt{BW}} = \partial_t f(m_t, \Sigma_t)\Big|_{t=0}$$

Identifying each component, we achieve the following result of Wasserstein gradient flow in Bures–Wasserstein space as

$$\nabla_{\mathtt{BW}} f = (\nabla_m f, 2\nabla_\Sigma f), \tag{71}$$

where ∇_m and ∇_Σ denote Euclidean gradient. We refer to Appendix A of Altschuler et al. (2021) and Appendix B Lambert et al. (2022) for further geometric properties and discussion for the BW space.

Table 5: Hyperparameters.

	2D	**EOT**	**MSCI**	**MNIST (Pixel)**	**MNIST (Latent)**	**FFHQ**
Dimension d	2	$\{2, 16, 64, 128\}$	$\{50, 100, 1000\}$	784	128	512
Modality K	$\{8, 20, 50\}$	$[5, 100]$	50	{256, 1024, 4096}	{256, 1024}	10
Volatility ε	0.1	$\{0.1, 1, 10\}$	0.1	10^{-4}	10^{-3}	$\{0.1, 0.5, 1.0, 10.0\}$
Total steps (τ)	20,000	30,000	10,000	100,000	30,000	20,000
OMD steps (t)	400	600	200	1000	375	400

C Experimental Details

C.1 Rationales of the GMM parameterization for VMSB

Our parameterization choice follows LightSB (Korotin et al., 2024) because of the following two key reasons. First, GMMs ensure that the model space satisfies certain measure concentration, which is suitable for analyzing theoretical properties of SB models (Conforti et al., 2023). Firstly, we analyzed the regret under the log Sobolev inequality in Proposition 2. Enforcing the LightSB parameterization will automatically satisfy Assumption 1. Secondly, VMSB requires tractable gradient computation of Wasserstein gradient flow in § 4.3. As shown in Proposition 3, we can perform VMSB using the variational inference in the WFR geometry of the GMM parameterization.

C.2 Hyperparameters.

The hyperparameters are displayed in Table 5. For step size scheduling, we followed the theoretical result in Theorem 1 and Proposition 1, and chose $\eta_1 = 1$ and $\eta_T \in \{0.05, 0.1\}$ with harmonic sequences, as illustrated in Fig. 5. For high dimensional tasks in MSCI (1000d), MNIST-EMNIST (784d), and latent FFHQ Image-to-Image transfer tasks (512d), the initial *warm up* steps for 10% of the total learning helped starting a training sequence from a reasonable starting point as this set $\eta_t = 1$ as verified in Fig. 6 (c).

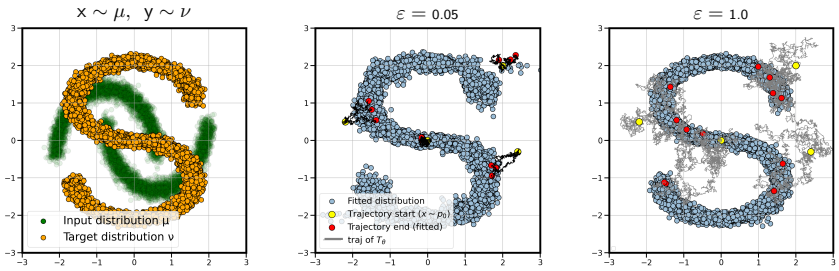

Figure 10: SB in 2D synthetic datasets. SB processes $\mathcal{T}_\theta$ with different volatility ε.

C.3 2D Synthetic datasets and the online learning setup

Fig. 10 demonstrates that our method achieved the SB model for the various volatility ε. For various configurations, most of baseline SB algorithms are capable of learning in the 2D space (10). In order to align our theoretical arguments for online learning, we selectively offered with a rotating filter that only 12.5% of the samples to the SB solvers based on the angles measured from the origin. For instance, we provided data for angle of $[0, \pi/4]$ for first $t \in [0, 25)$ steps, and so on. This partial observability is periodically rotated through the data stream, thereby testing the algorithm's ability to learn robustly under sparse and shifting information. Since this requires 200 batches for the full rotation of the filter, the problem became substantially more challenging, and LightSB and LightSB-M algorithms oftentimes failed on this online learning setting.

C.4 Entropic optimal transport benchmark

Our hyperparameter for the EOT benchmarks choices mostly follow the official repositories of the LightSB[4] and LightSB-M[5]. Since it is known that initial distribution μ is the standard Gaussian distribution (Gushchin et al., 2024b), we only trained v_θ using the variational MD algorithm. Due to the huge number of configurations, some hyperparameter settings were not clearly reported. Thus, we conducted our own examination on these cases; we replicated better performance than the reported numbers by carefully dealing each benchmark configuration.

C.5 SB learning with adversarial networks

Suppose a discriminator network, denoted as D, is equipped with useful architectural properties for discriminating images. The discriminator outputs a binary classification regarding authenticity through sigmoidal outputs, *i.e.*, $D(x) \in [0,1]\ \forall x \in \mathbb{R}^{28\times 28\times 1}$. For image samples $\mathbf{x} = \{x^1, \dots, x^N\} \sim \mu$, we trained the discriminator D with the logistic regression:

$$\underset{D}{\text{maximize}}\ \frac{1}{N}\sum_{n=1}^{N} \log D(y^n) + \frac{1}{B}\sum_{m=1}^{M} \log(1 - D(\hat{y}_\phi^m)), \tag{72}$$

where $\hat{y}_\phi^m$ in the right-hand side denotes a sample from an SB model parameterized by ϕ, generated using an input x^m. From our experiment setting, we use the SB distribution ρ_ϕ which is generated by $\vec{\pi}_\phi$ from samples of the marginal μ. This makes the objective of adversarial learning of training the law of SB process at time $t = 1$. For a completely separable metric space, it is well known that the discriminator converges at $D(x) = \frac{\nu(x)}{\nu(x)+\rho_\phi(x)}$ (Goodfellow et al., 2014).

In the adversarial learning technique, retaining a fully differentiable computation path from the input pixels to the discriminator outputs is essential. Therefore, we implemented a differentiable inference function using the categorical reparameterization trick with Gumbel-softmax (Jang et al., 2016), as well as the Gaussian reparameterization trick. These reparameterization tricks enabled learning with samples generated through LightSB-adv-K, directly by maximizing

$$\tilde{\mathcal{J}}(\phi) = \frac{1}{M}\sum_{m=1}^{M} \log D(y_\phi^m) - \log(1 - D(y_\phi^m)),$$

Table 6: A simple discriminator D.

Layer Type	Shape
Input Layer	(-1, 28, 28, 1)
Conv Layer 1	(-1, 14, 14, 64)
Conv Layer 2	(-1, 7, 7, 128)
Batch Norm	(-1, 7, 7, 128)
Flatten	(-1, 6272)
Dense	(-1, 1024)
Dense	(-1, 1)

where the term essentially represents the *logit* function $\text{logit}(D(y)) = \log \frac{D(y)}{1-D(y)}$. When D approaches the equilibrium, we can approximate the following KL learning

$$\tilde{\mathcal{J}}(\phi) \approx \int \log \frac{\nu(y)}{\rho_\phi(y)} \rho_\phi(y) \mathrm{d}y = \mathrm{KL}(\rho_\phi \| \nu),$$

where the KL functional directly corresponds to the divergence minimization of the SB problems (4) and (24), under the disintegration theorem of Schrödinger bridge (Léonard, 2014).

In the MNIST-EMNIST image transfer tasks, we set one of the baseline as the aforementioned adversarial learning as the baseline for training the SB model for the pixel space. Among our attempts, while the LightSB-adv method successfully generated learning signals to train GMM-based models, the losses proposed by LightSB (Korotin et al., 2024) and LightSB-M (Gushchin et al., 2024a) failed to generate relevant images with high fidelity. For the discriminator, we used the DCGAN (Radford, 2015) architecture shown in Table 6, and this can be replaced with more complex architecture for more realistic images with high fidelity. We fixed the covariance after warm-ups in 10,000 steps, and we used the entropy coefficient $\varepsilon = 10^{-4}$ based on our hyperparameter search.

[4]https://github.com/ngushchin/LightSB
[5]https://github.com/SKholkin/LightSB-Matching

C.6 Latent diffusion experiments

For the latent space, we pretrained ALAE (Pidhorskyi et al., 2020) model using the both MNIST and EMNIST (first ten letters) datasets. The ALAE is a high-fidelity autoencoder internally use an adversarial learning to generate high-fidelity images. For the encoder network, as well as decoder network, we mostly adopt the DCGAN architecture. Therefore, the encoder is mostly identical to Table 6 except the point the final layer is 128 dimension instead of 1, and the decoder is a convolutional neural network with four convolutional layers.

Following the latent SB setting (Korotin et al., 2024), we assessed our method by utilizing the ALAE model (Pidhorskyi et al., 2020) for generating 1024×1024 images of the FFHQ dataset (Karras et al., 2019). The base generative model has a latent embedding layer which represent 512-dimensional embedding space. The goal is to transport a point latent space to another, performing unpaired image-to-image translation tasks for four distinct cases: *Adult*$\rightarrow$ *Child*, *Child*$\rightarrow$ *Adult*, *Female*$\rightarrow$ *Male*, and *Male*$\rightarrow$ *Female*. We conducted a quantitative analysis using the ED on the predefined ALAE embedding as a metric for evaluation.

D Discussion on Implementation of VMSB

Limitations. GMM-based SB models, due to the lack of deep structural processing, tend to focus on *instance-level* associations of images in EOT couplings rather than the *subinstance-* or *feature-level* associations that are intrinsic to deep generative models. As a result, while VMSB produces statistically valid representations of optimal transportation within the given architectural constraints, these outcomes may be perceived as somewhat "synthetic." Nevertheless, GMM-based models still hold an irreplaceable role in numerous problems such as latent diffusion and variational methods, due to their simplicity and distinctive properties (Korotin et al., 2024). As we successfully demonstrated in two distinct ways of interacting with neural networks for solving unpaired image transfer, we hope our theoretical and empirical findings help novel neural architecture studies.

Computation. For fast computation, we utilized the JAX automatic differentiation library (Bradbury et al., 2018) for computing gradients and Hessians in Proposition 3. For each input, the computational of VMSB requires quadratic time for computing the Wasserstein gradient flow (asymptotically $\mathcal{O}(K^2 n_y)$) and the memory footprint for estimating with internal Gaussian particles is linear (asymptotically $\mathcal{O}(K n_y)$). There are inherent trade-offs between accuracy and computational efficiency when choosing between LightSB and VMSB; nevertheless, VMSB remains significantly more manageable and computationally tractable compared to deep learning methods for moderate settings. For instance, we have presented performance regarding efficiency and scalability up to 1,000 dimensions in the experiments. Driven by parallel nature of Gaussian particles, we observed that the computation of Proposition 3 favors vectorized instructions, and the expected speed enhancement from using GPUs is much more evident in neural network cases. In Table 7, we report the wall-clock time for a 100-dimensional single-cell data problem Vargas et al. (2021); Korotin et al. (2024), where the performance is reported in Table 3. Additionally, training time in the MNIST-EMNIST translation is reported in Table 11 in the ablation study. This property also holds for generation, allowing practitioners to deploy the model much faster on GPUs. In Table 8, we also report that generating 100 MNIST samples from 4096 Gaussian particles, equipped with competitive performance, can be done 1,854 times faster under the same hardware. Since VMSB a simulation-free, the GMM generation process does not suffer from discretization errors of SDE.

Table 7: Training time for the 100-dimension single-cell data problem.

Sinkhorn (IPF)	LightSB	VMSB
8m (GPU)	66s (CPU)	32s (GPU) / 22m (CPU)

Reproducibility statement. Comprehensive justification and theoretical background are presented in Appendices A and B. Since the primary contributions of this paper pertain to the learning methodology, we ensured that all architectures and hyperparameters remained consistent across the LightSB variants.

Table 8: Generation time for the 784-dimension MNIST pixel data.

	$K = 64$	$K = 256$	$K = 1024$	$K = 4096$	NN (SDE)
GPU	721μs	726μs	739μs	740μs	1.372s
CPU	60.140ms	133.333ms	428.433ms	1.527s	–

All datasets utilized in this study are available for download alongside the training scripts. Please refer to Appendix C for more information on the experimental setups.

Table 9: EOT Benchmark scores of $\mathrm{B}\mathbb{W}_2^2$-UVP $\downarrow$ (%).

Type	Solver	$\varepsilon = 0.1$				$\varepsilon = 1$				$\varepsilon = 10$			
		$d = 2$	$d = 16$	$d = 64$	$d = 128$	$d = 2$	$d = 16$	$d = 64$	$d = 128$	$d = 2$	$d = 16$	$d = 64$	$d = 128$
Classical solvers (best)†		0.016	0.05	0.25	0.22	0.005	0.09	0.56	0.12	0.01	0.02	0.15	0.23
Bridge-M	DSBM (Shi et al.)‡	0.03	0.18	0.7	2.26	0.04	0.09	1.9	7.3	0.26	102	3563	15000
Bridge-M	SF^{2}M-Sink (Tong et al.)‡	0.04	0.18	0.39	1.1	0.07	0.3	4.5	17.7	0.17	4.7	316	812
rev. KL	LightSB (Korotin et al.)	0.004 ± 0.004	0.009 ± 0.004	0.023 ± 0.003	0.036 ± 0.003	0.004 ± 0.005	0.009 ± 0.003	0.016 ± 0.002	0.035 ± 0.003	0.009 ± 0.004	0.013 ± 0.007	0.034 ± 0.004	0.066 ± 0.008
Bridge-M	LightSB-M (Gushchin et al.)	0.005 ± 0.003	0.012 ± 0.004	0.034 ± 0.003	0.063 ± 0.002	0.005 ± 0.001	0.027 ± 0.007	0.057 ± 0.010	0.108 ± 0.004	0.004 ± 0.002	0.017 ± 0.007	0.133 ± 0.010	0.409 ± 0.042
EMA	LightSB-EMA	0.004 ± 0.002	0.014 ± 0.003	0.021 ± 0.003	0.044 ± 0.001	0.004 ± 0.003	0.009 ± 0.004	0.013 ± 0.001	0.032 ± 0.004	0.004 ± 0.001	0.008 ± 0.003	0.023 ± 0.013	0.010 ± 0.002
Var-MD	VMSB (ours)	$\mathbf{0.003 \pm 0.001}$	$\mathbf{0.007 \pm 0.003}$	$\mathbf{0.018 \pm 0.002}$	$\mathbf{0.039 \pm 0.001}$	$\mathbf{0.002 \pm 0.002}$	$\mathbf{0.004 \pm 0.001}$	$\mathbf{0.009 \pm 0.001}$	$\mathbf{0.023 \pm 0.003}$	$\mathbf{0.005 \pm 0.007}$	$\mathbf{0.006 \pm 0.004}$	$\mathbf{0.011 \pm 0.010}$	$\mathbf{0.011 \pm 0.004}$
Var-MD	VMSB-M (ours)	$\mathbf{0.002 \pm 0.001}$	$\mathbf{0.010 \pm 0.067}$	$\mathbf{0.031 \pm 0.004}$	$\mathbf{0.056 \pm 0.005}$	$\mathbf{0.003 \pm 0.004}$	$\mathbf{0.005 \pm 0.002}$	$\mathbf{0.032 \pm 0.006}$	$\mathbf{0.077 \pm 0.018}$	$\mathbf{0.003 \pm 0.003}$	$\mathbf{0.011 \pm 0.004}$	$\mathbf{0.117 \pm 0.012}$	0.429 ± 0.748

Table 10: EOT scores of $\mathrm{cB}\mathbb{W}_2^2$-UVP, the fully extended version of Table 2.

Type	Solver	$\varepsilon = 0.1$				$\varepsilon = 1$				$\varepsilon = 10$			
		$d = 2$	$d = 16$	$d = 64$	$d = 128$	$d = 2$	$d = 16$	$d = 64$	$d = 128$	$d = 2$	$d = 16$	$d = 64$	$d = 128$
Classical solvers (best)†		1.94	13.67	11.74	11.4	1.04	9.08	18.05	15.23	1.40	1.27	2.36	1.31
Bridge-M	DSBM (Shi et al.)‡	5.2	10.8	37.3	35	0.3	1.1	9.7	31	3.7	105	3557	15000
Bridge-M	SF^{2}M-Sink (Tong et al.)‡	0.54	3.7	9.5	10.9	0.2	1.1	9	23	0.31	4.9	319	819
rev. KL	LightSB (Korotin et al.)	0.007 ± 0.005	0.040 ± 0.023	0.100 ± 0.013	0.140 ± 0.003	0.014 ± 0.003	0.026 ± 0.002	0.060 ± 0.004	0.140 ± 0.003	0.019 ± 0.005	0.027 ± 0.005	0.052 ± 0.002	0.092 ± 0.001
Bridge-M	LightSB-M (Gushchin et al.)	0.017 ± 0.004	0.088 ± 0.014	0.204 ± 0.036	0.346 ± 0.036	0.020 ± 0.007	0.069 ± 0.016	0.134 ± 0.014	0.294 ± 0.017	0.014 ± 0.001	0.029 ± 0.004	0.207 ± 0.005	0.747 ± 0.028
EMA	LightSB-EMA	0.005 ± 0.002	0.040 ± 0.014	0.078 ± 0.007	0.149 ± 0.006	0.012 ± 0.002	0.022 ± 0.003	0.051 ± 0.001	0.127 ± 0.002	0.017 ± 0.003	0.021 ± 0.003	0.025 ± 0.002	0.042 ± 0.002
Var-MD	VMSB (ours)	$\mathbf{0.004 \pm 0.001}$	$\mathbf{0.012 \pm 0.002}$	$\mathbf{0.038 \pm 0.002}$	$\mathbf{0.101 \pm 0.002}$	$\mathbf{0.010 \pm 0.001}$	$\mathbf{0.018 \pm 0.001}$	$\mathbf{0.044 \pm 0.001}$	$\mathbf{0.114 \pm 0.001}$	$\mathbf{0.013 \pm 0.001}$	$\mathbf{0.019 \pm 0.001}$	$\mathbf{0.021 \pm 0.008}$	$\mathbf{0.040 \pm 0.001}$
Var-MD	VMSB-M (ours)	$\mathbf{0.015 \pm 0.016}$	$\mathbf{0.067 \pm 0.036}$	$\mathbf{0.108 \pm 0.020}$	$\mathbf{0.253 \pm 0.107}$	$\mathbf{0.010 \pm 0.001}$	$\mathbf{0.019 \pm 0.001}$	$\mathbf{0.094 \pm 0.010}$	$\mathbf{0.222 \pm 0.033}$	$\mathbf{0.013 \pm 0.001}$	$\mathbf{0.029 \pm 0.003}$	$\mathbf{0.193 \pm 0.015}$	0.748 ± 0.036

E Additional Experimental Results

E.1 Additional results on the EOT benchmark

We present the full results of EOT benchmark experiments. Tables 9 and 10 show comprehensive statistics on the EOT benchmark with more SB solvers. As mentioned in § 6.2, the VMSB and VMSB-M solvers consistently brought better performance with low standard deviations of scores for $\mathrm{cB}\mathbb{W}_2^2$-UVP and $\mathrm{B}\mathbb{W}_2^2$-UVP measures. We note that the experiment was conducted in a highly controlled setting with identical model configurations; with all other aspects controlled and outcomes differing only by learning methods, the consistent performance gains of our work were a well-anticipated result from our theoretical analysis.

E.2 Additional image generation results

In the unpaired EMNIST-to-MNIST translation task for the raw 784 pixel, we measured FID scores for various K for the SB parameterization. We considered $K \in \{64, 256, 1024, 4096\}$ with $\varepsilon = 10^{-4}$ for our VMSB algorithm. Our observations, both qualitative and quantitative, indicate that higher modalities yield higher-quality samples. In every case of K, VMSB-adv outperformed its counterpart. For instance, Fig. 12 demonstrates that VMSB generates more diverse samples with high fidelity. Notably, we achieved the competitive FID score of 15.471 using a standard neural network discriminator with relatively low MSD similarity scores. As the latent VMSB model for 128-dimensional embeddings also achieved the considerably low FID score of 9.558 (Table 4), we concluded that VMSB showed promising quality improvements for the both case, and this strongly supports the generality of our theory.

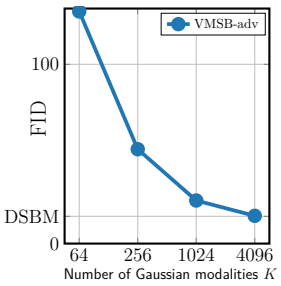

Figure 12: FID vs. modality.

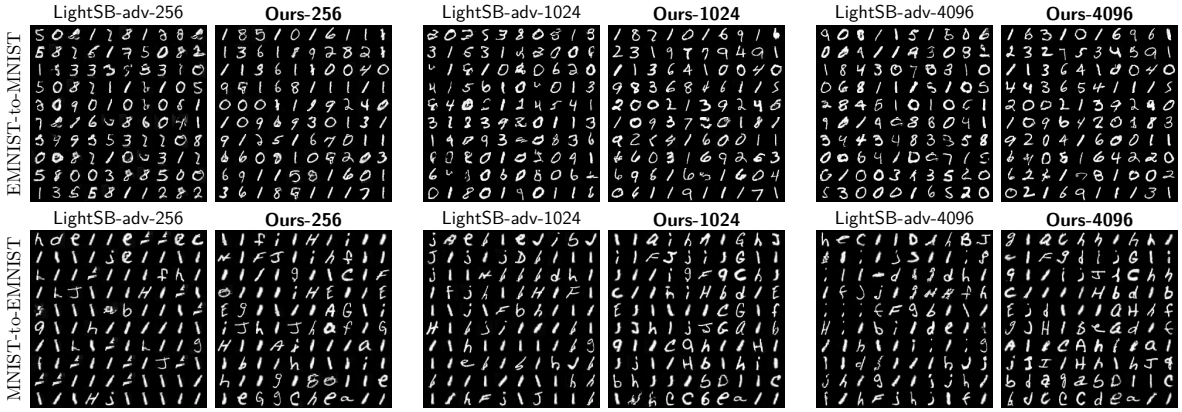

Figure 11: Generation results for unpaired image-to-image translation in the raw pixel space. We considered image data from MNIST and EMNIST (containing the first ten letters), sized as 28×28 pixels. For comparison, we trained GMM-based models with adversarial learning using a simple logistic discriminator (Table 6). This was used as both a benchmark and a tractable target SB model (LightSB-adv-K). Our method in the raw pixel domain, denoted as Ours-K, demonstrated qualitative improvements in terms of diversity and clarity of image samples.

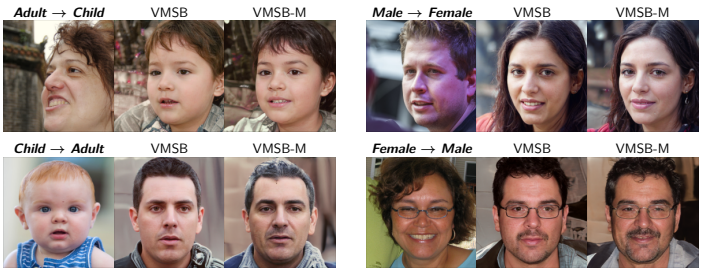

Figure 13: Image-to-Image translation on a latent space for the VMSB and VMSB-M algorithms.

Fig. 11 demonstrates that VMSB generated more diverse samples with high fidelity. Note that the proposed method suffers less from mode collapse than LightSB method (especially on the transfer MNIST-to-EMNIST), with the same Gaussian mixture setting. This result is especially a good point where the difference only lies in the learning methodology, which aligns with our theory. Tables 11 and 12 effectively show the statistics and FID scores on both the train and the test datasets. The quantitative results highlight that the VMSB solver is more performant with less overfitting than its counterpart. Consequently, our claim regarding the stability of SB solution acquisition is verified by additional experiments involving pixel spaces.

We present Embedding-ED scores (Jayasumana et al., 2024) and some qualitative generation results in Table 13, which is visualized in Fig. 9. SF^{2}M-Sink For quantitative results, we calculated statistics from ED scores on embeddings of the ALAE model (Pidhorskyi et al., 2020), for the four different unpaired image-to-

Table 11: MNIST transfer statistics.

	FID	Time	Parameters
LightSB-256	61.257	30m	0.4M
LightSB-1024	26.487	53m	1.6M
LightSB-4096	20.017	135m	6.4M
VMSB-256	52.634	76m	0.4M
VMSB-1024	24.022	203m	1.6M
VMSB-4096	15.471	44h	6.4M
DSBM-IMF	11.429	42h	6.6M

Table 12: FID scores and differences for generated MNIST.

	FID (Train)	FID (Test)	Diff. (test − train).
LightSB-adv-256	60.746	61.604	0.858
LightSB-adv-1024	25.934	26.569	0.635
LightSB-adv-4096	19.960	20.196	0.237
VMSB-adv-256	51.684	52.283	0.599
VMSB-adv-1024	23.853	24.053	0.200
VMSB-adv-4096	15.508	15.496	−0.012

image translation tasks. The results show that VMSB is capable of translating an arbitrary representation, which is closer to target domain than baselines. In Fig. 13, as well as Fig. 9, we can see that VMSB and VMSB-M algorithms generate FFHQ data with a given translation task. To qualitatively verify these generation results, we generated images using LightSB and VMSB in Figures 14 and 15. Since these improvements are purely based on information geometry and learning theory, we anticipate that following works on the variational principle application across various fields such as image processing, natural language processing, and control systems (Caron et al., 2020; Liu et al., 2023; Alvarez-Melis & Jaakkola, 2018; Chen et al., 2022).

Table 13: ALAE Embedding-ED scores. To evaluate the performance, we computed averages and standard deviations of the ED scores across four different transfer tasks.

	$\varepsilon = 0.1$	$\varepsilon = 0.5$	$\varepsilon = 1.0$	$\varepsilon = 10.0$
SF^{2}M-Sink	0.02916 ± 0.00145	0.04112 ± 0.00191	0.05670 ± 0.00249	0.06641 ± 0.00441
DSBM-IMF	0.02275 ± 0.00101	0.03358 ± 0.00142	0.04866 ± 0.00168	0.06474 ± 0.00381
LightSB	0.01086 ± 0.00045	0.02382 ± 0.00093	0.03462 ± 0.00148	0.05376 ± 0.00273
LightSB-M	0.01066 ± 0.00055	0.02366 ± 0.00107	0.03519 ± 0.00153	0.05975 ± 0.00298
VMSB	$\mathbf{0.01002 \pm 0.00055}$	$\mathbf{0.02288 \pm 0.00101}$	$\mathbf{0.03396 \pm 0.00174}$	$\mathbf{0.05315 \pm 0.00307}$
VMSB-M	$\mathbf{0.00997 \pm 0.00054}$	$\mathbf{0.02298 \pm 0.00106}$	$\mathbf{0.03391 \pm 0.00140}$	$\mathbf{0.05351 \pm 0.00241}$

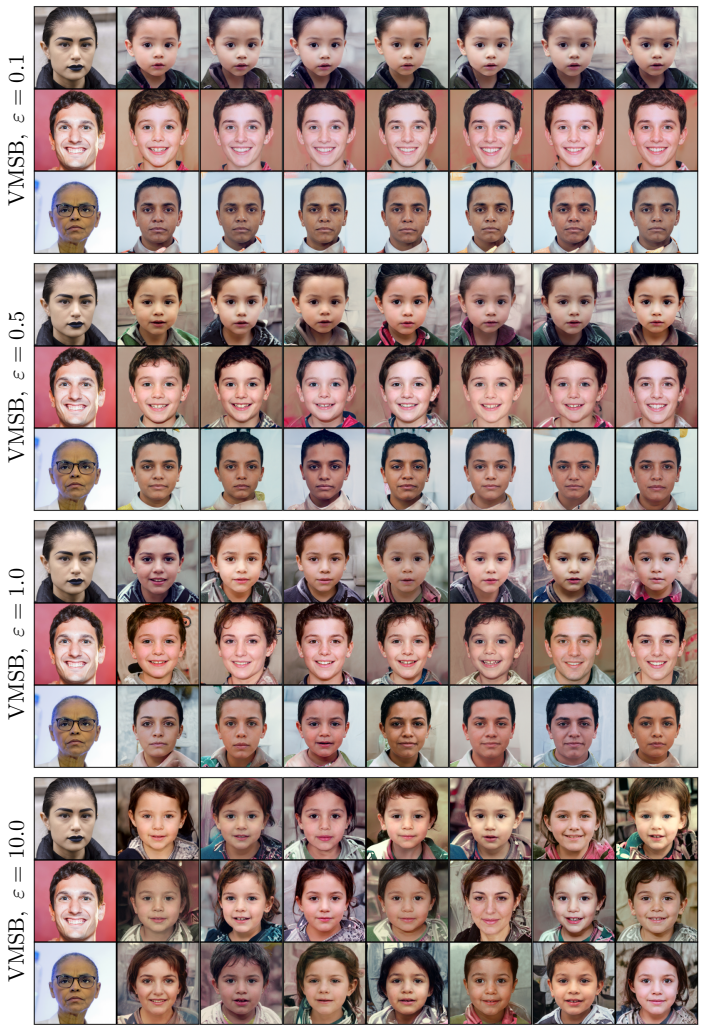

Figure 14: Generation results of VMSB (*Adult* $\rightarrow$ *Child*) with different volatility settings

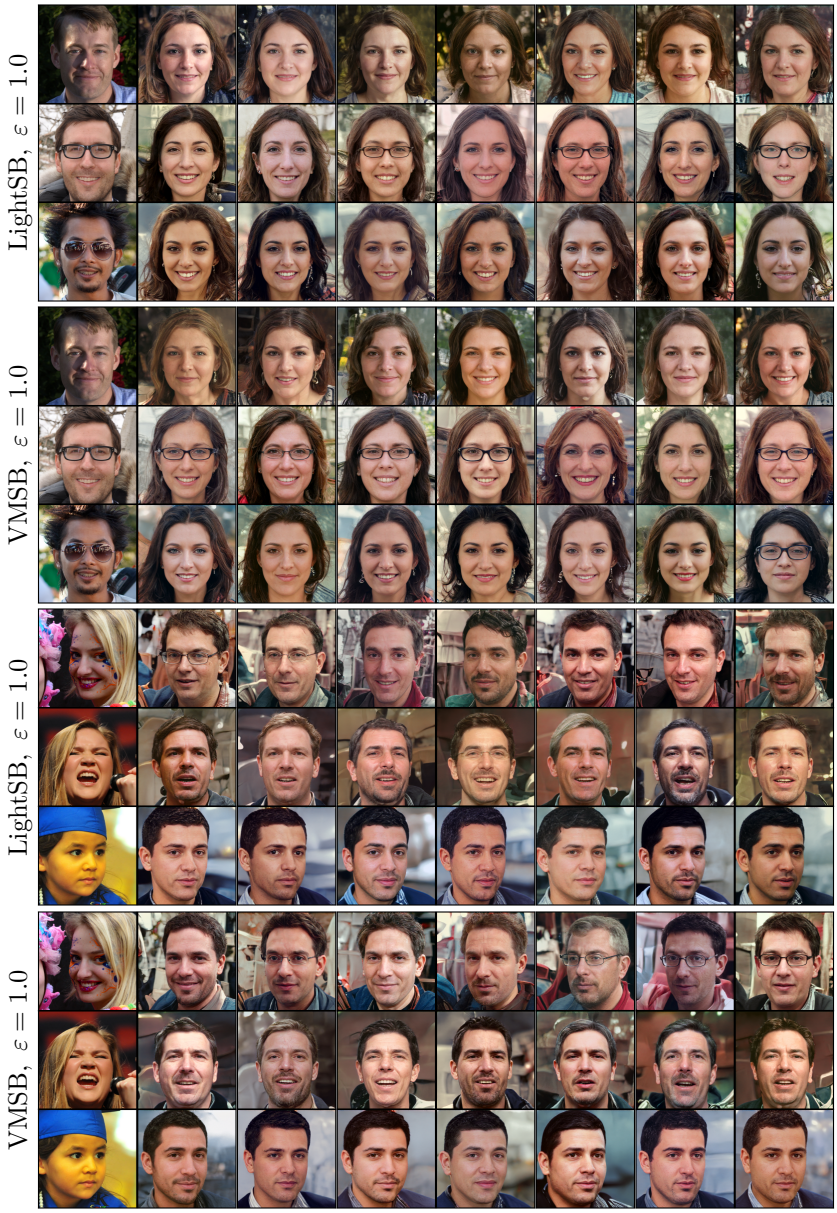

Figure 15: Qualitative comparison between LightSB and VMSB for relatively high volatility, $\varepsilon = 1.0$. Top (*Male* $\rightarrow$ *Female*): We find that VSBM has preserved more facial details, such as wearing glasses, than LightSB. Bottom (*Adult* $\rightarrow$ *Child*): VSBM was stable at retaining facial position even with high ε.