Action Matching: Learning Stochastic Dynamics from Samples

Kirill Neklyudov [1] **Rob Brekelmans** [1] **Daniel Severo** [1 2] **Alireza Makhzani** [1 2]

Abstract

Learning the continuous dynamics of a system from snapshots of its temporal marginals is a problem which appears throughout natural sciences and machine learning, including in quantum systems, single-cell biological data, and generative modeling. In these settings, we assume access to cross-sectional samples that are uncorrelated over time, rather than full trajectories of samples. In order to better understand the systems under observation, we would like to learn a model of the underlying process that allows us to propagate samples in time and thereby simulate entire individual trajectories. In this work, we propose Action Matching, a method for learning a rich family of dynamics using only independent samples from its time evolution. We derive a tractable training objective, which does not rely on explicit assumptions about the underlying dynamics and does not require back-propagation through differential equations or optimal transport solvers. Inspired by connections with optimal transport, we derive extensions of Action Matching to learn stochastic differential equations and dynamics involving creation and destruction of probability mass. Finally, we showcase applications of Action Matching by achieving competitive performance in a diverse set of experiments from biology, physics, and generative modeling.

1. Introduction

Understanding the time evolution of systems of particles or individuals is a fundamental problem appearing across machine learning and the natural sciences. In many scenarios, it is expensive or even physically impossible to observe entire individual trajectories. For example, in quantum mechanics, the act of measurement at a given point collapses the wave function (Griffiths & Schroeter, 2018), while in biological applications, single-cell RNA- or ATAC- sequencing techniques destroy the cell in question (Macosko et al., 2015; Klein et al., 2015; Buenrostro et al., 2015).

Instead, from 'cross-sectional' or independent samples at various points in time, we would like to learn a model which simulates particles such that their density matches that of the observed samples. The problem of learning stochastic dynamics from marginal samples is variously referred to as learning *population dynamics* (Hashimoto et al., 2016) or as *trajectory inference* (Lavenant et al., 2021), in contrast to time series modeling where entire trajectories are assumed to be available. Learning such models to predict entire trajectories holds the promise of facilitating simulation of complex chemical or physical systems (Vázquez, 2007; Noé et al., 2020) and understanding developmental processes or treatment effects in biology (Schiebinger et al., 2019; Tong et al., 2020; Schiebinger, 2021; Bunne et al., 2021).

Furthermore, recent advances in generative modeling have been built upon learning stochastic dynamics which interpolate between the data distribution and a prior distribution. In particular, score-based diffusion models (Song et al., 2020b; Ho et al., 2020) construct a stochastic differential equation (SDE) to move samples from the data distribution to a prior distribution, while score matching (Hyvärinen & Dayan, 2005) is used to learn a reverse SDE which models the gradients of intermediate distributions. However, these methods rely on analytical forms of the SDEs and/or the tractability of intermediate Gaussian distributions (Lipman et al., 2022). Since our proposed method can learn dynamics which simulate an arbitrary path of marginal distributions, it can also be applied in the context of generative modeling. Namely, we can approach generative modeling by constructing an interpolating path between the data and an arbitrary prior distribution, and learning to model the resulting dynamics.

In this work, we propose *Action Matching*, a method for learning population dynamics from samples of their temporal marginals q_t. Our contributions are as follows:

- In Theorem 2.1, we establish the existence of a unique gradient field ∇s_t^* which traces any given time-

[1]Vector Institute [2]University of Toronto. Correspondence to: <k.necludov@gmail.com>, <makhzani@vectorinstitute.ai>.

Proceedings of the 40th International Conference on Machine Learning, Honolulu, Hawaii, USA. PMLR 202, 2023. Copyright 2023 by the author(s).

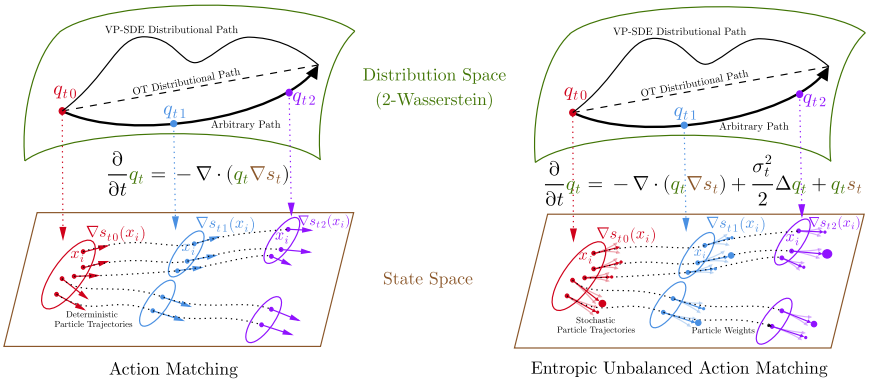

Figure 1. Action Matching, and its entropic (eAM) and unbalanced (uAM) variants, can learn to trace any arbitrary distributional path. For a given path, AM learns deterministic trajectories, eAM learns stochastic trajectories, and uAM learns weighted trajectories.

continuous distributional path q_t. Notably, our restriction to gradient fields is without loss of expressivity for this class of q_t. To learn this gradient field, we propose the tractable Action Matching training objective in Theorem 2.2.

- In Sec. 3.1-3.3, we extend the above approach in several ways: an 'entropic' version which can approximate ground-truth dynamics involving stochasticity, an 'unbalanced' version which allows for creation and destruction of probability mass, and a version which can minimize an arbitrary convex cost function in the Action Matching objective.
- We discuss the close relationship between Action Matching and dynamical optimal transport, along with other related works in Sec. 5 and App. B.
- Since Action Matching relies only on samples and does not require tractable intermediate densities or knowledge of the underlying stochastic dynamics, it is applicable in a wide variety of problem settings. In particular, we demonstrate competitive performance of Action Matching in a number of experiments, including trajectory inference in biological data (Sec. 4.1), evolution of quantum systems (Sec. 4.2), and a variety of tasks in generative modeling (Sec. 4.3).

2. Action Matching

2.1. Continuity Equation

Suppose we have a set of particles in space $\mathcal{X} \subset \mathbb{R}^d$, initially distributed as $q_{t=0}$. Let each particle follow a time-dependent ODE (continuous flow) with the velocity field $v : [0,1] \times \mathcal{X} \to \mathbb{R}^d$ as follows

$$\frac{d}{dt}x(t) = v_t(x(t))\,, \quad x(t=0) = x\,. \tag{1}$$

The continuity equation describes how the density of the particles q_t evolves in time t, i.e.,

$$\frac{\partial}{\partial t}q_t = -\nabla \cdot (q_t v_t)\,, \tag{2}$$

which holds in the distributional sense, where $\nabla\cdot$ denotes the divergence operator. Under mild conditions, the following theorem shows that *any continuous dynamics can be modeled by the continuity equation*, and moreover any continuity equation results in a continuous dynamics.

Theorem 2.1 (Adapted from Theorem 8.3.1 of Ambrosio et al. (2008))**.** *Consider a continuous dynamic with the density evolution of q_t, which satisfies mild conditions (absolute continuity in the 2-Wasserstein space of distributions $\mathcal{P}_2(\mathcal{X})$). Then, there exists a unique (up to a constant) function $s_t^*(x)$, called the "action", such that vector field $v_t^*(x) = \nabla s_t^*(x)$ and q_t satisfies the continuity equation*

$$\frac{\partial}{\partial t}q_t = -\nabla \cdot (q_t \nabla s_t^*(x))\,. \tag{3}$$

In other words, the ODE $\frac{d}{dt}x(t) = \nabla s_t^*(x)$ can be used to move samples in time such that the marginals are q_t.

Using Theorem 2.1, the problem of learning the dynamics can be boiled down to learning the unique vector field ∇s_t^*, only using samples from q_t. Motivated by this, we restrict our search space to the family of curl-free vector fields

$$\mathcal{S}_t = \{\nabla s_t \,|\, s_t : \mathcal{X} \to \mathbb{R}\}\,. \tag{4}$$

We use a neural network to parameterize the set of functions $s_t(x;\theta)$, and will propose the Action Matching objective in Sec. 2 to learn parameters θ such that $s_t(x;\theta)$ approximates $s_t^*(x)$. Once we have learned the vector field, we can move samples forward or backward in time by simulating the ODE in Eq. (1) with the velocity ∇s_t. The continuity equation ensures that for ∇s_t^*, samples at any given time $t \in [0,1]$ are distributed according to q_t.

Note that, even though we arrived at the continuity equation and ground truth vector field $\nabla s_t^*(x)$ using ODEs, the continuity equation can describe a rich family of density evolutions in a wide range of stochastic processes, including those of SDEs (see Equation 37 of Song et al. (2020b)), or

even those of the porous medium equations (Otto, 2001) which are more general than SDEs. Since these processes also define an absolutely continuous curve in the density space, Theorem 2.1 applies. Thus, for the task of modeling the marginal evolution of q_t, our restriction to ODEs using curl-free vector fields does not sacrifice expressivity.

2.2. Action Matching Loss

The main development of this paper is the Action Matching method, which allows us to recover the *true action* s_t^* while having access only to samples from q_t. With this action in hand, we can simulate the continuous dynamics whose evolution matches q_t using the vector field ∇s_t^* (see Fig. 1).

In order to do so, we define the *variational action* $s_t(x)$ parameterized by a neural network, which approximates $s_t^*(x)$ by minimizing the "ACTION-GAP" objective

$$\text{ACTION-GAP}(s, s^*) \coloneqq \frac{1}{2}\int_0^1 \mathbb{E}_{q_t(x)} \|\nabla s_t(x) - \nabla s_t^*(x)\|^2 dt\,. \quad (5)$$

Note that this objective is intractable, as we do not have access to ∇s^*. However as the following proposition shows, we can still derive a tractable objective for minimizing the action gap.

Theorem 2.2. *For an arbitrary variational action s, the* ACTION-GAP(s, s^*) *can be decomposed as the sum of an intractable constant $\mathcal{K}$, and a tractable term $\mathcal{L}_{AM}(s)$*

$$\text{ACTION-GAP}(s_t, s_t^*) = \mathcal{K} + \mathcal{L}_{\text{AM}}(s_t)\,. \quad (6)$$

where $\mathcal{L}_{\text{AM}}(s)$ is the Action Matching objective, which we minimize

$$\begin{aligned}\mathcal{L}_{\text{AM}}(s) \coloneqq\ & \mathbb{E}_{q_0(x)}\big[s_0(x)\big] - \mathbb{E}_{q_1(x)}\big[s_1(x)\big] \\ & + \int_0^1 \mathbb{E}_{q_t(x)}\left[\frac{1}{2}\|\nabla s_t(x)\|^2 + \frac{\partial s_t}{\partial t}(x)\right] dt\end{aligned} \quad (7)$$

See App. A.1 for the proof. The term $\mathcal{L}_{\text{AM}}$ is tractable, since we can use the samples from marginals q_t to obtain an unbiased low variance estimate. We show in App. A.1 that the intractable constant $\mathcal{K}$ is the *kinetic energy* of the distributional path, defined as $\mathcal{K}(\nabla s_t^*) \coloneqq \frac{1}{2}\int_0^1 \mathbb{E}_{q_t(x)}\|\nabla s_t^*(x)\|^2 dt$, and thus minimizing $\mathcal{L}_{AM}(s)$ can be viewed as maximizing a variational lower bound on the kinetic energy.

Connection with Optimal Transport In App. B.1, we show that the optimal dynamics of AM along the curve is also optimal in the sense of optimal transport with the 2-Wasserstein cost. More precisely, at any given time t, the optimal vector field in the AM objective defines a mapping between two infinitesimally close distributions q_t and q_{t+h}, which is of the form $x \mapsto x + h\nabla s_t^*(x)$. This mapping is indeed the same as the Brenier map (Brenier, 1987) in optimal transport, which is of the form $x \mapsto x + \nabla\varphi_t(x)$, where φ_t is the (c-convex) Kantorovich potential.

Algorithm 1 Action Matching [1]

Require: data $\{x_t^j\}_{j=1}^{N_t},\ x_t^j \sim q_t(x)$
Require: parametric model $s_t(x, \theta)$
for learning iterations **do**
  get batch of samples from boundaries:
    $\{x_0^i\}_{i=1}^n \sim q_0(x),\ \{x_1^i\}_{i=1}^n \sim q_1(x)$
  sample times $\{t^i\}_{i=1}^n \sim \text{Uniform}[0, 1]$
  get batch of intermediate samples $\{x_{t^i}^i\}_{i=1}^n \sim q_t(x)$
  $$\mathcal{L}_{\text{AM}}(\theta) = \frac{1}{n}\sum_{i=1}^n \Big[s_0(x_0^i, \theta) - s_1(x_1^i, \theta) + \tfrac{1}{2}\big\|\nabla s_{t^i}(x_{t^i}^i, \theta)\big\|^2 + \tfrac{\partial s_{t^i}(x_{t^i}^i, \theta)}{\partial t}\Big]$$
  update the model $\theta \leftarrow \text{Optimizer}(\theta, \nabla_\theta \mathcal{L}_{\text{AM}}(\theta))$
end for
output trained model $s_t(x, \theta^*)$

Finally, in App. A.1, we adapt reasoning from Albergo & Vanden-Eijnden (2022) to show that the 2-Wasserstein distance between the ground truth marginals and those simulated using our learned $\nabla s_t(x)$ can be upper bounded in terms of ACTION-GAP(s_t, s_t^*).

2.3. Learning, Sampling, and Likelihood Evaluation

Learning We provide pseudo-code for learning with the Action Matching objective in Algorithm 1. With our learned $\nabla s_t(x, \theta)$, we now describe how to simulate the dynamics and evaluate likelihoods when the initial density q_0 is known.

Sampling We sample from the target distribution via the trained function $s_t(x(t), \theta^*)$ by solving the following ODE forward in time:

$$\frac{d}{dt}x(t) = \nabla_x s_t(x(t), \theta^*), \quad x(t = 0) \sim q_0(x). \quad (8)$$

Recall that this sampling process is justified by Eq. (3), where $s_t(x(t), \theta^*)$ approximates $s_t^*(x(t))$.

Evaluating the Log-Likelihood When the density for q_0 is available, we can evaluate the log-likelihood of a sample $x \sim q_1$ using the continuous change of variables formula (Chen et al., 2018). Integrating the ODE backward in time,

$$\begin{aligned}\log q_1(x) &= \log q_0(x(0)) - \int_0^1 dt\ \Delta s_t^*(x(t)), \\ \frac{d}{dt}x(t) &= \nabla_x s_t^*(x(t)), \quad x(t = 1) = x,\end{aligned} \quad (9)$$

[1] Notebooks with pedagogical examples of AM are given at github.com/necludov/jam#tutorials.

where $\frac{d}{dt}\log q_t = -\Delta s_t^*$ can be confirmed using a simple calculation and we approximate $s_t^*(x(t))$ by $s_t(x(t), \theta^*)$.

3. Extensions of Action Matching [2]

In this section, we propose several extensions of Action Matching, which can be used to learn dynamics which include stochasticity (Sec. 3.1), allow for teleportation of probability mass (Sec. 3.2), and minimize alternative kinetic energy costs (Sec. 3.3).

3.1. Entropic Action Matching

In this section, we propose *entropic* Action Matching (eAM), which can recover the ground-truth dynamics arising from diffusion processes with curl-free drift term and known diffusion term. This setting takes place in biological applications studying the Brownian motion of cells in a medium (Schiebinger et al., 2019; Tong et al., 2020). We will show in Prop. 3.1 that, at optimality, entropic AM can *also* learn to trace any absolutely continuous distributional path under mild conditions, so that the choice between entropic AM and deterministic AM should be made based on prior knowledge of the true underlying dynamics.

Consider the stochastic differential equation

$$dx(t) = v_t(x)dt + \sigma_t dW_t\,, \quad x(t=0) = x\,. \tag{10}$$

where W_t is the Wiener process. We know that the evolution of density of this diffusion process is described by the Fokker–Planck equation:

$$\frac{\partial}{\partial t}q_t = -\nabla\cdot(q_t v_t) + \frac{\sigma_t^2}{2}\Delta q_t\,, \tag{11}$$

In the following proposition, we extend Theorem 2.1 and prove that any continuous distributional path, regardless of ground truth generating dynamics, can be modeled with the diffusion dynamics in the state-space.

Proposition 3.1. *Consider a continuous dynamic with the density evolution of q_t, and suppose σ_t is given. Then, there exists a unique (up to a constant) function $\tilde{s}_t^*(x)$, called the "entropic action", such that vector field $v_t^*(x) = \nabla\tilde{s}_t^*(x)$ and q_t satisfies the Fokker-Planck equation*

$$\frac{\partial}{\partial t}q_t = -\nabla\cdot(q_t\nabla\tilde{s}_t^*) + \frac{\sigma_t^2}{2}\Delta q_t\,, \tag{12}$$

See App. A.2 for the proof. This proposition indicates that the we can use the the SDE $dx(t) = \nabla\tilde{s}_t^* dt + \sigma_t dW_t$ to move samples in time such that the marginals are q_t.

[2]This section can be skipped without loss of understanding of our core contributions, although our experiments also evaluate the entropic AM method from Sec. 3.1.

Entropic AM objective aims to recover the unique $\tilde{s}_t^*(x)$, as described by the above proposition. In order to learn the diffusion velocity vector, we define the *variational action* $s_t(x)$, parameterized by a neural network, that approximates $\tilde{s}_t^*(x)$, by minimizing the "E-ACTION-GAP" objective

$$\text{E-ACTION-GAP}(s, \tilde{s}^*) := \frac{1}{2}\int_0^1 \mathbb{E}_{q_t(x)}\|\nabla s_t(x) - \nabla\tilde{s}_t^*(x)\|^2 dt\,.$$

Note that while the E-ACTION-GAP is similar to the original ACTION-GAP objective, it minimizes the distance to $\tilde{s}_t^*$, which is different than s_t^*. As in AM, this objective is intractable since we do not have access to $\nabla\tilde{s}_t^*$. However, we derive a tractable objective in the following proposition.

Proposition 3.2. *For an arbitrary variational action s, the* E-ACTION-GAP(s, s^*) *can be decomposed as the sum of an intractable constant $\mathcal{K}$, and a tractable term $\mathcal{L}_{\text{eAM}}(s)$ which can be minimized:*

$$\text{E-ACTION-GAP}(s, \tilde{s}^*) = \mathcal{L}_{\text{eAM}}(s) + \mathcal{K}_{\text{eAM}}\,,$$

where $\mathcal{L}_{\text{eAM}}(s)$ is the entropic Action Matching objective, which we minimize

$$\mathcal{L}_{\text{eAM}}(s) := \mathbb{E}_{q_0(x)}[s_0(x)] - \mathbb{E}_{q_1(x)}[s_1(x)] + \int_0^1 \mathbb{E}_{q_t(x)}\left[\frac{1}{2}\|\nabla s_t(x)\|^2 + \frac{\partial s_t}{\partial t}(x) + \frac{\sigma_t^2}{2}\Delta s_t\right]dt \tag{13}$$

See App. A.2 for the proof. The constant $\mathcal{K}_{\text{eAM}}$ is the entropic kinetic energy, discussed in App. A.2.

Connection with Entropic Optimal Transport In App. B.3, we describe connections between the eAM objective and dynamical formulations of entropy-regularized optimal transport (Cuturi, 2013) or Schrödinger Bridge (Léonard, 2014; Chen et al., 2016; 2021) problems.

3.2. Unbalanced Action Matching

In this section, we further extend the scope of underlying dynamics which can be learned by Action Matching by allowing for the creation and destruction of probability mass via a growth rate $g_t(x)$. This term is useful to account for cell growth and death in trajectory inference for single-cell biological dynamics (Schiebinger et al., 2019; Tong et al., 2020; Baradat & Lavenant, 2021; Lübeck et al., 2022; Chizat et al., 2022), and is well-studied in relation to *unbalanced* optimal transport problems (Chizat et al., 2018a;b;c; Liero et al., 2016; 2018; Kondratyev et al., 2016).

To introduce *unbalanced* Action Matching (uAM), consider the following ODE, which attaches importance weights to each sample and updates the weights according to a growth rate $g_t(x)$ while transporting the samples in space,

$$\frac{d}{dt}x(t) = v_t(x(t))\,, \quad x(t=0) = x\,. \tag{14}$$

$$\frac{d}{dt}\log w_t(x(t)) = g_t(x(t))\,, \quad w(t=0) = w\,. \tag{15}$$

where v_t is the vector field moving particles, similar to continuity equation, $w_t(x)$ is the importance weight of particles, and $g_t(x(t))$ the growth rate of particles. These importance weights can grow or shrink over time, allowing the particles to be destroyed or create mass probability without needing to transport the particles. The evolution of density governing the importance weighted ODE is given by the following continuity equation:

$$\frac{\partial}{\partial t} q_t = -\nabla \cdot (q_t v_t) + q_t g_t \,. \tag{16}$$

In the following proposition, we extend Theorem 2.1 to show that any distributional path (under mild conditions), regardless of how it was generated in the state-space, can be modeled with the importance weighted ODE.

Proposition 3.3. *Consider a continuous dynamic with density evolution q_t satisfying mild conditions. Then, there exists a unique function $\hat{s}_t^*(x)$, called the "unbalanced action", such that velocity field $v_t^*(x) = \nabla \hat{s}_t^*(x)$ and growth term $g_t^*(x) = \hat{s}_t^*(x)$ satisfy the importance weighted continuity equation:*

$$\frac{\partial}{\partial t} q_t = -\nabla \cdot (q_t \nabla \hat{s}_t^*) + q_t \hat{s}_t^* \,, \tag{17}$$

See App. A.3 for the proof. This proposition indicates that we can use the importance weighted ODE

$$\frac{d}{dt} x(t) = \nabla \hat{s}_t^*(x(t)) \,, \quad x(t=0) = x \,, \tag{18}$$

$$\frac{d}{dt} \log w_t(x(t)) = \hat{s}_t^*(x(t)) \,, \quad w(t=0) = w \,, \tag{19}$$

to move the particles and update their weights in time, such that the marginals are q_t.

Remarkably, the optimal velocity vector field $v_t^* = \nabla \hat{s}_t^*$ and growth rate $g_t^* = \hat{s}_t^*$ in Prop. 3.3 are linked to a single action function $\hat{s}_t^*(x)$. Thus, for learning the variational action $s_t(x)$ in unbalanced AM, we add a term to the "UNBALANCED-ACTION-GAP" objective which encourages s_t to match $\hat{s}_t^*$,

$$\text{U-ACTION-GAP}(s, \hat{s}^*) := \frac{1}{2} \int_0^1 \mathbb{E}_{q_t(x)} \|\nabla s_t(x) - \nabla \hat{s}_t^*(x)\|^2 dt + \frac{1}{2} \int_0^1 \mathbb{E}_{q_t(x)} \|s_t(x) - \hat{s}_t^*(x)\|^2 dt \,.$$

As before, U-ACTION-GAP$(s, \hat{s}^*)$ objective is intractable since we do not have access to $\hat{s}_t^*$. However, as the following proposition shows, we can still derive a tractable objective.

Proposition 3.4. *For an arbitrary variational action s, the* U-ACTION-GAP$(s, \hat{s}^*)$ *can be decomposed as the sum of intractable constants $\mathcal{K}$ and $\mathcal{G}$, and a tractable term $\mathcal{L}_{\text{uAM}}(s)$*

$$\text{U-ACTION-GAP}(s, \hat{s}^*) = \mathcal{K}_{\text{uAM}} + \mathcal{G}_{\text{uAM}} + \mathcal{L}_{\text{uAM}}(s)$$

where $\mathcal{L}_{\text{uAM}}(s)$ is the unbalanced Action Matching objective, which we minimize

$$\mathcal{L}_{\text{uAM}}(s) := \mathbb{E}_{q_0(x)}[s_0(x)] - \mathbb{E}_{q_1(x)}[s_1(x)] + \int_0^1 \mathbb{E}_{q_t(x)} \left[\frac{1}{2} \|\nabla s_t(x)\|^2 + \frac{\partial s_t}{\partial t}(x) + \frac{1}{2} s_t^2 \right] dt \,. \tag{20}$$

See App. A.3 for the proof. The constants $\mathcal{K}_{\text{uAM}}$ and $\mathcal{G}_{\text{uAM}}$ are the unbalanced kinetic and growth energy, defined in App. A.3 and B.4. We note that the entropic and unbalanced extensions of Action Matching can also be combined, as is common in biological applications (Schiebinger et al., 2019; Chizat et al., 2022). To showcase how uAM can handle creation and destruction of mass, without transporting particles, we provide a mixture of Gaussians example in App. E.3.

Connection with Unbalanced Optimal Transport In App. B.4, we show that at any given time t, the optimal dynamics of uAM along the curve is optimal in the sense of the unbalanced optimal transport (Chizat et al., 2018a; Liero et al., 2016; Kondratyev et al., 2016) between two infinitesimally close distributions q_t and q_{t+h}.

3.3. Action Matching with Convex Costs

In App. B.2, we further extend AM to minimize kinetic energies defined using an arbitrary strictly convex cost $c(v_t)$ (Villani (2009, Ch. 7)). For a given path q_t, consider

$$\mathcal{K}_{\text{cAM}} := \inf_{v_t} \int_0^1 \mathbb{E}_{q_t(x)}[c(v_t)] dt \quad \text{s.t.} \quad \frac{\partial}{\partial t} q_t = -\nabla \cdot (q_t v_t).$$

In this case, the unique vector field tracing the density evolution of q_t becomes $v_t^* = \nabla c^*(\nabla \bar{s}_t^*)$, where c^* is the convex conjugate of the c. The corresponding action gap becomes an integral of the Bregman divergence generated by c^*,

$$\text{ACTION-GAP}_{c^*}(s_t, \bar{s}_t^*) := \int_0^1 \int D_{c^*}[\nabla s_t : \nabla \bar{s}_t^*] \, dx_t dt.$$

In practice, we can minimize ACTION-GAP$_{c^*}$ using the following c-Action Matching loss:

$$\mathcal{L}_{\text{cAM}}(s_t) := \int s_0(x_0) \, q_0(x_0) dx_0 - \int s_1(x_1) \, q_1(x_1) dx_1 + \int_0^1 \int \left[c^*(\nabla s_t(x_t)) + \frac{\partial s_t(x_t)}{\partial t} \right] q_t(x_t) dx_t dt \,.$$

For $c(\cdot) = c^*(\cdot) = \frac{1}{2} \| \cdot \|^2$, we recover standard AM. Importantly, the continuity equation for this formulation is

$$\frac{\partial}{\partial t} q_t = -\nabla \cdot (q_t \nabla c^*(\nabla \bar{s}_t^*)) \,. \tag{21}$$

Thus, the sampling is done by integrating the following vector field

$$\frac{d}{dt} x(t) = \nabla c^*(\nabla \bar{s}_t^*, t), \quad x(t=0) = x \,. \tag{22}$$

4. Applications of Action Matching [3]

In this section, we discuss and empirically study applications of Action Matching. We first consider the scenario when the samples from the dynamics q_t are given as a dataset, which is the case for applications in biology and physics. Furthermore, we demonstrate applications of Action Matching in generative modeling, where we would like to learn a tractable model of a target distribution represented by a dataset of samples $x_1 \sim q_1$ without given intermediate samples or densities.

4.1. Population Dynamics in Biology

Action Matching is a natural approach to population dynamics inference in biology. Namely, given the marginal distribution of cells at several timesteps one wants to learn the model of the dynamics of cells (Tong et al., 2020; Bunne et al., 2021; Huguet et al., 2022; Koshizuka & Sato, 2022). It is crucial that Action Matching does not rely on the trajectories of cells since the measurement process destroys cells, thus the individual trajectories are unavailable. We use entropic Action Matching for this task since the ground truth processes are guided by the Brownian motion. Moreover, if the diffusion coefficient is known from the experiment conditions and the drift term is a gradient flow, entropic Action Matching recovers the ground truth drift term.

Synthetic Data We consider synthetic data from (Huguet et al., 2022), which simulates natural dynamics arising in cellular differentiation, including branching and merging of cells. We generate three datasets with $5, 10, 15$ time steps and learn entropic Action Matching with a fixed diffusion coefficient. In Fig. 2, we compare our method with MIOFlow (Huguet et al., 2022) by measuring the distance between the empirical marginals from the dataset and the generated distributions. We see that Action Matching outperforms MIOFlow both in terms of Wasserstein-2 distance and Maximum Mean Discrepancy (MMD) (Gretton et al., 2012). Moreover, unlike MIOFlow, Action Matching performance does not degrade with more timesteps or finer granulation of the dynamics in time. In Fig. 3, we demonstrate that the learned trajectories for Action Matching stay closer to the data marginals compared to the trajectories of MIOFlow.

Embryoid scRNA-Seq Data For a real data example, we consider a embryoid body single-cell RNA sequencing dataset from Moon et al. (2019). We follow the experimental setup from NLSB (Koshizuka & Sato, 2022) and learn entropic Action Matching where the sample space is a 5-dimensional PCA decomposition of the original data. For Action Matching, we interpolate the data using mixtures between given time steps to obtain data which is more dense in time. In Table 1, we compare our method with

[3]The code is available at github.com/necludov/jam

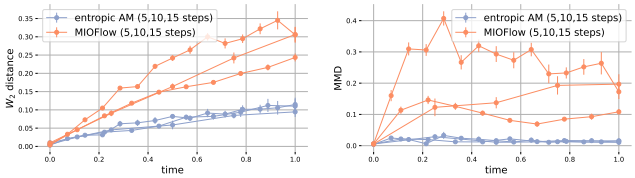

Figure 2. Performance of entropic Action Matching and MIOFlow on the synthetic data. We simulate the dynamics starting from the initial data distribution and estimate Wasserstein-2 distance and MMD between generated distributions and dataset marginals.

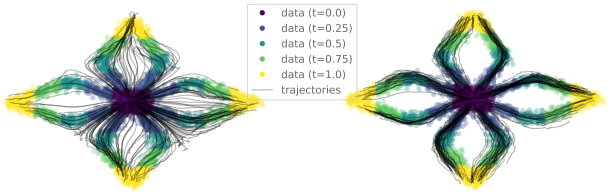

Figure 3. Generated trajectories by MIOFlow (left) and entropic Action Matching (right). The training data is scattered.

OT-flow (Onken et al., 2021), Trajectory-NET (Tong et al., 2020), Iterative Proportional Fitting (IPF) (De Bortoli et al., 2021), Neural SDE (Li et al., 2020), and Neural Lagrangian Schrodinger Bridge (NLSB) (Koshizuka & Sato, 2022). We find that eAM performs on par with NLSB and outperforms other methods.

4.2. Quantum System Simulation

In this section, we apply Action Matching to learn the dynamics of a quantum system evolving according to the Schrödinger equation. The Schrödinger equation describes the evolution of many quantum systems, and in particular, it describes the physics of molecular systems. Here, for the ground truth dynamics, we take the dynamics of an excited state of the hydrogen atom, which is described by the following equation

$$i\frac{\partial}{\partial t}\psi(x,t) = -\frac{1}{\|x\|}\psi(x,t) - \frac{1}{2}\nabla^2\psi(x,t). \tag{23}$$

The function $\psi(x,t) : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{C}$ is called a wavefunction and completely describes the state of the quantum system. In particular, it defines the distribution of the coordinates x by the density $q_t(x) := |\psi(x,t)|^2$, which is defined through the evolution of $\psi(x,t)$ in Eq. (23).

For the baseline, we take Annealed Langevin Dynamics (ALD) as considered in (Song & Ermon, 2019). It approximates the ground truth dynamics using only scores of the distributions by running the approximate MCMC method (which does not have access to the densities) targeting the intermediate distributions of the dynamics (see Algorithm 3). For the estimation of scores, we consider Score Matching (SM) (Hyvärinen & Dayan, 2005), Sliced Score Matching (SSM) (Song et al., 2020a), and additionally evaluate the

Distance↓	OT-flow	Trajectory-Net	IPF	Neural SDE	NLSB	eAM (ours)
$W_2(q_{t_1}, \hat{q}_{t_1})$	0.75	0.64	0.65 ± 0.016	0.62 ± 0.016	0.63 ± 0.015	$\mathbf{0.58 \pm 0.015}$
$W_2(q_{t_2}, \hat{q}_{t_2})$	0.93	0.87	0.78 ± 0.021	0.78 ± 0.021	$\mathbf{0.75 \pm 0.017}$	0.77 ± 0.016
$W_2(q_{t_3}, \hat{q}_{t_3})$	0.93	0.78	0.76 ± 0.018	0.77 ± 0.017	0.75 ± 0.014	$\mathbf{0.72 \pm 0.007}$
$W_2(q_{t_4}, \hat{q}_{t_4})$	0.88	0.89	0.76 ± 0.017	0.75 ± 0.017	$\mathbf{0.72 \pm 0.010}$	0.74 ± 0.017

Table 1. Performance for embryoid body scRNA-seq data as measured by the W_2 distance computed between test data marginals and predicted marginals from the previous test data marginals. Numbers for concurrent methods are taken from (Koshizuka & Sato, 2022).

Method	Average MMD↓
AM (ours)	$\mathbf{5.7 \cdot 10^{-4} \pm 3.1 \cdot 10^{-4}}$
ALD + SM	$4.8 \cdot 10^{-2} \pm 4.8 \cdot 10^{-3}$
ALD + Sliced SM	$4.7 \cdot 10^{-2} \pm 4.0 \cdot 10^{-3}$
ALD + True Scores	$3.6 \cdot 10^{-2} \pm 4.1 \cdot 10^{-4}$

Table 2. Performance of Action Matching and ALD for the Schrödinger equation simulation. We report average MMD over time between the data and the generated samples. For ALD, we use different estimations of scores: Score Matching (SM), Sliced SM, ground true scores.

baseline using the ground truth scores. For further details, we refer the reader to App. E.1 and the supplemented code.

Action Matching outperforms both Score Matching and Sliced Score Matching, precisely simulating the true dynamics (see Table 2). Despite that both SM and SSM accurately recover the ground truth scores for the marginal distributions (see the right plot in Fig. 5 of App. E.1), one cannot efficiently use them for the sampling from the ground truth dynamics. Note, that even using the ground truth scores in ALD does not match the performance of Action Matching (see Table 2) since it is itself an approximation to the Metropolis-Adjusted Langevin Algorithm. Finally, we provide animations of the learned dynamics for different methods (see github.com/necludov/action-matching) to illustrate the performance difference.

4.3. Generative Modeling

In addition to learning stochastic dynamics in natural sciences, Action Matching has a wide range of applications in generative modeling. The key feature of Action Matching for generative modeling is the flexibility to choose *any* dynamics q_t that starts from the prior distribution q_0 and arrives to the given data q_1. After using Action Matching to learn a vector field $\nabla s^\star$ which simulates the appropriate dynamics, we can sample and evaluate likelihoods using Eq. (8) and Eq. (9). We consider a flexible approach for specifying dynamics through interpolants in the sample-space in Eq. (24), and instantiate it for specific problem settings in the following paragraphs. We defer analysis of experimental results in each setting to Sec. 4.3.1.

There are many degrees of freedom in the definition of the density path between a given q_0 and q_1 = dataset. For instance, we consider interpolating between samples $x_0 \sim q_0$ and $x_1 \sim q_1$ from the prior and data distributions using

$$x_t = \alpha_t(x_0) + \beta_t(x_1), \quad x_0 \sim q_0(x), \quad x_1 \sim q_1(x)\,, \quad (24)$$

where α_t, β_t are some continuous transformations of the prior and the data correspondingly. To respect the endpoint marginals q_0, q_1, we select α_t and β_t such that $\alpha_0(x_0) = x_0, \beta_0(x_1) = 0$ and $\alpha_1(x_0) = 0, \beta_1(x_1) = x_1$. The path in density space q_t is implicitly defined via these samples.

Unconditional Generation A simple choice of α_t, β_t is the linear interpolation between a simple prior and the data,

$$x_t = (1-t) \cdot x_0 + t \cdot x_1, \quad x_0 \sim \mathcal{N}(0,1), \quad x_1 \sim \text{dataset}\,,$$

We demonstrate below that these dynamics can be learned with Action Matching, which allows us to perform unconditional image generation and log-likelihood evaluation.

Conditional Generation We can also consider conditional generation tasks by simulating the dynamics only for the subset of dimensions. For instance, we can lower the resolution of images in the dataset and learn the dynamics that performs super-resolution using

$$x_t = \text{mask} \cdot x_1 + (1 - \text{mask}) \cdot ((1-t) \cdot x_0 + t \cdot x_1), \\ x_0 \sim \mathcal{N}(0,1), \quad x_1 \sim \text{dataset}\,,$$

where `mask` is the binary valued vector that defines the subset of dimensions to condition on, e.g., which pixels we generate. Note that we still can learn the dynamics for all the dimensions of vector x_t. For super-resolution, we keep one pixel in each 2×2 block, while the remaining pixels are gradually transformed to the standard normal random variable.

Conditional Coupled Generation In some applications, the conditions and the predicted variables might be coupled, or share common dimensions. We still can perform conditional generation by considering the marginal distribution of conditions as a prior. To showcase this scenario, we consider image colorization, for which we define the following dynamics

$$x_t = (1-t) \cdot (10^{-1} \cdot x_0 + \text{gray}(x_1)) + t \cdot x_1, \\ x_0 \sim \mathcal{N}(0,1), \quad x_1 \sim \text{dataset}\,,$$

where the function $\text{gray}(x_1)$ returns the grayscale version of image x_1. We found that minor distortion of the inputs

Model	BPD↓	FID↓	IS↑	NFE↓
VP-SDE (uses extra information)	3.25	3.71	9.12	199
Flow Matching (uses extra information)	2.99	6.35	–	142
Baseline (ALD + SSM)	-	86.29	5.43	1090
AM - generation (ours)	3.54	10.04	8.60	132
Baseline (ALD + SSM)	–	n/a	n/a	n/a
AM - superres (ours)	–	1.44	10.93	166
AM - color (ours)	–	2.47	9.88	89

Table 3. Model performance for CIFAR-10. When possible, we report log-likelihood in bits per dimension (BPD). For all tasks, we report FID and IS evaluated for 50k generated images.

with Normal noise stabilizes the training significantly, while barely destroying the visual information of the image (see Fig. 6). We also can restore this information by passing the non-distorted grayscale image to the model.

4.3.1. Empirical Study for Generative Modeling

We demonstrate the performance of Action Matching for the aforementioned generative modeling tasks. For evaluation, we choose the CIFAR-10 dataset of natural images. Despite the fact that we do not consider image generation to be the main application of Action Matching, we argue that it demonstrates the scalability and applicability of the proposed method for the following reasons: (i) the data is high-dimensional and challenging; (ii) the quality of generated samples is easy to assess; (iii) we use known deep learning architectures, hence, remove this component from the scope of our study.

For the sake of proper comparison, we consider a baseline model using Annealed Langevin Dynamics (ALD) (Song & Ermon, 2019) and Sliced Score Matching (SSM) (Song et al., 2020a), which estimates the scores of the marginals $\nabla \log q_t$ *without knowledge* of the underlying dynamics. As stronger baselines which rely on a known form for the dynamics, we consider the variance-preserving diffusion model (VP-SDE) from Song et al. (2020b) and Flow Matching from (Lipman et al., 2022). We train all models using the same architecture for 500k iterations and evaluate the negative log-likelihood, FID (Heusel et al., 2017) and IS (Salimans et al., 2016). For Flow Matching, we report the numbers from (Lipman et al., 2022).

In Table 3, we compare the performance of all models, finding that Action Matching performs much closer to VP-SDE and Flow Matching (which rely on the knowledge of the dynamics) than to the baseline, with fewer function evaluations (NFE). Moreover, for the conditional generation, the ALD+SSM baseline fails to learn any meaningful scores. We argue that this is due to the dynamics starting from a complicated prior distribution, whose scores are difficult to learn without the annealing process to the noise distribution (Song & Ermon, 2019). Although VP-SDE achieves better FID and IS scores than AM in Table 3, in Fig. 4, we find only minor qualitative differences in comparing images generated from the corresponding ODEs starting from the same noise values. We provide all generations in App. F and discuss implementation details in App. C and E.2.

5. Related Works

A large body of previous work on continuous normalizing flows (Chen et al., 2018) has considered regularization inspired by dynamical optimal transport. In particular, Finlay et al. (2020); Tong et al. (2020) regularize the kinetic energy $\frac{1}{2}\|v_t(x)\|^2$ (and its Jacobian) across time, while Yang & Karniadakis (2020); Onken et al. (2021); Koshizuka & Sato (2022) parameterize v_t as the gradient of a potential function s_t and *additionally* regularize $|\frac{1}{2}\|\nabla s_t(x)\|^2 + \frac{\partial s_t}{\partial t}|$ to be close to zero. However, these methods all require simulation of and backpropagation through the differential equations during training. Further, the loss functions for these include maximum likelihood or adversarial losses at the endpoints or entropic OT losses to intermediate distributions. By contrast, Action Matching is simulation-free and evaluates the action function at the endpoints. The Action Matching objective in Eq. (7) naturally includes minimization of $\frac{1}{2}\|\nabla s_t(x)\|^2 + \frac{\partial s_t}{\partial t}$, while we show in App. B.1 that the minimum of $\frac{1}{2}\|v_t(x)\|^2$ is *also* achieved by the action-minimizing $\nabla s_t(x)$. In fact, these are dual optimizations, as shown in App. B.2 or Mikami & Thieullen (2006). Thus, including *both* regularizers is redundant from the perspective of the Action Matching loss.

Existing methods for trajectory inference often optimize an entropy-regularized (unbalanced) optimal transport loss between the observed samples q_t and samples simulated from a current model $\hat{q}_t$, which requires backpropagation through OT solvers (Cuturi, 2013; Chizat et al., 2018b; Cuturi et al., 2022). In particular, Tong et al. (2020); Koshizuka & Sato (2022); Huguet et al. (2022) sample $\hat{q}_t$ using continuous normalizing flows with the regularization described above. The JKONET scheme of Bunne et al. (2022) seeks to learn a time-independent potential function $\mathcal{F}[q] = \int q(x)s(x)$ for which the observed q_t path is a gradient flow, and generates samples from $\hat{q}_t$ using learned optimal transport maps. As we discuss in App. B.5, Action Matching can be viewed as learning a time-dependent potential via $s_t(x)$, without simulation of $\hat{q}_t$ during training, using a simple objective.

5.1. Connection with Flow Matching Methods

Recent related work on Flow Matching (FM) (Lipman et al., 2022; Pooladian et al., 2023; Tong et al., 2023), stochastic interpolants (Albergo & Vanden-Eijnden, 2022), and Rectified Flow (Liu et al., 2022; Liu, 2022) have recently been understood under the umbrella of *bridge matching* (BM) methods (Shi et al., 2023; Peluchetti, 2023; Liu et al., 2023). We highlight several distinctions between Action Matching and flow matching methods in this section.

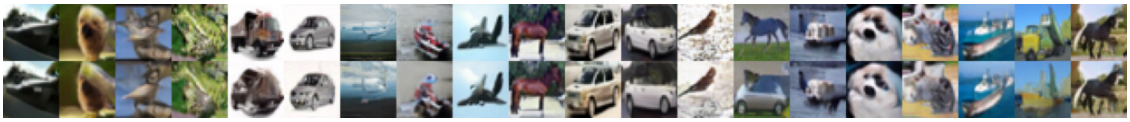

Figure 4. Generated samples for CIFAR-10 using VP-SDE model (top row) and Action Matching (bottom row).

In particular, consider a deterministic *bridge* which produces intermediate samples $x_t = f_t(x_0, x_1)$ with a tractable conditional vector field $u_t(x_t|x_0, x_1)$. A natural example is linear interpolation of the endpoint samples. A marginal vector field $u_t^\theta(x_t)$, which does not condition on a data sample x_1 and thus can be used for unconditional generation, is learned to match $u_t(x_t|x_0, x_1)$ using a squared error loss

$$\mathcal{L}_{\text{FM}}(u_t^\theta) = \inf_{u_t} \int_0^1 \mathbb{E}_{q(x_0,x_1)} \left[\|u(x_t|x_0, x_1) - u_t^\theta(x_t)\|^2 \right] dt.$$

Liu (2022) justifies other Bregman divergence losses, which parallels our c-Action Matching method in Sec. 3.3.

The $\mathcal{L}_{\text{FM}}$ objective can be decomposed as in (Banerjee et al., 2005),

$$\begin{aligned} &\mathbb{E}_{q_{0,1}} \left[\|u(x_t|x_0, x_1) - u_t^\theta(x_t)\|^2 \right] = \\ &\mathbb{E}_{q_{0,1}} [\|u(x_t|x_0, x_1) - u_t^*(x_t)\|^2] + \mathbb{E}_{q_t} [\|u_t^*(x_t) - u_t^\theta(x_t)\|^2], \end{aligned} \tag{25}$$

where $u_t^*(x_t)$ is the unique minimizer of $\mathcal{L}_{\text{FM}}$ given by

$$u_t^*(x_t) = \mathbb{E}_{q(x_0,x_1)} [u(x_t|x_0, x_1)] \,. \tag{26}$$

We highlight several key differences between FM and AM.

Target Vector Fields We can see that the ACTION-GAP (Eq. (5)) in AM is analogous to the second term on the right-hand side of Eq. (25), using a different target and parameterization. While both the FM target vector field u_t^* and the AM target vector field ∇s_t^* yield the same marginals q_t (Theorem 2.1), they have several key differences that can be understood through the Helmholtz decomposition (Ambrosio et al., 2008, Lemma 8.4.2). This theorem states that any vector field, such as u_t^*, can be uniquely decomposed as $u_t^* = \nabla s^* + w_t$, where ∇s^* is the gradient-field component, and w_t is the divergence-free component, $\nabla \cdot (q_t w_t) = 0$.

Among all vector fields respecting marginals q_t, the AM target ∇s_t^* is the unique vector field without a divergence-free component, moving the particles optimally in the sense of optimal transport. By contrast, the FM target u_t^* may contain a divergence-free component, moving the particles in a way that reflects the underlying *path measure* (Shi et al., 2023) defined through the endpoint distribution $q(x_0, x_1)$ and the bridge process. Indeed, the gradient-field component of u_t^* is identical to ∇s_t^*, and the divergence-free component of u_t^* does not influence the marginals. As the result, the optimal ∇s_t^* has a lower kinetic energy or dynamical cost

$$\frac{1}{2}\mathbb{E}_{q_t(x_t)} \|\nabla s_t^*\|^2 \leq \frac{1}{2}\mathbb{E}_{q_t(x_t)} \|u_t^*\|^2 \quad \forall\, t \,. \tag{27}$$

Applications Starting from the ACTION-GAP, Action Matching derives a tractable dual objective which removes the need for a known conditional vector field $u_t(x_t|x_0, x_1)$. Thus, Action Matching may be applied in settings where linear interpolation is not a reasonable inductive bias, or no bridge vector field $u_t(x_t|x_0, x_1)$ is available. In particular, in Sec. 4, we applied AM to scientific applications where the dynamics of interest are only given via a dataset of samples from the temporal marginals.

Optimization We note that FM methods are able to optimize the simple L_2 loss $\mathbb{E}[\|u(x_t|x_0, x_1) - u_t^\theta(x_t)\|^2]$ due to the assumption that the target conditional vector field $u(x_t|x_0, x_1)$ is available in closed form. Furthermore, parametrizing $u_t^\theta : [0, 1] \times \mathbb{R}^d \to \mathbb{R}^d$ directly using a neural network, rather than parametrizing $s_t : [0, 1] \times \mathbb{R}^d \to \mathbb{R}$ and differentiating to obtain $\nabla s_t \in \mathbb{R}^d$ in AM, provides computational benefits by saving a backward pass. Nevertheless, we have previously noted the benefits of the gradient-field parametrization and the generality of Action Matching.

6. Conclusion

In this work, we have presented a family of Action Matching methods for learning continuous dynamics from samples along arbitrary absolutely continuous paths on the 2-Wasserstein space. We propose the tractable optimization objective that relies *only* on samples from intermediate distributions. We further derived three extensions of Action Matching: Entropic AM, that can model stochasticity in the state space dynamics; Unbalanced AM, that allows for creation and destruction of probability mass; and c-AM, that incorporates convex cost functions on the state space.

The key property of the proposed objective is that it can be efficiently optimized for a wide range of applications. We demonstrate this empirically in the physical and natural sciences, where snapshots of samples are often given. Further, we demonstrated competitive performance of Action Matching for generative modeling of images, where a prior distribution and density path can be flexibly constructed depending on the task of interest.

Acknowledgement The authors thank Viktor Oganesyan for helpful discussions regarding reproducing VP-SDE results. AM acknowledges support from the Canada CIFAR AI Chairs program.

References

Albergo, M. S. and Vanden-Eijnden, E. Building normalizing flows with stochastic interpolants. *arXiv preprint arXiv:2209.15571*, 2022.

Ambrosio, L., Gigli, N., and Savaré, G. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2008.

Banerjee, A., Merugu, S., Dhillon, I. S., Ghosh, J., and Lafferty, J. Clustering with bregman divergences. *Journal of machine learning research*, 6(10), 2005.

Baradat, A. and Lavenant, H. Regularized unbalanced optimal transport as entropy minimization with respect to branching brownian motion. *arXiv preprint arXiv:2111.01666*, 2021.

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

Brenier, Y. Décomposition polaire et réarrangement monotone des champs de vecteurs. *CR Acad. Sci. Paris Sér. I Math.*, 305:805–808, 1987.

Buenrostro, J. D., Wu, B., Litzenburger, U. M., Ruff, D., Gonzales, M. L., Snyder, M. P., Chang, H. Y., and Greenleaf, W. J. Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature*, 523(7561): 486–490, 2015.

Bunne, C., Stark, S. G., Gut, G., del Castillo, J. S., Lehmann, K.-V., Pelkmans, L., Krause, A., and Rätsch, G. Learning single-cell perturbation responses using neural optimal transport. *bioRxiv*, 2021.

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

Chen, R. T., Rubanova, Y., Bettencourt, J., and Duvenaud, D. K. Neural ordinary differential equations. *Advances in neural information processing systems*, 31, 2018.

Chen, Y., Georgiou, T. T., and Pavon, M. On the relation between optimal transport and Schrödinger bridges: A stochastic control viewpoint. *Journal of Optimization Theory and Applications*, 169(2):671–691, 2016.

Chen, Y., Georgiou, T. T., and Pavon, M. Stochastic control liaisons: Richard sinkhorn meets gaspard monge on a schrodinger bridge. *SIAM Review*, 63(2):249–313, 2021.

Chizat, L., Peyré, G., Schmitzer, B., and Vialard, F.-X. An interpolating distance between optimal transport and fisher–rao metrics. *Foundations of Computational Mathematics*, 18(1):1–44, 2018a.

Chizat, L., Peyré, G., Schmitzer, B., and Vialard, F.-X. Scaling algorithms for unbalanced optimal transport problems. *Mathematics of Computation*, 87(314):2563–2609, 2018b.

Chizat, L., Peyré, G., Schmitzer, B., and Vialard, F.-X. Unbalanced optimal transport: Dynamic and Kantorovich formulations. *Journal of Functional Analysis*, 274(11): 3090–3123, 2018c.

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

Cuturi, M. Sinkhorn distances: Lightspeed computation of optimal transport. *Advances in neural information processing systems*, 26, 2013.

Cuturi, M., Meng-Papaxanthos, L., Tian, Y., Bunne, C., Davis, G., and Teboul, O. Optimal transport tools (ott): A jax toolbox for all things Wasserstein. *arXiv preprint arXiv:2201.12324*, 2022. URL `https://github.com/ott-jax/ott`.

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

Figalli, A. and Glaudo, F. *An Invitation to Optimal Transport, Wasserstein Distances, and Gradient Flows*. 2021.

Finlay, C., Jacobsen, J.-H., Nurbekyan, L., and Oberman, A. How to train your neural ode: the world of jacobian and kinetic regularization. In *International conference on machine learning*, pp. 3154–3164. PMLR, 2020.

Gangbo, W. and McCann, R. J. The geometry of optimal transportation. *Acta Mathematica*, 177(2):113–161, 1996.

Gretton, A., Borgwardt, K. M., Rasch, M. J., Schölkopf, B., and Smola, A. A kernel two-sample test. *The Journal of Machine Learning Research*, 13(1):723–773, 2012.

Griffiths, D. J. and Schroeter, D. F. *Introduction to quantum mechanics*. Cambridge university press, 2018.

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

Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., and Hochreiter, S. Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in neural information processing systems*, 30, 2017.

Ho, J., Jain, A., and Abbeel, P. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33:6840–6851, 2020.

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

Hyvärinen, A. and Dayan, P. Estimation of non-normalized statistical models by score matching. *Journal of Machine Learning Research*, 6(4), 2005.

Klein, A. M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D. A., and Kirschner, M. W. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell*, 161(5):1187–1201, 2015.

Kondratyev, S., Monsaingeon, L., and Vorotnikov, D. A new optimal transport distance on the space of finite radon measures. *Advances in Differential Equations*, 21(11/12): 1117–1164, 2016.

Koshizuka, T. and Sato, I. Neural lagrangian Schrödinger bridge, 2022. URL `https://arxiv.org/abs/2204.04853`.

Lavenant, H., Zhang, S., Kim, Y.-H., and Schiebinger, G. Towards a mathematical theory of trajectory inference. *arXiv preprint arXiv:2102.09204*, 2021.

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

Li, X., Wong, T.-K. L., Chen, R. T., and Duvenaud, D. Scalable gradients for stochastic differential equations. In *International Conference on Artificial Intelligence and Statistics*, pp. 3870–3882. PMLR, 2020.

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

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

Lipman, Y., Chen, R. T., Ben-Hamu, H., Nickel, M., and Le, M. Flow matching for generative modeling. *arXiv preprint arXiv:2210.02747*, 2022.

Liu, Q. Rectified flow: A marginal preserving approach to optimal transport. *arXiv preprint arXiv:2209.14577*, 2022.

Liu, X., Gong, C., and Liu, Q. Flow straight and fast: Learning to generate and transfer data with rectified flow. *arXiv preprint arXiv:2209.03003*, 2022.

Liu, X., Wu, L., Ye, M., et al. Learning diffusion bridges on constrained domains. In *The Eleventh International Conference on Learning Representations*, 2023.

Lübeck, F., Bunne, C., Gut, G., del Castillo, J. S., Pelkmans, L., and Alvarez-Melis, D. Neural unbalanced optimal transport via cycle-consistent semi-couplings. *arXiv preprint arXiv:2209.15621*, 2022.

Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A. R., Kamitaki, N., Martersteck, E. M., et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell*, 161(5):1202–1214, 2015.

Mikami, T. and Thieullen, M. Duality theorem for the stochastic optimal control problem. *Stochastic processes and their applications*, 116(12):1815–1835, 2006.

Mikhailov, V. *Partial differential equations*. 1976.

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

Noé, F., Tkatchenko, A., Müller, K.-R., and Clementi, C. Machine learning for molecular simulation. *Annual review of physical chemistry*, 71:361–390, 2020.

Onken, D., Fung, S. W., Li, X., and Ruthotto, L. Ot-flow: Fast and accurate continuous normalizing flows via optimal transport. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 9223–9232, 2021.

Otto, F. The geometry of dissipative evolution equations: the porous medium equation. 2001.

Peluchetti, S. Diffusion bridge mixture transports, schr\" odinger bridge problems and generative modeling. *arXiv preprint arXiv:2304.00917*, 2023.

Pooladian, A.-A., Ben-Hamu, H., Domingo-Enrich, C., Amos, B., Lipman, Y., and Chen, R. Multisample flow matching: Straightening flows with minibatch couplings. *arXiv preprint arXiv:2304.14772*, 2023.

Ronneberger, O., Fischer, P., and Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241. Springer, 2015.

Salimans, T. and Ho, J. Should ebms model the energy or the score? In *Energy Based Models Workshop-ICLR 2021*, 2021.

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., and Chen, X. Improved techniques for training gans. *Advances in neural information processing systems*, 29, 2016.

Schiebinger, G. Reconstructing developmental landscapes and trajectories from single-cell data. *Current Opinion in Systems Biology*, 27:100351, 2021.

Schiebinger, G., Shu, J., Tabaka, M., Cleary, B., Subramanian, V., Solomon, A., Gould, J., Liu, S., Lin, S., Berube, P., et al. Optimal-transport analysis of single-cell gene expression identifies developmental trajectories in reprogramming. *Cell*, 176(4):928–943, 2019. URL https://www.cell.com/cms/10.1016/j.cell.2019.01.006/attachment/d251b88d-a356-436a-a9a7-b7e04b56b41f/mmc8.pdf.

Shi, Y., De Bortoli, V., Campbell, A., and Doucet, A. Diffusion schr\" odinger bridge matching. *arXiv preprint arXiv:2303.16852*, 2023.

Song, Y. and Ermon, S. Generative modeling by estimating gradients of the data distribution. *Advances in Neural Information Processing Systems*, 32, 2019.

Song, Y., Garg, S., Shi, J., and Ermon, S. Sliced score matching: A scalable approach to density and score estimation. In *Uncertainty in Artificial Intelligence*, pp. 574–584. PMLR, 2020a.

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

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

Tong, A., Malkin, N., Huguet, G., Zhang, Y., Rector-Brooks, J., Fatras, K., Wolf, G., and Bengio, Y. Conditional flow matching: Simulation-free dynamic optimal transport. *arXiv preprint arXiv:2302.00482*, 2023.

Vázquez, J. L. *The porous medium equation: mathematical theory*. Oxford University Press on Demand, 2007.

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

Yang, L. and Karniadakis, G. E. Potential flow generator with l2 optimal transport regularity for generative models. *IEEE Trans. Neural Networks Learn. Syst.*, 33(2):528–538, 2020. doi: 10.1109/TNNLS.2020.3028042. URL https://doi.org/10.1109/TNNLS.2020.3028042.

A. Proofs

A.1. Action Matching Proofs

Theorem 2.1 (Adapted from Theorem 8.3.1 of Ambrosio et al. (2008))**.** *Consider a continuous dynamic with the density evolution of q_t, which satisfies mild conditions (absolute continuity in the 2-Wasserstein space of distributions $\mathcal{P}_2(\mathcal{X})$). Then, there exists a unique (up to a constant) function $s_t^*(x)$, called the "action", such that vector field $v_t^*(x) = \nabla s_t^*(x)$ and q_t satisfies the continuity equation*

$$\frac{\partial}{\partial t} q_t = -\nabla \cdot \left(q_t \nabla s_t^*(x)\right). \tag{3}$$

In other words, the ODE $\frac{d}{dt}x(t) = \nabla s_t^*(x)$ can be used to move samples in time such that the marginals are q_t.

Proof of existence and uniqueness. The existence of and uniqueness of the solution could be argued by observing that

$$\begin{aligned} \frac{\partial}{\partial t} q_t &= -\nabla \cdot (q_t \nabla s_t^*) && \text{in } X, \\ \langle \nabla s_t^*, \mathbf{n} \rangle &= 0 && \text{on } \partial X, \end{aligned}$$

where $\mathbf{n}$ is the surface normal, is an elliptic PDE with the Neumann boundary condition, and that it is a classical fact that these PDEs have a solution under mild conditions on q_t (Mikhailov, 1976). See Ambrosio et al. (2008) for a proof in more general settings. □

Theorem 2.2. *For an arbitrary variational action s, the* ACTION-GAP(s, s^*) *can be decomposed as the sum of an intractable constant $\mathcal{K}$, and a tractable term $\mathcal{L}_{AM}(s)$*

$$\text{ACTION-GAP}(s_t, s_t^*) = \mathcal{K} + \mathcal{L}_{\text{AM}}(s_t). \tag{6}$$

where $\mathcal{L}_{\text{AM}}(s)$ is the Action Matching objective, which we minimize

$$\begin{aligned} \mathcal{L}_{\text{AM}}(s) \coloneqq\ & \mathbb{E}_{q_0(x)}\left[s_0(x)\right] - \mathbb{E}_{q_1(x)}\left[s_1(x)\right] \\ & + \int_0^1 \mathbb{E}_{q_t(x)}\left[\frac{1}{2}\|\nabla s_t(x)\|^2 + \frac{\partial s_t}{\partial t}(x)\right] dt \end{aligned} \tag{7}$$

Proof.

$$\begin{aligned}
&\text{ACTION-GAP}(s_t, s_t^*) \\
&= \frac{1}{2}\int_0^1 \omega_t \mathbb{E}_{q_t(x)} \|\nabla s_t - \nabla s_t^*\|^2 dt \\
&= \frac{1}{2}\int_0^1 \int_X \omega_t q_t(x) \|\nabla s_t - \nabla s_t^*\|^2 dx dt \\
&= \frac{1}{2}\int_0^1 \int_X \omega_t q_t(x) \|\nabla s_t\|^2 dx dt - \int_0^1 \omega_t \int_X q_t(x) \langle \nabla s_t(x), \nabla s_t^*(x) \rangle dx dt + \overbrace{\frac{1}{2}\int \mathbb{E}_{q_t(x)} \|\nabla s_t^*\|^2 dt}^{\mathcal{K}_{\text{AM}}} \\
&= \frac{1}{2}\int_0^1 \int_X \omega_t q_t(x) \|\nabla s_t\|^2 dx dt - \int_0^1 \omega_t \int_X \langle \nabla s_t(x), q_t(x) \nabla s_t^*(x) \rangle dx dt + \mathcal{K}_{\text{AM}} \\
&\overset{(1)}{=} \frac{1}{2}\int_0^1 \int_X \omega_t q_t(x) \|\nabla s_t\|^2 dx dt + \int_0^1 \omega_t \int_X s_t(x) [\nabla \cdot (q_t(x) \nabla s_t^*(x))] dx dt - \int_0^1 \omega_t \oint_{\partial X} q_t(x) s_t(x) \cancelto{0}{\langle \nabla s_t^*, d\mathbf{n} \rangle} dt + \mathcal{K}_{\text{AM}} \\
&= \frac{1}{2}\int_0^1 \int_X \omega_t q_t(x) \|\nabla s_t\|^2 dx dt - \int_0^1 \left(\int_X \omega_t s_t(x) \frac{\partial q_t(x)}{\partial t} dx \right) dt + \mathcal{K}_{\text{AM}} \\
&\overset{(2)}{=} \int_0^1 \omega_t \mathbb{E}_{q_t(x)} \left[\frac{1}{2} \|\nabla s_t(x)\|^2 \right] dt - \left(\omega_t \mathbb{E}_{q_t(x)} [s_t(x)] \Big|_{t=0}^{t=1} - \int_X \mathbb{E}_{q_t(x)} \left[s_t(x) \frac{d\omega_t}{dt} + \omega_t \frac{\partial s_t(x)}{\partial t} \right] dt \right) + \mathcal{K}_{\text{AM}} \\
&= \int_0^1 \omega_t \mathbb{E}_{q_t(x)} \left[\frac{1}{2} \|\nabla s_t(x)\|^2 + \frac{\partial s_t(x)}{\partial t} + s_t(x) \frac{d \log \omega_t}{dt} \right] dt - \omega_1 \mathbb{E}_{q_1(x)} [s_1(x)] + \omega_0 \mathbb{E}_{q_0(x)} [s_0(x)] + \mathcal{K}_{\text{AM}} \\
&= \mathcal{L}_{\text{AM}}(s) + \mathcal{K}_{\text{AM}}
\end{aligned}$$

where in (1), we have used integration by parts for divergence operator $\int_X \langle \nabla g, \mathbf{f} \rangle dx = \oint_{\partial X} \langle \mathbf{f} g, d\mathbf{n} \rangle - \int_X g(\nabla \cdot \mathbf{f}) dx$ and that $\langle \nabla s_t^*, d\mathbf{n} \rangle|_{\partial X} = 0$ due to the Neumann boundary condition (see proof of Theorem 2.1 above), and in (2) we have used integration by parts. □

For each distributional path q_t, the kinetic energy term only depends on the true actions s_t^* and is defined as

$$\mathcal{K}(\nabla s_t^*) \coloneqq \frac{1}{2} \int_0^1 \mathbb{E}_{q_t(x)} \|\nabla s_t^*(x)\|^2 dt \,. \tag{28}$$

Thus, minimizing the $\mathcal{L}_{\text{AM}}(s)$ can be interpreted as maximizing a variational lower bound on kinetic energy.

Proposition A.1 (Adapted from Albergo & Vanden-Eijnden (2022))**.** *Let $\nabla s_t(x)$ be a learned vector field, continuously differentiable in (t, x) and Lipschitz in x uniformly on $[0, 1] \times \mathbb{R}^d$ with Lipschitz constant K. Let $\hat{q}_t$ denote the density induced by $\frac{\partial x}{\partial t} = \nabla s_t(x)$ with $x_0 \sim q_0$. Then, the squared Wasserstein-2 distance between $\hat{q}_t$ and the ground truth q_t (induced by $\nabla s_t^*(x)$) at each time $\tau \in [0, 1]$ is bounded by*

$$W_2^2(\hat{q}_\tau, q_\tau) \leq e^{(1+2K)\tau} \int_0^\tau \mathbb{E}_{q_t(x)} \|\nabla s_t(x) - \nabla s_t^*(x)\|^2 dt$$

where the right-hand side includes the action gap in Eq. (5)*.*

Proof. Consider two flows $\varphi_t(x)$ and $\varphi_t^*(x)$ defined by

$$\frac{\partial \varphi_t(x)}{\partial t} = \nabla s_t(\varphi_t(x)) \;\text{ and }\; \frac{\partial \varphi_t^*(x)}{\partial t} = \nabla s_t^*(\varphi_t^*(x)) \tag{29}$$

correspondingly. Consider the quantity

$$Q_t = \int dx \; q_0(x) \|\varphi_t(x) - \varphi_t^*(x)\|^2. \tag{30}$$

Note that $W_2^2(\hat{q}_\tau, q_\tau) \leq Q_\tau$, since $(\varphi_\tau)_\# q_0 = \hat{q}_\tau$ and $(\varphi_\tau^*)_\# q_0 = q_\tau$ have the correct marginals and thus constitute an admissible coupling for the W_2 distance. The time derivative of Q_t is

$$\begin{aligned}
\frac{\partial Q_t}{\partial t} &= 2 \int dx \; q_0(x) \left\langle \varphi_t(x) - \varphi_t^*(x), \frac{\partial \varphi_t(x)}{\partial t} - \frac{\partial \varphi_t^*(x)}{\partial t} \right\rangle && (31)\\
&= 2 \int dx \; q_0(x) \langle \varphi_t(x) - \varphi_t^*(x), \nabla s_t(\varphi_t(x)) - \nabla s_t^*(\varphi_t^*(x)) \rangle && (32)\\
&= 2 \int dx \; q_0(x) \langle \varphi_t(x) - \varphi_t^*(x), \nabla s_t(\varphi_t(x)) - \nabla s_t(\varphi_t^*(x)) \rangle && (33)\\
&\quad + 2 \int dx \; q_0(x) \langle \varphi_t(x) - \varphi_t^*(x), \nabla s_t(\varphi_t^*(x)) - \nabla s_t^*(\varphi_t^*(x)) \rangle.
\end{aligned}$$

By the Lipshitzness of $\nabla s_t(x)$, we bound the first term as

$$2 \langle \varphi_t(x) - \varphi_t^*(x), \nabla s_t(\varphi_t(x)) - \nabla s_t(\varphi_t^*(x)) \rangle \leq 2K \|\varphi_t(x) - \varphi_t^*(x)\|^2. \tag{34}$$

The second term is bounded as follows

$$\|\varphi_t(x) - \varphi_t^*(x)\|^2 - 2 \langle \varphi_t(x) - \varphi_t^*(x), \nabla s_t(\varphi_t^*(x)) - \nabla s_t^*(\varphi_t^*(x)) \rangle + \|\nabla s_t(\varphi_t^*(x)) - \nabla s_t^*(\varphi_t^*(x))\|^2 \geq 0, \tag{35}$$

$$2 \langle \varphi_t(x) - \varphi_t^*(x), \nabla s_t(\varphi_t^*(x)) - \nabla s_t^*(\varphi_t^*(x)) \rangle \leq \|\varphi_t(x) - \varphi_t^*(x)\|^2 + \|\nabla s_t(\varphi_t^*(x)) - \nabla s_t^*(\varphi_t^*(x))\|^2. \tag{36}$$

Thus,

$$\frac{\partial Q_t}{\partial t} \leq (1 + 2K) Q_t + \int dx \; q_0(x) \|\nabla s_t(\varphi_t^*(x)) - \nabla s_t^*(\varphi_t^*(x))\|^2, \tag{37}$$

and by Gronwall's lemma, we have

$$Q_\tau \leq Q_0 \exp(\tau(1 + 2K)) \int_0^\tau dt \int dx \; q_0(x) \|\nabla s_t(\varphi_t^*(x)) - \nabla s_t^*(\varphi_t^*(x))\|^2. \tag{38}$$

Clearly, $Q_0 = 0$. Using Eq. (38) and the fact that $W_2^2(\hat{q}_\tau, q_\tau) \leq Q_\tau$ from above, the proposition is proven. □

A.2. Entropic Action Matching Proofs

Proposition 3.1. *Consider a continuous dynamic with the density evolution of q_t, and suppose σ_t is given. Then, there exists a unique (up to a constant) function $\tilde{s}_t^*(x)$, called the "entropic action", such that vector field $v_t^*(x) = \nabla \tilde{s}_t^*(x)$ and q_t satisfies the Fokker-Planck equation*

$$\frac{\partial}{\partial t} q_t = -\nabla \cdot (q_t \nabla \tilde{s}_t^*) + \frac{\sigma_t^2}{2} \Delta q_t \,, \tag{12}$$

Proof of existence and uniqueness. Consider the PDE

$$\begin{aligned} \frac{\partial}{\partial t} q_t &= -\nabla \cdot (q_t \nabla \tilde{s}_t^*) + \frac{\sigma_t^2}{2} \Delta q_t && \text{in } X \\ \langle \nabla \tilde{s}_t^*, \mathbf{n} \rangle &= \frac{\sigma_t^2}{2} \langle \nabla \log q_t, \mathbf{n} \rangle && \text{on } \partial X \end{aligned}$$

with the reparametrization $s_t^* = \tilde{s}_t^* - \frac{\sigma_t^2}{2} \log q_t$, we can write this PDE as

$$\begin{aligned} \frac{\partial}{\partial t} q_t &= -\nabla \cdot (q_t \nabla \tilde{s}_t^*) + \frac{\sigma_t^2}{2} \Delta q_t \\ &= -\nabla \cdot \left[q_t \left(\nabla \tilde{s}_t^* - \frac{\sigma_t^2}{2} \nabla \log q_t \right) \right] \\ &= -\nabla \cdot \left[q_t \left(\nabla \left(\tilde{s}_t^* - \frac{\sigma_t^2}{2} \log q_t \right) \right) \right] \\ &= -\nabla \cdot (q_t \nabla s_t^*) \end{aligned}$$

with the boundary condition of

$$\left\langle \nabla \left(\tilde{s}^* - \frac{\sigma_t^2}{2} \log q_t \right), \mathbf{n} \right\rangle = \langle \nabla s_t^*, \mathbf{n} \rangle = 0$$

From App. A.1, we know s_t^* exists and is unique (up to a constant), thus $\tilde{s}_t^*$ also exists and is unique (up to a constant).

□

Proposition 3.2. *For an arbitrary variational action s, the* E-ACTION-GAP(s, s^*) *can be decomposed as the sum of an intractable constant $\mathcal{K}$, and a tractable term $\mathcal{L}_{\text{eAM}}(s)$ which can be minimized:*

$$\text{E-ACTION-GAP}(s, \tilde{s}^*) = \mathcal{L}_{\text{eAM}}(s) + \mathcal{K}_{\text{eAM}} \,,$$

where $\mathcal{L}_{\text{eAM}}(s)$ is the entropic Action Matching objective, which we minimize

$$\begin{aligned} \mathcal{L}_{\text{eAM}}(s) \coloneqq\ & \mathbb{E}_{q_0(x)}[s_0(x)] - \mathbb{E}_{q_1(x)}[s_1(x)] \\ & + \int_0^1 \mathbb{E}_{q_t(x)} \left[\frac{1}{2} \|\nabla s_t(x)\|^2 + \frac{\partial s_t}{\partial t}(x) + \frac{\sigma_t^2}{2} \Delta s_t \right] dt \end{aligned} \tag{13}$$

Proof.

$$
\begin{aligned}
&\text{E-ACTION-GAP}(s_t, \tilde{s}_t^*) \\
&= \frac{1}{2}\int_0^1 \mathbb{E}_{q_t(x)}\|\nabla s_t - \nabla \tilde{s}_t^*\|^2 dt \\
&= \frac{1}{2}\int_0^1 \int_X q_t(x)\|\nabla s_t - \nabla \tilde{s}_t^*\|^2 dx dt \\
&= \frac{1}{2}\int_0^1 \int_X q_t(x)\|\nabla s_t\|^2 dx dt - \int_0^1 \int_X q_t(x)\langle \nabla s_t(x), \nabla \tilde{s}_t^*(x)\rangle dx dt + \overbrace{\frac{1}{2}\int \mathbb{E}_{q_t(x)}\|\nabla \tilde{s}_t^*\|^2 dt}^{\mathcal{K}_{\text{eAM}}} \\
&= \frac{1}{2}\int_0^1 \int_X q_t(x)\|\nabla s_t\|^2 dx dt - \int_0^1 \int_X \langle \nabla s_t(x), q_t(x)\nabla \tilde{s}_t^*(x)\rangle dx dt + \mathcal{K}_{\text{eAM}} \\
&\overset{(1)}{=} \frac{1}{2}\int_0^1 \int_X q_t(x)\|\nabla s_t\|^2 dx dt + \int_0^1 \int_X s_t(x)[\nabla \cdot (q_t(x)\nabla \tilde{s}_t^*(x))] dx dt - \frac{\sigma_t^2}{2}\int_0^1 \oint_{\partial X} s_t \langle \nabla q_t, d\mathbf{n}\rangle dt + \mathcal{K}_{\text{eAM}} \\
&\overset{(2)}{=} \frac{1}{2}\int_0^1 \int_X q_t(x)\|\nabla s_t\|^2 dx dt - \int_0^1 \left(\int_X s_t(x)\frac{\partial}{\partial t} q_t(x) dx\right) dt + \frac{\sigma_t^2}{2}\int_0^1 \left(\int_X s_t(x)\Delta q_t dx\right) dt - \frac{\sigma_t^2}{2}\int_0^1 \oint_{\partial X} s_t \langle \nabla q_t, d\mathbf{n}\rangle dt + \mathcal{K}_{\text{eAM}} \\
&\overset{(3)}{=} \int_0^1 \mathbb{E}_{q_t(x)}\left[\frac{1}{2}\|\nabla s_t(x)\|^2\right] dt - \left(\mathbb{E}_{q_t(x)}[s_t(x)]\Big|_{t=0}^{t=1} - \int_X \mathbb{E}_{q_t(x)}\left[\frac{\partial s_t(x)}{\partial t}\right] dt\right) + \frac{\sigma_t^2}{2}\int_0^1 \left(\int_X q_t(x)\Delta s_t dx\right) dt \\
&\quad - \frac{\sigma_t^2}{2}\int_0^1 \oint_{\partial X} q_t \langle \nabla s_t, d\mathbf{n}\rangle dt + \mathcal{K}_{\text{eAM}} \\
&= \int_0^1 \mathbb{E}_{q_t(x)}\left[\frac{1}{2}\|\nabla s_t(x)\|^2 + \frac{\partial s_t(x)}{\partial t} + \frac{\sigma_t^2}{2}\Delta s_t\right] dt - \mathbb{E}_{q_1(x)}[s_1(x)] + \mathbb{E}_{q_0(x)}[s_0(x)] - \frac{\sigma_t^2}{2}\int_0^1 \oint_{\partial X} q_t \langle \nabla s_t, d\mathbf{n}\rangle dt + \mathcal{K}_{\text{eAM}} \\
&= \mathcal{L}_{\text{eAM}}(s) + \mathcal{K}_{\text{eAM}}
\end{aligned}
$$

where in (1), we have used the integration by parts for divergence operator $\int_X \langle \nabla g, \mathbf{f}\rangle dx = \oint_{\partial X}\langle \mathbf{f} g, d\mathbf{n}\rangle - \int_X g(\nabla \cdot \mathbf{f}) dx$, and the Neumann boundary condition from App. A.2

$$
\begin{aligned}
\int_X \langle \nabla s_t(x), q_t(x)\nabla \tilde{s}_t^*(x)\rangle dx &= \oint_{\partial X} q_t(x) s_t(x)\langle \nabla \tilde{s}_t^*, d\mathbf{n}\rangle - \int_X s_t(x)\nabla \cdot (q_t(x)\nabla \tilde{s}_t^*(x)) dx && \text{integration by parts} \\
&= \frac{\sigma_t^2}{2}\oint_{\partial X} s_t q_t \langle \nabla \log q_t, d\mathbf{n}\rangle - \int_X s_t(x)\nabla \cdot (q_t(x)\nabla \tilde{s}_t^*(x)) dx && \text{boundary condition} \\
&= \frac{\sigma_t^2}{2}\oint_{\partial X} s_t \langle \nabla q_t, d\mathbf{n}\rangle - \int_X s_t(x)\nabla \cdot (q_t(x)\nabla \tilde{s}_t^*(x)) dx
\end{aligned}
$$

In (2), we have used the Fokker-Planck equation: $\nabla \cdot (q_t \nabla s_t^*(x)) = -\frac{\partial}{\partial t} q_t + \frac{\sigma_t^2}{2}\Delta q_t$.

To derive (3), we have

$$
\begin{aligned}
\int_X s_t(x)\Delta q_t dx - \int_{\partial X} s_t \langle \nabla q_t, d\mathbf{n}\rangle &= \int_X s_t(x)[\nabla \cdot (\nabla q_t)] dx - \int_{\partial X} s_t \langle \nabla q_t, d\mathbf{n}\rangle \\
&= -\int_X \langle \nabla s_t, \nabla q_t\rangle dx && \text{integration by parts for divergence theorem} \\
&= -\int_X \langle \nabla q_t, \nabla s_t\rangle dx && \text{by symmetry} \\
&= \int_X q_t(x)\Delta s_t dx - \int_{\partial X} q_t \langle \nabla s_t, d\mathbf{n}\rangle && \text{following the reverse of the first 2 steps}
\end{aligned}
$$

Finally, note that the term $\frac{\sigma_t^2}{2}\int_0^1 \oint_{\partial X} q_t \langle \nabla s_t, d\mathbf{n}\rangle dt$ can be dropped if q_t vanishes on the boundary ∂X. □

For each distributional path q_t, the *entropic kinetic energy* term only depends on the true entropic action $\tilde{s}_t^*$ and is defined as

$$
\mathcal{K}_{\text{eAM}}(\nabla s_t^*) \coloneqq \frac{1}{2}\int_0^1 \mathbb{E}_{q_t(x)}\|\nabla \tilde{s}_t^*(x)\|^2 dt\,. \tag{39}
$$

Thus, minimizing $\mathcal{L}_{\text{eAM}}(s)$ can be interpreted as maximizing a variational lower bound on the entropic kinetic energy.

A.3. Unbalanced Action Matching Proofs

Proposition 3.3. *Consider a continuous dynamic with density evolution q_t satisfying mild conditions. Then, there exists a unique function $\hat{s}_t^*(x)$, called the "unbalanced action", such that velocity field $v_t^*(x) = \nabla \hat{s}_t^*(x)$ and growth term $g_t^*(x) = \hat{s}_t^*(x)$ satisfy the importance weighted continuity equation:*

$$\frac{\partial}{\partial t} q_t = -\nabla \cdot (q_t \nabla \hat{s}_t^*) + q_t \hat{s}_t^* \,, \tag{17}$$

Proof of existence and uniqueness. The existence and uniqueness of the solution could be argued by observing that

$$\begin{aligned} \frac{\partial}{\partial t} q_t &= -\nabla \cdot (q_t \nabla \hat{s}_t^*) + \hat{s}_t^* q_t && \text{in } X \\ \langle \nabla \hat{s}_t^*, \mathbf{n} \rangle &= 0 && \text{on } \partial X \end{aligned}$$

is an elliptic PDE with the Neumann boundary condition, and that it is a classical fact that these PDEs have a solution under mild conditions on q_t (Mikhailov, 1976).

□

Proposition 3.4. *For an arbitrary variational action s, the* U-ACTION-GAP$(s, \hat{s}^*)$ *can be decomposed as the sum of intractable constants $\mathcal{K}$ and $\mathcal{G}$, and a tractable term $\mathcal{L}_{\text{uAM}}(s)$*

$$\text{U-ACTION-GAP}(s, \hat{s}^*) = \mathcal{K}_{\text{uAM}} + \mathcal{G}_{\text{uAM}} + \mathcal{L}_{\text{uAM}}(s)$$

where $\mathcal{L}_{\text{uAM}}(s)$ is the unbalanced Action Matching objective, which we minimize

$$\begin{aligned} \mathcal{L}_{\text{uAM}}(s) \coloneqq\ & \mathbb{E}_{q_0(x)}[s_0(x)] - \mathbb{E}_{q_1(x)}[s_1(x)] \\ & + \int_0^1 \mathbb{E}_{q_t(x)} \left[\frac{1}{2} \|\nabla s_t(x)\|^2 + \frac{\partial s_t}{\partial t}(x) + \frac{1}{2} s_t^2 \right] dt \,. \end{aligned} \tag{20}$$

Proof.

$$\begin{aligned}
&\text{U-ACTION-GAP}(s_t, \hat{s}_t^*) \\
&= \frac{1}{2} \int_0^1 \mathbb{E}_{q_t(x)} \|\nabla s_t - \nabla \hat{s}_t^*\|^2 dt + \frac{1}{2} \int_0^1 \mathbb{E}_{q_t(x)} \|s_t - \hat{s}_t^*\|^2 dt \\
&= \frac{1}{2} \int_0^1 \int_X q_t(x) \|\nabla s_t - \nabla \hat{s}_t^*\|^2 dx dt + \frac{1}{2} \int_0^1 \int_X q_t(x) \|s_t - \hat{s}_t^*\|^2 dx dt \\
&= \int_0^1 \int_X q_t(x) \left[\frac{1}{2} \|\nabla s_t\|^2 + \frac{1}{2} s_t^2 \right] dx dt - \int_0^1 \int_X q_t(x) \langle \nabla s_t(x), \nabla \hat{s}_t^*(x) \rangle dx dt - \int_0^1 \int_X q_t(x) s_t(x) \hat{s}_t^*(x) dx dt \\
&\quad + \overbrace{\frac{1}{2} \int \mathbb{E}_{q_t(x)} \left\| \nabla \hat{s}_t^{*2} \right\|^2 dt}^{\mathcal{K}_{\text{uAM}}} + \overbrace{\frac{1}{2} \int \mathbb{E}_{q_t(x)} \hat{s}_t^{*2} dt}^{\mathcal{G}_{\text{uAM}}} \\
&= \int_0^1 \int_X q_t(x) \left[\frac{1}{2} \|\nabla s_t\|^2 + \frac{1}{2} s_t^2 \right] dx dt - \int_0^1 \int_X \langle \nabla s_t(x), q_t(x) \nabla \hat{s}_t^*(x) \rangle dx dt - \int_0^1 \int_X q_t(x) s_t(x) \hat{s}_t^*(x) dx dt + \mathcal{K}_{\text{uAM}} + \mathcal{G}_{\text{uAM}} \\
&\overset{(1)}{=} \int_0^1 \int_X q_t(x) \left[\frac{1}{2} \|\nabla s_t\|^2 + \frac{1}{2} s_t^2 \right] dx dt + \int_0^1 \int_X s_t(x) [\nabla \cdot (q_t(x) \nabla \hat{s}_t^*(x)) - q_t(x) \hat{s}_t^*(x)] dx dt - \int_0^1 \oint_{\partial X} q_t(x) s_t(x) \cancelto{0}{\langle \nabla s_t^*, d\mathbf{n} \rangle} dt \\
&\quad + \mathcal{K}_{\text{uAM}} + \mathcal{G}_{\text{uAM}} \\
&\overset{(2)}{=} \int_0^1 \int_X q_t(x) \left[\frac{1}{2} \|\nabla s_t\|^2 + \frac{1}{2} s_t^2 \right] dx dt - \int_0^1 \left(\int_X s_t(x) \frac{\partial}{\partial t} q_t(x) dx \right) dt + \mathcal{K}_{\text{uAM}} + \mathcal{G}_{\text{uAM}} \\
&\overset{(3)}{=} \int_0^1 \mathbb{E}_{q_t(x)} \left[\frac{1}{2} \|\nabla s_t(x)\|^2 + \frac{1}{2} s_t^2 \right] dt - \left(\mathbb{E}_{q_t(x)}[s_t(x)] \Big|_{t=0}^{t=1} - \int_X \mathbb{E}_{q_t(x)} \left[\frac{\partial s_t(x)}{\partial t} \right] dt \right) + \mathcal{K}_{\text{uAM}} + \mathcal{G}_{\text{uAM}} \\
&= \int_0^1 \mathbb{E}_{q_t(x)} \left[\frac{1}{2} \|\nabla s_t(x)\|^2 + \frac{\partial s_t(x)}{\partial t} + \frac{1}{2} s_t^2 \right] dt - \mathbb{E}_{q_1(x)}[s_1(x)] + \mathbb{E}_{q_0(x)}[s_0(x)] + \mathcal{K}_{\text{uAM}} + \mathcal{G}_{\text{uAM}} \\
&= \mathcal{L}_{\text{uAM}}(s) + \mathcal{K}_{\text{uAM}} + \mathcal{G}_{\text{uAM}}
\end{aligned}$$

where in (1), we have used integration by parts for divergence operator $\int_X \langle \nabla g, \mathbf{f} \rangle dx = \oint_{\partial X} \langle \mathbf{f} g, d\mathbf{n} \rangle - \int_X g(\nabla \cdot \mathbf{f}) dx$ and that $\langle \nabla s_t^*, d\mathbf{n} \rangle|_{\partial X} = 0$ due to the Neumann boundary condition (see App. A.3), in (2) we have used: $\frac{\partial}{\partial t} q_t = -\nabla \cdot (q_t \nabla s^*) + s^* q_t$, and in (3) we have integration by parts. □

For each distributional path q_t, the *unbalanced kinetic energy* and *unbalanced growth energy* terms only depend on the true unbalanced action $\hat{s}_t^*$ and are defined as

$$\mathcal{K}_{\text{uAM}}(\nabla \hat{s}_t^*) \coloneqq \frac{1}{2} \int_0^1 \mathbb{E}_{q_t(x)} \|\nabla \hat{s}_t^*(x)\|^2 dt \,. \qquad \mathcal{G}_{\text{uAM}}(\hat{s}_t^*) \coloneqq \frac{1}{2} \int_0^1 \mathbb{E}_{q_t(x)} \hat{s}_t^*(x)^2 dt \,. \tag{40}$$

Thus, minimizing the $\mathcal{L}_{\text{uAM}}(s)$ can be interpreted as maximizing a variational lower bound on the summation of unbalanced kinetic energy and growth energy.

B. Action Matching and Optimal Transport

In this section, we describe various connections between Action Matching and optimal transport. First, we describe how Action Matching (App. B.1), entropic AM (App. B.3), and unbalanced AM (App. B.4) can be understood as solving 'local' versions of optimal transport problems between infinitesimally close distributions, where connections with dynamical OT formulations play a key role (Benamou & Brenier, 2000; Chen et al., 2016; Chizat et al., 2018a; Liero et al., 2016). In App. B.2, we generalize AM by considering Lagrangian action cost functions beyond the squared Euclidean OT cost and standard kinetic energy underlying the other results in this paper. Finally, in App. B.5, we interpret AM as learning a linear potential functional whose Wasserstein-2 gradient flow traces the given density path q_t.

B.1. Action Matching as Infinitesimal Optimal Transport

Using Ambrosio et al. (2008) Prop. 8.4.6 (see also Villani (2009) Remark 13.10), we can interpret ∇s_t^* in action matching as learning the optimal transport map between two infinitesimally close distributions on the given curve q_t, under the squared Euclidean cost. In particular, by Ambrosio et al. (2008) Prop. 8.4.6, we have

$$\nabla s_t^* = \lim_{dt \to 0} \frac{1}{dt} (T^*(q_t, q_{t+dt}) - \text{id}). \tag{41}$$

where id is the identity mapping and $T^*(q_t, q_{t+dt}) = x_t + \nabla \varphi_{t,dt}^*(x_t)$ is the unique transport map (Gangbo & McCann (1996) Thm. 4.5) solving the Monge formulation of the Wasserstein-2 distance between neighboring densities q_t, q_{t+dt}

$$W_2^2(q_t, q_{t+dt}) = \inf_T \left\{ \int \|T(x) - x\|^2 q_t(x) dx \;\middle|\; T_\# q_t = q_{t+dt} \right\} = \frac{1}{2} \mathbb{E}_{q_t(x)} \left[\|\nabla \varphi_{t,dt}^*(x)\|^2 \right] \;\text{ s.t. }\; (x + \nabla \varphi_{t,dt}^*)_\# q_t = q_{t+dt} \tag{42}$$

and $T_\# q_t$ is the pushforward density of q_t under the map T. Note that the continuity equation suggests the pushforward map $x + dt \nabla s_t^*(x)$ for small dt. Thus, among all vector fields, the Action Matching objective finds the one that satisfies the continuity equation while minimizing the infinitesimal displacement of samples according to the squared Euclidean cost. We can also compare Eq. (42) to Eq. (44) below.

Kinetic Energy and Dynamical Optimal Transport To further understand this result, we observe that Action Matching is closely related to the dynamical formulation of optimal transport due to Benamou & Brenier (2000). The squared Wasserstein-2 distance between given densities p_0 and p_1 with densities q_0 and q_1 can be expressed as,

$$W_2^2(\mu, \nu) = \inf_{q_t} \inf_{v_t} \frac{1}{2} \int_0^1 \mathbb{E}_{q_t(x)} \|v_t(x)\|^2 dt \qquad \text{s.t. } \frac{\partial}{\partial t} q_t = -\nabla \cdot (q_t v_t),\; q_0 = p_0, \text{ and } q_1 = p_1. \tag{43}$$

However, in Action Matching, the intermediate densities or distributional path q_t are *fixed* via the given samples. In this case, we can interpret AM as learning a *local* or infinitesimal optimal transport map as in Eq. (41), although the given q_t may not trace the *global* optimal transport path between q_0 and q_1. Nevertheless, it can be shown that (Ambrosio et al. (2008))

$$\nabla s_t^* = \arg\min_{v_t} \frac{1}{2} \int_0^1 \mathbb{E}_{q_t(x)} \left[\|v_t(x)\|^2 \right] dt \qquad \text{s.t. } \frac{\partial}{\partial t} q_t = -\nabla \cdot (q_t v_t). \tag{44}$$

To show this directly, we introduce a Lagrange multiplier s_t to enforce the continuity equation and integrate by parts with respect to x and t (also see Eq. (58)),

$$\mathcal{L}(v_t, s_t) = \sup_{s_t} \inf_{v_t} \frac{1}{2} \int_0^1 \int \mathbb{E}_{q_t(x)} \left[\|v_t(x)\|^2 \right] dt + \int_0^1 \int s_t(x_t) \Big(\frac{\partial q_t}{\partial t}(x_t) + \nabla \cdot \big(q_t(x_t) v_t(x_t) \big) \Big) dx_t dt \tag{45}$$

$$= \sup_{s_t} \inf_{v_t} \frac{1}{2} \int_0^1 \mathbb{E}_{q_t(x)} \left[\|v_t(x)\|^2 \right] dt + \mathbb{E}_{q_t(x_t)} [s_t(x_t)] \Big|_{t=0}^{t=1} - \int_0^1 \int \mathbb{E}_{q_t(x_t)} \left[\frac{\partial s_t(x_t)}{\partial t} + \langle v_t(x_t), \nabla s_t(x_t) \rangle \right] dt \tag{46}$$

Eliminating the optimization with respect to v_t, we obtain the optimality condition

$$v_t(x) = \nabla s_t(x) \tag{47}$$

Plugging this condition into the Lagrangian in Eq. (46), we obtain the AM objective from Theorem 2.2, $\inf_{s_t} \mathcal{L}_{\text{AM}}(s_t)$, where $\mathcal{L}_{\text{AM}}(s_t) = \int_0^1 \mathbb{E}_{q_t(x)} \left[\frac{1}{2} \|\nabla s_t(x)\|^2 + \frac{\partial s_t(x_t)}{\partial t} \right] dt - \mathbb{E}_{q_t(x_t)} [s_t(x_t)] \big|_{t=0}^{t=1}$. This objective is an upper bound on the optimal kinetic energy $\mathcal{K}(\nabla s_t^*)$, which is evaluated using the optimal vector field $v_t^* = \nabla s_t^*$ in Eq. (44).

Metric Derivative Finally, the AM objective and Eq. (44) are related to the *metric derivative* in the 2-Wasserstein space $\mathcal{P}_2(\mathcal{X})$. By Ambrosio et al. (2008) Thm. 8.3.1, for q_t, q_{t+dt} along an absolutely continuous density path with $\partial_t q_t = -\nabla \cdot (q_t v_t^*)$, we have

$$|q_t'|^2 = \left(\lim_{dt \to 0} \frac{W_2(q_t, q_{t+dt})}{dt} \right)^2 = \mathbb{E}_{q_t(x)} \left[\|v_t^*(x)\|^2 \right] = \mathbb{E}_{q_t(x)} \left[\|\nabla s_t^*(x)\|^2 \right]. \tag{48}$$

The optimal value of the AM objective at each t thus reflects the infinitesimal squared W2 distance along the given path.

B.2. Action Matching with Lagrangian Costs

Starting from the Benamou-Brenier dynamical formulation of optimal transport in Eq. (43), we can also formulate action matching with more general Lagrangian action cost functions. For a Lagrangian $L(x_t, v_t, t)$ which is strictly convex in the velocity v_t, we define the dynamical optimal transport problem

$$\mathcal{T}(p_0, p_1) = \inf_{q_t} \inf_{v_t} \int_0^1 \int L(x_t, v_t, t)\, q_t(x_t) dx_t dt \quad \text{subj:} \quad \frac{\partial q_t}{\partial t}(x_t) = \nabla \cdot \big(q_t(x_t) v_t(x_t) \big) \quad q_0 = p_0 \text{ and } q_1 = p_1. \tag{49}$$

See Villani (2009) Ch. 7 for in-depth analysis of Eq. (49) as an optimal transport problem, where $L(x_t, v_t, t) = \frac{1}{2}\|v_t\|^2$ recovers the Wasserstein-2 distance. For action matching, assume that the intermediate path of q_t is given via its samples. We now seek to learn a vector field which minimizes the Lagrangian while satifying the continuity equation,

$$\mathcal{K}_{\ell\text{AM}}^* \coloneqq \inf_{v_t} \int_0^1 \int L(x_t, v_t, t)\, q_t(x_t) dx_t dt \quad \text{subj:} \quad \frac{\partial q_t}{\partial t} = -\nabla \cdot \big(q_t v_t \big). \tag{50}$$

For this Lagrangian cost, we prove the following analogue of Theorem 2.2, which is recovered for $L(x_t, v_t, t) = \frac{1}{2}\|v_t\|^2$. We first state a definition.

Definition B.1. The convex conjugate of a Lagrangian $L(x_t, v_t, t)$ which is strictly convex in v_t is defined as

$$H(x_t, a_t, t) = \sup_{v_t} \langle v_t, a_t \rangle - L(x_t, v_t, t). \tag{51}$$

Solving for the optimizing argument yields the dual correspondence $v_t = \nabla_{a_t} H(x_t, a_t, t)$ and $a_t = \nabla_{v_t} L(x_t, v_t, t)$.

Proposition B.2. *For a Lagrangian $L(x_t, v_t, t)$ which is strictly convex in the velocity v_t, the optimization defining $\mathcal{K}^*_{\ell\text{AM}}$ in Eq.* (50) *can be expressed via the dual optimization,*

$$\mathcal{K}^*_{\ell\text{AM}} = \inf_{s_t} \mathcal{L}_{\ell\text{AM}}(s_t), \tag{52}$$

$$\textit{where} \quad \mathcal{L}_{\ell\text{AM}}(s_t) := \int s_0(x_0)\, q_0(x_0) dx_0 - \int s_1(x_1)\, q_1(x_1) dx_1 + \int_0^1 \int \left[H(x_t, \nabla s_t(x_t), t) + \frac{\partial s_t(x_t)}{\partial t} \right] q_t(x_t) dx_t dt$$

For a variational s_t, the corresponding action gap between s_t and the optimal $\bar{s}^*_t$, is written using a Bregman divergence

$$\mathcal{K}^*_{\ell\text{AM}} + \mathcal{L}_{\ell\text{AM}}(s_t) \;=\; \text{ACTION-GAP}_H(s_t, \bar{s}^*_t) \tag{53}$$

$$\textit{where} \quad \text{ACTION-GAP}_H(s_t, \bar{s}^*_t) := \int_0^1 \int D_H[\nabla s_t(x_t) : \nabla \bar{s}^*_t(x_t)]\; q_t(x_t)\; dx_t dt \tag{54}$$

where the Bregman divergence D_H is defined as

$$D_H[\nabla s_t(x_t) : \nabla \bar{s}^*_t(x_t)] = H(x_t, \nabla s_t, t) - H(x_t, \nabla \bar{s}^*_t, t) - \langle v^*_t(x_t), \nabla s_t(x_t) - \nabla \bar{s}^*_t(x_t) \rangle. \tag{55}$$

Using Definition B.1 and the fact that v^*_t and $\nabla \bar{s}^*_t$ are duals related by $v^*_t = \nabla_{a_t} H(x_t, \nabla \bar{s}^*_t, t)$, the Bregman divergence may also be written in mixed parameterization $D_H[\nabla s_t : \nabla \bar{s}^*_t] = D_{L,H}[v^*_t : \nabla s_t]$, with

$$D_{L,H}[v^*_t : \nabla s_t(x_t)] = L(x_t, v^*_t, t) + H(x_t, \nabla s_t, t) - \langle v^*_t(x_t), \nabla s_t(x_t) \rangle. \tag{56}$$

For the case of $L(x_t, v_t, t) = \frac{1}{2}\|v_t\|^2$, the Hamiltonian is $H(x_t, \nabla s_t, t) = \frac{1}{2}\|\nabla s_t\|^2$ and the two parameterizations are self-dual with $v_t = \nabla s_t$. Using this transformation, the Bregman divergence is simply half the squared Euclidean norm,

$$D_{L,H}[v^*_t : \nabla s_t] = \frac{1}{2}\|v^*_t\|^2 + \frac{1}{2}\|\nabla s_t\|^2 - \langle v^*_t, \nabla s_t \rangle = \frac{1}{2}\|\nabla \bar{s}^*_t\|^2 + \frac{1}{2}\|\nabla s_t\|^2 - \langle \nabla \bar{s}^*_t, \nabla s_t \rangle = \frac{1}{2}\|\nabla \bar{s}^*_t - \nabla s_t\|^2. \tag{57}$$

From Definition B.1, note that in general, the optimality condition translating between v_t and ∇s_t is $v_t = \nabla_{a_t} H(x_t, \nabla s_t, t)$ or $\nabla s_t = \nabla_{v_t} L(x_t, v_t, t)$.

Proof. Introducing a Lagrange multipler $s_t(x)$ to enforce the continuity equation constraint, we integrate by parts in both x and t using the assumption of the boundary condition $\langle v_t, \mathbf{n} \rangle = 0$,

$$\begin{aligned}
\mathcal{K}^*_{\ell\text{AM}} &= \inf_{v_t} \int_0^1 \int L(x_t, v_t, t)\; q_t(x_t) dx_t dt \qquad \text{subj:} \quad \frac{\partial q_t}{\partial t}(x_t) = -\nabla \cdot \big(q_t(x_t) v_t(x_t)\big) \\
&= \sup_{s_t} \inf_{v_t} \int_0^1 \int L(x_t, v_t, t) q_t(x_t) dx_t dt + \int_0^1 \int s_t(x_t) \Big(\frac{\partial q_t}{\partial t}(x_t) + \nabla \cdot \big(q_t(x_t) v_t(x_t)\big) \Big) dx_t dt \\
&= \sup_{s_t} \inf_{v_t} \int_0^1 \int L(x_t, v_t, t)\; q_t(x_t) dx_t dt + \int s_t(x_t) q_t(x_t) dx_t \Big|_{t=0}^{t=1} - \int_0^1 \int \frac{\partial s_t(x_t)}{\partial t} q_t(x_t) dx_t dt \\
&\qquad \int_0^1 \oint_{\partial X} q_t(x_t) s_t(x_t) \cancelto{0}{\langle v_t(x_t), d\mathbf{n} \rangle} dt - \int_0^1 \int \langle v_t(x_t), \nabla s_t(x_t) \rangle\; q_t(x_t) dx_t dt \\
&= \sup_{s_t} \int_0^1 \int -\Big(\sup_{v_t} \langle v_t(x_t), \nabla s_t(x_t) \rangle - L(x_t, v_t, t) \Big) q_t(x_t) dx_t dt - \int_0^1 \int \frac{\partial s_t(x_t)}{\partial t} q_t(x_t) dx_t dt + \int s_t(x_t) q_t(x_t) dx_t \Big|_{t=0}^{t=1}
\end{aligned} \tag{58}$$

We can recognize the highlighted terms as a Legendre transform, where we define the Hamiltonian H as the convex conjugate $H(x_t, \nabla s_t(x_t), t) = \sup_{v_t} \langle v_t(x_t), \nabla s_t(x_t)\rangle - L(x_t, v_t, t)$ of the Lagrangian L. This finally results in

$$\begin{aligned}\mathcal{K}^*_{\ell\text{AM}} &\coloneqq \sup_{s_t} \int s_1(x_1)\, q_1(x_1) dx_1 - \int s_0(x_0)\, q_0(x_0) dx_0 - \int_0^1 \int \Big(H(x_t, \nabla s_t(x_t), t) + \frac{\partial s_t(x_t)}{\partial t} \Big)\, q_t(x_t) dx_t dt. \\ &= -\mathcal{L}_{\ell\text{AM}}(s_t^*)\end{aligned} \tag{59}$$

where Eq. (59) can be used to define a corresponding action matching objective $\mathcal{L}_{\ell\text{AM}}(s_t)$ for a variational s_t (see Eq. (52)). To calculate the action gap, we subtract $\mathcal{L}_{\ell\text{AM}}(s_t)$ from the optimal value of the 'kinetic energy' in Eq. (50), using v_t^*

$$\begin{aligned}&\text{ACTION-GAP}_H(s_t, \bar{s}_t^*) \\ &\quad\coloneqq \mathcal{K}^*_{\ell\text{AM}} + \mathcal{L}_{\ell\text{AM}}(s_t) \\ &\quad= \mathcal{L}_{\ell\text{AM}}(s_t) - \mathcal{L}_{\ell\text{AM}}(\bar{s}_t^*) \qquad \text{(using Eq. (59))} \\ &\quad= \int s_0(x_0)\, q_0(x_0) dx_0 - \int s_1(x_1)\, q_1(x_1) dx_1 + \int_0^1 \int \Big(H(x_t, \nabla s_t(x_t), t) + \frac{\partial s_t(x_t)}{\partial t} \Big)\, q_t(x_t) dx_t dt \\ &\qquad - \left(\int \bar{s}_0^*(x_0)\, q_0(x_0) dx_0 - \int \bar{s}_1^*(x_1)\, q_1(x_1) dx_1 + \int_0^1 \int \Big(H(x_t, \nabla \bar{s}_t^*(x_t), t) + \frac{\partial \bar{s}_t^*(x_t)}{\partial t} \Big)\, q_t(x_t) dx_t dt \right) \\ &\quad\overset{(1)}{=} \int_0^1 \int \Big(H(x_t, \nabla s_t(x_t), t) - H(x_t, \nabla \bar{s}_t^*(x_t), t) \Big)\, q_t(x_t) dx_t dt + \int_0^1 \int \frac{\partial q_t(x_t)}{\partial t} \big(- s_t(x_t) + \bar{s}_t^*(x_t) \big) dx_t dt \\ &\quad\overset{(2)}{=} \int_0^1 \int \Big(H(x_t, \nabla s_t(x_t), t) - H(x_t, \nabla \bar{s}_t^*(x_t), t) \Big)\, q_t(x_t) dx_t dt - \int_0^1 \int \big(-\nabla \cdot \big(q_t v_t^*(x_t) \big) \big) \cdot \big(s_t(x_t) - \bar{s}_t^*(x_t) \big) dx_t dt \\ &\quad\overset{(3)}{=} \int_0^1 \int \Big(H(x_t, \nabla s_t(x_t), t) - H(x_t, \nabla \bar{s}_t^*(x_t), t) - \langle v_t^*(x_t), \nabla s_t(x_t) - \nabla \bar{s}_t^*(x_t) \rangle \Big)\, q_t(x_t) dx_t dt \\ &\quad= \int_0^1 \int D_H[\nabla s_t(x_t) : \nabla \bar{s}_t^*(x_t)]\, q_t(x_t) dx_t dt\end{aligned}$$

where in (1) we use the fact that $\int_0^1 \int -\frac{\partial q_t}{\partial t} s_t dx_t dt = \int s_0\, q_0 dx_0 - \int s_1 q_1 dx_1 + \int_0^1 \int \frac{\partial s_t}{\partial t}\, q_t dx_t dt$ by integration by parts. In (2), we use the fact that v_t^* satisfies the continuity equation for q_t from Eq. (50). In (3), we integrate by parts with respect to x and recognize the resulting expression as the definition of the Bregman divergence from Eq. (55). □

B.3. Entropic Action Matching and Entropy-Regularized Optimal Transport

Consider the dynamical formulation of entropy-regularized optimal transport (Léonard, 2014; Chen et al., 2016; 2021), which involves the same kinetic energy minimization as in Eq. (43) but modifying the continuity equation to account for stochasticity,

$$\frac{1}{2} W_\epsilon(p_0, p_1) = \inf_{v_t} \inf_{q_t} \int_0^1 \frac{1}{2} \mathbb{E}_{q_t(x)} \|v_t(x)\|^2 dt, \quad \text{s.t. } \frac{\partial q_t(x)}{\partial t} = -\nabla \cdot \big(q_t(x) v_t(x) \big) + \frac{\sigma_t^2}{2} \Delta q_t(x) \text{ and } q_0 = p_0,\ q_1 = p_1\,. \tag{60}$$

Since we fix the density path q_t in our Action Matching formulation problem, we again omit the optimization over q_t and the marginal constraints,

$$\mathcal{K}_{\text{eAM}} \coloneqq \inf_{v_t} \int_0^1 \frac{1}{2} \mathbb{E}_{q_t(x)} \|v_t(x)\|^2 dt, \quad \text{s.t. } \frac{\partial q_t(x)}{\partial t} = -\nabla \cdot \big(q_t(x) v_t(x) \big) + \frac{\sigma_t^2}{2} \Delta q_t(x) \tag{61}$$

Introducing Lagrange multiplier s_t to enforce the Fokker-Planck constraint leads to

$$\mathcal{K}_{\text{eAM}} = \sup_{s_t} \inf_{v_t} \int_0^1 \frac{1}{2}\mathbb{E}_{q_t(x)}\|v_t(x)\|^2 dt + \int_0^1 \int s_t(x_t)\Big(\frac{\partial q_t}{\partial t}(x_t) + \nabla \cdot \big(q_t(x_t)v_t(x_t)\big) - \frac{\sigma_t^2}{2}\Delta q_t(x)\Big)dx_t dt.$$

Compared to Eq. (46) and Eq. (58), note that the additional $-s_t(x_t)\frac{\sigma_t^2}{2}\Delta q_t(x)$ term does not depend on v_t. Thus, integrating by parts and eliminating v_t as above yields the condition that v_t^* is a gradient field,

$$v_t(x) = \nabla s_t^*(x) \tag{62}$$

Substituting into Eq. (61) and following derivations as in the proof of Prop. 3.2 in App. A.2 yields the entropic AM objective,

$$\mathcal{L}_{\text{eAM}} \coloneqq \inf_{s_t(x)} \int dx\ s_0(x)p_0(x) - \int dx\ s_1(x)p_1(x) + \int_0^1 \int_{\mathcal{X}} dt dx\ q_t(x)\left[\frac{1}{2}\|\nabla s_t\|^2 + \frac{\sigma_t^2}{2}\Delta s_t + \frac{\partial s_t}{\partial t}\right]. \tag{63}$$

Since $\nabla \tilde{s}_t^*$ is the unique gradient field which satisfies the Fokker-Planck equation for the distributional path of q_t (Prop. 3.1), and the solution v_t^* minimizing the kinetic energy in Eq. (61) is a gradient field satisfying the Fokker-Planck equation, we conclude that these vector fields are the same, $v_t^* = \nabla \tilde{s}_t^*$.

B.4. Unbalanced Action Matching and Unbalanced Optimal Transport

To account for growth or destruction of probability mass across time or optimal transport between positive measures with unequal normalization constant, (Chizat et al., 2018a;c; Kondratyev et al., 2016; Liero et al., 2016; 2018) analyze optimal transport problems involving a growth rate $g_t(x)$. In particular, the 'Wasserstein Fisher-Rao' distance (Chizat et al., 2018a) is defined by adding a term involving the norm of the growth rate g_t to the Benamou & Brenier (2000) dynamical OT formulation in Eq. (43) and accounting for the growth term in the modified continuity equation (see Eq. (15)-16)

$$WFR_\lambda(p_0, p_1) \coloneqq \inf_{v_t} \inf_{g_t} \inf_{q_t} \int_0^1 \mathbb{E}_{q_t(x)}\left[\frac{1}{2}\|v_t(x)\|^2 + \frac{\lambda}{2}g_t(x)^2\right]dt, \quad \text{subj. to} \tag{64}$$

$$\frac{\partial q_t(x)}{\partial t} = -\nabla \cdot \big(q_t(x)v_t(x)\big) + \lambda g_t(x)q_t(x)\,, \quad \text{and } q_0 = p_0,\ q_1 = p_1. \tag{65}$$

where the growth term may also be scaled by a multiplier λ. If we again fix the path q_t, as in Action Matching, we define

$$\mathcal{K}^*_{\text{uAM}} + \mathcal{G}^*_{\text{uAM}} \coloneqq \inf_{v_t} \inf_{g_t} \int_0^1 \mathbb{E}_{q_t(x)}\left[\frac{1}{2}\|v_t(x)\|^2 + \frac{\lambda}{2}g_t(x)^2\right]dt \ \text{ s.t. } \frac{\partial q_t(x)}{\partial t} = -\nabla \cdot \big(q_t(x)v_t(x)\big) + \lambda g_t(x)q_t(x) \tag{66}$$

Note that Eq. (64) and Eq. (66) are each convex optimizations after the change of variables $(q_t, v_t, g_t) \mapsto (q_t,\ q_t v_t,\ q_t g_t)$ (Chizat et al., 2018a). We slightly abuse notation to write integrals as expectations with respect to $q_t(x)$, which may not be a normalized probability measure in general (see below, Liero et al. (2016)).

Introducing s_t as a Lagrange multiplier enforcing the modified continuity equation, we obtain the following Lagrangian

$$\begin{aligned}
\mathcal{K}^*_{\text{uAM}} + \mathcal{G}^*_{\text{uAM}} &= \sup_{s_t} \inf_{v_t} \inf_{g_t} \int_0^1 \mathbb{E}_{q_t(x)}\left[\frac{1}{2}\|v_t(x)\|^2 + \frac{\lambda}{2}g_t(x)^2\right]dt + \int_0^1 \int s_t(x)\Big(\frac{\partial q_t}{\partial t}(x) + \nabla \cdot \big(q_t(x)v_t(x)\big) - q_t(x)g_t(x)\Big)dxdt \\
&\overset{(1)}{=} \sup_{s_t} \inf_{v_t} \inf_{g_t} \int_0^1 \int \left[\frac{1}{2}\|v_t(x)\|^2 + \frac{\lambda}{2}g_t(x)^2\right]q_t(x)dxdt + \int s_t(x)q_t(x)dx\Big|_{t=0}^{t=1} - \int_0^1 \int \frac{\partial s_t(x)}{\partial t}q_t(x)dxdt \\
&\quad + \int_0^1 \int \Big[-\lambda s_t(x)\, g_t(x) - \langle v_t(x), \nabla s_t(x)\rangle\Big]\, q_t(x)dx
\end{aligned} \tag{67}$$

where integrate by parts wrt x and t in (1). Finally, eliminating v_t and g_t yields the optimality conditions

$$v_t^*(x) = \nabla s_t(x) \qquad g_t^*(x) = s_t(x). \tag{68}$$

These optimality conditions show that the action function s_t links the problems of transporting mass via the vector field v_t and creating or destroying mass via the growth rate g_t. Substituting back into Eq. (67) and simplifying, we obtain the unbalanced objective $\mathcal{L}_{\text{uAM}}(s_t)$, which is to be minimized as a function of the variational action

$$\mathcal{L}_{\text{uAM}_\lambda}(s_t) := \mathbb{E}_{q_0(x)}[s_0(x)] - \mathbb{E}_{q_1(x)}[s_1(x)] + \int \mathbb{E}_{q_t(x)}\left[\frac{1}{2}\|\nabla s_t(x)\|^2 + \frac{\partial s_t}{\partial t}(x) + \frac{\lambda}{2}s_t^2\right]dt \tag{69}$$

As in Prop. 3.4, we can define an appropriate action gap involving the growth term, with the optimal $\hat{s}_t^*$ evaluating the kinetic energy $\mathcal{K}^*_{\text{uAM}}(\nabla\hat{s}_t^*) := \frac{1}{2}\int_0^1 \mathbb{E}_{q_t(x)}\|\nabla\hat{s}_t^*(x)\|^2 dt$ and growth energy $\mathcal{G}^*_{\text{uAM}}(\hat{s}_t^*) := \frac{\lambda}{2}\int_0^1 \mathbb{E}_{q_t(x)}[\hat{s}_t^*(x)^2]dt$ of a given curve q_t.

Metric Derivative Finally, the uAM objective and Eq. (66) are related to the *metric derivative* in the space of finite positive Borel measures $\mathcal{M}(\mathcal{X})$ equipped with the Wasserstein Fisher-Rao metric distance in Eq. (64), where we consider measures with arbitrary mass due to the effect of the growth term. Liero et al. (2016) Thm. 8.16 and 8.17 are analogous to Ambrosio et al. (2008) Thm. 8.3.1 for this case. In particular, for q_t, q_{t+dt} along an absolutely continuous (AC) curve of positive measures with $\partial_t q_t = -\nabla\cdot(q_t v_t^*) + q_t g_t^*$, we have (Liero et al., 2016)

$$|q_t'|^2 = \left(\lim_{dt\to 0}\frac{WFR_1(q_t, q_{t+dt})}{dt}\right)^2 = \int \left[\|v_t^*(x)\|^2 + \|g_t^*(x)\|^2\right]q_t(x)dx = \int \left[\|\nabla\hat{s}_t^*(x)\|^2 + \|\hat{s}_t^*(x)\|^2\right]q_t(x)dx \tag{70}$$

The optimal uAM objective (Eq. (66)) at each t thus reflects the infinitesimal squared WFR distance along an AC curve.

B.5. Action Matching and Wasserstein Gradient Flows

We can also view dynamics learned by action matching as parameterizing the gradient flow of a time-dependent linear functional on the Wasserstein-2 manifold. This is in contrast JKOnet (Bunne et al., 2022), which is limited to learning gradient flows for class of time-homogenous linear functionals. We give a brief review of concepts related to gradient flows, but refer to e.g. Figalli & Glaudo (2021) Ch. 3-4 for more details. Throughout this section, we let $\mathcal{X} \subseteq \mathbb{R}^d$, consider the space of probability measures $\mathcal{P}_2(\mathcal{X})$ with finite second moment, and identify measures $\mu \in \mathcal{P}_2(\mathcal{X})$ with their densities $d\mu = pdx$.

In order to define a gradient flow, we first need to define an inner product on the tangent space. The seminal work of Otto (2001) defines the desired Riemannian metric on $T\mathcal{P}_2(\mathcal{X})$. Consider two curves $p_t^{(i)} : t \to \mathcal{P}_2(\mathcal{X})$ passing through a point $q \in \mathcal{P}_2(\mathcal{X})$, with $p_0^{(i)} = q$ and tangent vectors $\dot{p_0}^{(1)}, \dot{p_0}^{(2)} \in T_q\mathcal{P}_2(\mathcal{X})$ for $\dot{p}^{(i)} := \frac{\partial p_t^{(i)}}{\partial t}$. Using Theorem 2.1, each curve satisfies a continuity equation for a gradient field $\nabla\psi_t^{(i)}(x)$, e.g. $\dot{p_t}^{(i)}|_{t=0} = -\nabla\cdot\left(p_0^{(i)}\nabla\psi_0^{(i)}\right)$ at $t = 0$. The inner product on $T_q\mathcal{P}_2(\mathcal{X})$ is defined as

$$\langle \dot{p_t}^{(1)}, \dot{p_t}^{(2)}\rangle_q = \int \langle \nabla\psi_0^{(1)}(x), \nabla\psi_0^{(2)}(x)\rangle\, q(x)dx \qquad \left(\text{where} \quad \dot{p_t}^{(i)}|_{t=0} = -\nabla\cdot\left(p_0^{(i)}\,\nabla\psi_0^{(i)}\right)\right) \tag{71}$$

The gradient of a functional $\mathcal{F}[p_t]$ with respect to the Wasserstein-2 metric is defined as the tangent vector for which the inner product yields the directional derivative along a curve $p_t : t \to \mathcal{P}_2(\mathcal{X})$ at $t = 0$

$$\langle \text{grad}_{W_2}\mathcal{F}[p_t], \dot{p_t}\rangle_{p_0} = \frac{d}{dt}\Big|_{t=0}\mathcal{F}[p_t] \tag{72}$$

$$\text{or, more explictly, as:} \qquad \text{grad}_{W_2}\mathcal{F}[p_t] = -\nabla\cdot\left(p_t\,\nabla\frac{\delta\mathcal{F}[p_t]}{\delta p_t}\right), \tag{73}$$

where $\frac{\delta\mathcal{F}[p]}{\delta p}$ is the first variation. For example, the Wasserstein gradient of a *time-dependent* linear functional is given by

$$\mathcal{F}_t[p_t] = \int p_t(x)s_t(x)dx \qquad \text{grad}_{W_2}\mathcal{F}_t[p_t] = -\nabla\cdot\left(p_t\,\nabla s_t\right). \tag{74}$$

A negative gradient flow on the Wasserstein manifold is then given, in either continuous or discrete time, by

$$\frac{\partial p_t}{\partial t} = \text{grad}_{W_2}\mathcal{F}_t[p_t] = -\nabla\cdot\left(p_t\,\nabla s_t\right). \tag{75}$$

Thus, the negative gradient flow of a time-dependent linear functional on $\mathcal{P}_2(\mathcal{X})$ can be modeled using the continuity equation for a vector field ∇s_t. Action Matching can now be viewed as learning the functional $\mathcal{F}_t[p_t] = \int p_t(x)s_t(x)dx$ for which a given density path $p_t : t \to \mathcal{P}_2(\mathcal{X})$ (or q_t, in the main text) traces gradient flow on the Wasserstein manifold $\mathcal{P}_2(\mathcal{X})$.

Comparison with JKONET In discrete time, the gradient flow in Eq. (75) can be written as

$$p_{t+1} = \arg\min_{p} \mathcal{F}[p] + \frac{1}{2\tau} W_2^2(p, p_t). \tag{76}$$

where we write the Wasserstein distance between densities instead of measures. The JKONET method of Bunne et al. (2022) uses the above discrete-time approach to learn a potential function $\mathcal{F}[p]$ which drives the observed sample dynamics. However, they restrict attention to learning the parameters θ of a *time-homogeneous* linear functional $\mathcal{F}[p] = \int p(x)s(x;\theta)dx$ and their optimization methodology is notably more complex than action matching.

C. Generative Modeling in Practice

Algorithm 2 Generative Modeling using Action Matching (In Practice)

Require: dataset $\{x_t^j\}_{j=1}^{N_t},\ x_t^j \sim q_t(x)$, batch-size n
Require: parameteric model $s_t(x,\theta)$, weight schedule $\omega(t)$
  for learning iterations **do**
    get batch of samples from boundaries: $\{x_0^i\}_{i=1}^n \sim q_0(x),\ \{x_1^i\}_{i=1}^n \sim q_1(x)$
    sample times $\{t^i\}_{i=1}^n \sim p(t)$
    get batch of intermediate samples $\{x_{t^i}^i\}_{i=1}^n \sim q_t(x)$
    $\mathrm{L}_i = \left[s_0(x_0^i)\omega(0) - s_1(x_1^i)\omega(1) + \frac{1}{2}\left\|\nabla s_t(x_{t^i}^i)\right\|^2 \omega(t^i) + \frac{\partial s_t(x_{t^i}^i)}{\partial t}\omega(t^i) + s_t(x_{t^i}^i)\frac{\partial \omega(t^i)}{\partial t^i} \right]$
    $\mathrm{L} = \sum_{i=1}^n \frac{1}{p(t^i)}\mathrm{L}_i$
    update the model $\theta \leftarrow \mathrm{Optimizer}(\theta, \nabla_\theta \mathrm{L}_\theta)$
  end for
output trained model $s_t(x,\theta^*)$

In practice, we found that the naive application of Action Matching (Algorithm 1) for complicated dynamics such as image generation might exhibit poor convergence due to the large variance of objective estimate. Moreover, the optimization problem

$$\min_{s_t} \frac{1}{2} \int \mathbb{E}_{q_t(x)} \|\nabla s_t(x) - \nabla s_t^*(x)\|^2 dt \tag{77}$$

might be ill posed due to the singularity of the ground truth vector field ∇s_t^*. This happens when the data distribution q_0 is concentrated close to a low dimensional manifold, and the final distribution q_1 has a much higher intrinsic dimensionality (e.g., Gaussian distributions). In this case, the deterministic velocity vector field must be very large (infinite in the limit), so that it can pull apart the low dimensional manifold to transform it to higher dimensions.

We now discuss an example of this behavior, when the data distribution is a mixture of delta functions. Consider the sampling process

$$x_t = f_t(x_0) + \sigma_t \varepsilon, \quad x_0 \sim \pi(x), \quad \varepsilon \sim \mathcal{N}(x \,|\, 0, 1)\,, \tag{78}$$

where the target distribution is a mixture of delta-functions

$$\pi(x) = \frac{1}{N} \sum_i^N \delta(x - x^i). \tag{79}$$

Denoting the distribution of x_t as $q_t(x)$, we can solve the continuity equation

$$\frac{\partial q_t}{\partial t} = -\nabla \cdot (q_t v_t) \tag{80}$$

analytically (see Appendix D) by finding one of the many possible solutions

$$v_t = \frac{1}{\sum_i q_t^i(x)} \sum_i q_t^i(x) \left[(x - f_t(x^i)) \frac{\partial}{\partial t} \log \sigma_t + \frac{\partial f_t(x^i)}{\partial t} \right], \quad q_t^i(x) = \mathcal{N}(x \,|\, f_t(x^i), \sigma_t^2). \tag{81}$$

Note that v_t is not curl-free in general, and thus is not the solution of action matching. However, it can be written as

$$v_t(x) = \sum_i \frac{q_t^i(x)}{\sum_i q_t^i(x)} \nabla s_t^i(x), \quad \text{where} \quad s_t^i(x) = \frac{1}{2}(x - f_t(x^i))^2 \frac{\partial}{\partial t} \log \sigma_t + \left\langle \frac{\partial f_t(x^i)}{\partial t}, x \right\rangle.$$

Given that the density of Gaussian distributions drop exponentially fast, we can conclude that for small values of t around each x^i, $\frac{q_t^j(x)}{\sum_j q_t^j(x)}$ is close to 1 if $i = j$, and close to 0 if $i \neq j$. Thus, $v_t(x)$ around each x^i can be locally approximated with the curl-free vector field $\nabla s_t^i(x)$. Now suppose $\nabla s_t^*(x)$ is the solution of action matching, i.e., the unique curl-free vector field that solves the continuity equation in every region, including regions around each x^i. Given the uniqueness of curl-free vector fields that solve continuity equation, we can conclude that $\nabla s_t^i(x)$ locally matches $\nabla s_t^*(x)$ around each x^i.

For generative modeling, it's essential that $q_0 = \pi(x)$; hence, $\lim_{t\to 0} \sigma_t = 0$ and $\lim_{t\to 0} f_t(x) = x$. Assuming that σ_t^2 is continuous and differentiable at 0, in the limit, around each x^i, we have

$$\text{for } t \to 0\,, \quad \|\nabla s_t^*(x)\|^2 \propto \frac{1}{\sigma_t^2}, \quad \text{and} \quad \frac{1}{2}\mathbb{E}_{q_t(x)}\|\nabla s_t^*(x)\|^2 \propto \frac{1}{\sigma_t^2}. \tag{82}$$

Thus, the loss can be properly defined only on the interval $t \in (\delta, 1]$, where $\delta > 0$. In practice, we want to set δ as small as possible, i.e., we ideally want to learn s_t on the whole interval $t \in [0, 1]$. We can prevent learning the singularity functions just by re-weighting the objective in time as follows

$$\frac{1}{2}\int \mathbb{E}_{q_t(x)}\|\nabla s_t(x) - \nabla s_t^*(x)\|^2 dt \longrightarrow \frac{1}{2}\int \omega(t)\mathbb{E}_{q_t(x)}\|\nabla s_t(x) - \nabla s_t^*(x)\|^2 dt. \tag{83}$$

To give an example, we can take $\sigma_t = \sqrt{t}$ and $f_t(x) = x\sqrt{1-t}$, then $\omega(t) = (1-t)t^{3/2}$ cancels out the singularities at $t = 0$ and $t = 1$.

The second modification of the original Algorithm 1 is the sampling of time-steps for the estimation of the time integral. Namely, the optimization of Equation (83) is equivalent to the minimization of the following objective

$$\mathcal{L}_{\text{AM}}(s) = \underbrace{\omega(1)\mathbb{E}_{q_1(x)}[s_1(x)] - \omega(0)\mathbb{E}_{q_0(x)}[s_0(x)]}_{\text{weighted action-increment}} \tag{84}$$

$$+ \underbrace{\int_0^1 \mathbb{E}_{q_t(x)}\left[\frac{1}{2}\omega(t)\|\nabla s_t(x)\|^2 + \omega(t)\frac{\partial s_t(x)}{\partial t} + s_t(x)\frac{\partial \omega(t)}{\partial t}\right] dt}_{\text{weighted smoothness}}, \tag{85}$$

which consists of two terms. Estimation of the weighted action-increment involves only sampling from q_0 and q_1, while the weighted smoothness term estimate depends on the distribution of time samples $p(t)$, i.e.,

$$\int_0^1 \underbrace{\frac{p(t)}{p(t)}}_{=1} \mathbb{E}_{q_t(x)}\left[\frac{1}{2}\omega(t)\|\nabla s_t(x)\|^2 + \omega(t)\frac{\partial s_t(x)}{\partial t} + s_t(x)\frac{\partial \omega(t)}{\partial t}\right] dt \tag{86}$$

$$= \mathbb{E}_{t\sim p(t)}\mathbb{E}_{x\sim q_t(x)}\frac{1}{p(t)}\left[\frac{1}{2}\omega(t)\|\nabla s_t(x)\|^2 + \omega(t)\frac{\partial s_t(x)}{\partial t} + s_t(x)\frac{\partial \omega(t)}{\partial t}\right]. \tag{87}$$

Note that $p(t)$ can be viewed as a proposal importance sampling distribution, and thus every choice of it results in an unbiased estimate of the original objective function. Thus, we can design $p(t)$ to reduce the variance of the weighted smoothness term of the objective. In our experiments, we observed that simply taking $p(t)$ proportionally to the standard deviation of the corresponding integrand significantly reduces the variance, i.e.,

$$p(t) \propto \sqrt{\mathbb{E}_{x\sim q_t}(\zeta_t - \mathbb{E}_{x\sim q_t}\zeta_t)^2}, \quad \zeta_t = \frac{1}{2}\omega(t)\|\nabla s_t(x)\|^2 + \omega(t)\frac{\partial s_t(x)}{\partial t} + s_t(x)\frac{\partial \omega(t)}{\partial t}. \tag{88}$$

We implement sampling from this distribution by aggregating the estimated variances throughout the training with exponential moving average, and then followed by linear interpolation between the estimates.

D. Sparse Data Regime

In this section, we find velocity vector fields that satisfy the continuity equation in the case where the data distribution q_0 is a delta function or a mixture of delta functions; and the conditional $k_t(x_t \mid x)$ is a Gaussian distribution.

D.1. Delta Function Data Distribution

We start with the case where the dataset consists only of a single point $x_0 \in \mathbb{R}^d$

$$q_0(x) = \delta(x - x_0), \quad k_t(x_t \mid x) = \mathcal{N}(x_t \mid f_t(x), \sigma_t^2). \tag{89}$$

Then the distribution at time t is

$$q_t(x) = \int dx' \; q_0(x') k_t(x \mid x') = \mathcal{N}(x \mid f_t(x_0), \sigma_t^2). \tag{90}$$

The ground truth vector field v comes from the continuity equation

$$\frac{\partial q_t}{\partial t} = -\nabla \cdot (q_t v) \quad \Longrightarrow \quad \frac{\partial}{\partial t} \log q_t = -\langle \nabla \log q_t, v \rangle - \nabla \cdot (v). \tag{91}$$

For our dynamics, we have

$$\begin{aligned}
\frac{\partial}{\partial t} \log q_t = & \; \frac{\partial}{\partial t} \Big[-\frac{d}{2} \log(2\pi\sigma_t^2) - \frac{1}{2\sigma_t^2} \|x - f_t(x_0)\|^2 \Big] && (92)\\
= & \; -d \frac{\partial}{\partial t} \log \sigma_t + \frac{1}{\sigma_t^2} \|x - f_t(x_0)\|^2 \frac{\partial}{\partial t} \log \sigma_t + \frac{1}{\sigma_t^2} \Big\langle x - f_t(x_0), \frac{\partial f_t(x_0)}{\partial t} \Big\rangle && (93)\\
= & \; -d \frac{\partial}{\partial t} \log \sigma_t + \frac{1}{\sigma_t^2} \Big\langle x - f_t(x_0), (x - f_t(x_0)) \frac{\partial}{\partial t} \log \sigma_t + \frac{\partial f_t(x_0)}{\partial t} \Big\rangle; && (94)\\
\nabla \log q_t = & \; -\frac{1}{\sigma_t^2} (x - f_t(x_0)); && (95)\\
\frac{\partial}{\partial t} \log q_t = & \; -d \frac{\partial}{\partial t} \log \sigma_t - \Big\langle \nabla \log q_t, (x - f_t(x_0)) \frac{\partial}{\partial t} \log \sigma_t + \frac{\partial f_t(x_0)}{\partial t} \Big\rangle. && (96)
\end{aligned}$$

Matching the corresponding terms in the continuity equation, we get

$$v = (x - f_t(x_0)) \frac{\partial}{\partial t} \log \sigma_t + \frac{\partial f_t(x_0)}{\partial t}. \tag{97}$$

We note that since the above vector field is curl-free, it is the unique vector field that the action matching would recover.

D.2. Mixture of Delta Functions Data Distribution

For the mixture of delta-functions, we denote

$$q_0(x) = \frac{1}{N} \sum_i^N \delta(x - x^i), \quad q_t(x) = \frac{1}{N} \sum_i^N q_t^i(x), \quad q_t^i(x) = \mathcal{N}(x \mid f_t(x^i), \sigma_t^2). \tag{98}$$

Due to the linearity of the continuity equation w.r.t. q, we have

$$\sum_i \frac{\partial q_t^i}{\partial t} = \sum_i \nabla \cdot \big(q_t^i v\big) \quad \Longrightarrow \quad \sum_i q_t^i \Big(\frac{\partial}{\partial t} \log q_t^i + \langle \nabla \log q_t^i, v \rangle + \nabla \cdot (v) \Big) = 0. \tag{99}$$

We first solve the equation for $\frac{\partial f_t}{\partial t} = 0$, then for $\frac{\partial}{\partial t} \log \sigma_t = 0$ and join the solutions.

For $\frac{\partial f_t}{\partial t} = 0$, we look for the solution in the following form

$$v_\sigma = \frac{A}{\sum_i q_t^i} \sum_i \nabla q_t^i, \quad q_t^i(x) = \mathcal{N}(x \mid f_t^i(x^i), \sigma_t^2). \tag{100}$$

Then we have

$$\nabla \cdot (v_\sigma) = \left\langle \nabla \frac{A}{\sum_i q_t^i}, \sum_i \nabla q_t^i \right\rangle + \frac{A}{\sum_i q_t^i} \sum_i \nabla^2 q_t^i \tag{101}$$

$$= -\frac{A}{(\sum_i q_t^i)^2} \left\| \sum_i \nabla q_t^i \right\|^2 + \frac{A}{\sum_i q_t^i} \sum_i q_t^i \left[\left\| \nabla \log q_t^i \right\|^2 - \frac{d}{\sigma_t^2} \right], \tag{102}$$

$$(\sum_i q_t^i) \nabla \cdot (v_\sigma) = -\frac{A}{\sum_i q_t^i} \left\| \sum_i \nabla q_t^i \right\|^2 + A \sum_i q_t^i \left[\left\| \nabla \log q_t^i \right\|^2 - \frac{d}{\sigma_t^2} \right], \tag{103}$$

and from (99) we have

$$\sum_i q_t^i \left(-d \frac{\partial}{\partial t} \log \sigma_t + \left\langle \nabla \log q_t^i, v_\sigma + \sigma_t^2 \frac{\partial}{\partial t} \log \sigma_t \nabla \log q_t^i \right\rangle + \nabla \cdot (v_\sigma) \right) = 0. \tag{104}$$

From these two equations we have

$$\sum_i q_t^i \nabla \cdot (v_\sigma) = -\frac{A}{\sum_i q_t^i} \left\| \sum_i \nabla q_t^i \right\|^2 + A \sum_i q_t^i \left[\left\| \nabla \log q_t^i \right\|^2 - \frac{d}{\sigma_t^2} \right] = \tag{105}$$

$$= \sum_i q_t^i \left(d \frac{\partial}{\partial t} \log \sigma_t \right) - \frac{A}{\sum_i q_t^i} \left\| \sum_i \nabla q_t^i \right\|^2 - \sigma_t^2 \frac{\partial}{\partial t} \log \sigma_t \sum_i q_t^i \left\| \nabla \log q_t^i \right\|^2. \tag{106}$$

Thus, we have

$$A = -\sigma_t^2 \frac{\partial}{\partial t} \log \sigma_t. \tag{107}$$

For $\frac{\partial}{\partial t} \log \sigma_t = 0$, we simply check that the solution is

$$v_f = \frac{1}{\sum_i q_t^i} \sum_i q_t^i \frac{\partial f_t(x^i)}{\partial t}. \tag{108}$$

Indeed, the continuity equation turns into

$$\sum_i q_t^i \left(\left\langle \nabla \log q_t^i, v_f - \frac{\partial f_t(x^i)}{\partial t} \right\rangle + \nabla \cdot (v_f) \right) = 0. \tag{109}$$

From the solution and the continuity equation we write $\sum_i q_t^i \nabla \cdot (v_f)$ in two different ways.

$$\sum_i q_t^i \nabla \cdot (v_f) = -\frac{1}{\sum_i q_t^i} \left\langle \sum_i \nabla q_t^i, \sum_i q_t^i \frac{\partial f_t(x^i)}{\partial t} \right\rangle + \sum_i \left\langle \nabla q_t^i, \frac{\partial f_t(x^i)}{\partial t} \right\rangle \tag{110}$$

$$= -\left\langle \sum_i \nabla q_t^i, v_f \right\rangle + \sum_i \left\langle \nabla q_t^i, \frac{\partial f_t(x^i)}{\partial t} \right\rangle \tag{111}$$

Thus, we see that (108) is indeed a solution.

Finally, unifying v_σ and v_f, we have the full solution

$$v = -\left(\frac{\partial}{\partial t} \log \sigma_t \right) \frac{\sigma_t^2}{\sum_i q_t^i} \sum_i \nabla q_t^i + \frac{1}{\sum_i q_t^i} \sum_i q_t^i \frac{\partial f_t(x^i)}{\partial t}, \quad q_t^i(x) = \mathcal{N}(x \mid f_t(x^i), \sigma_t^2), \tag{112}$$

$$v = \frac{1}{\sum_i q_t^i} \sum_i q_t^i \left[(x - f_t(x^i)) \frac{\partial}{\partial t} \log \sigma_t + \frac{\partial f_t(x^i)}{\partial t} \right]. \tag{113}$$

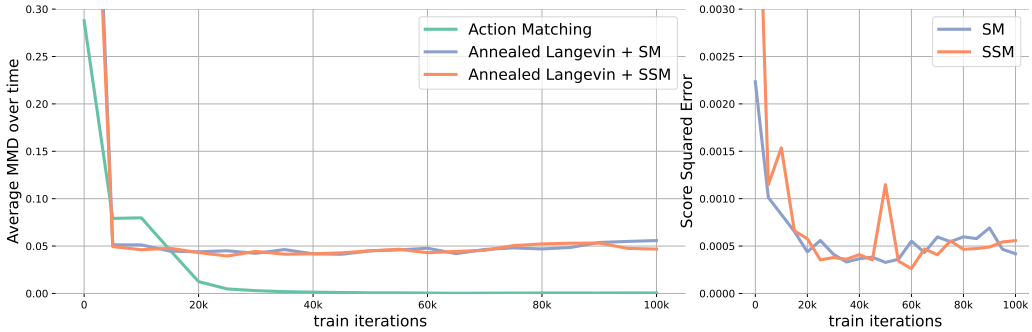

Figure 5. On the left, we demonstrate performance of various algorithms in terms of average MMD over the time of dynamics. The MMD is measured between generated samples and the training data. On the right, we report squared error of the score estimation for the score-based methods.

E. Experiments Details

E.1. Schrödinger Equation Simulation

For the initial state of the dynamics

$$i\frac{\partial}{\partial t}\psi(x,t) = -\frac{1}{\|x\|}\psi(x,t) - \frac{1}{2}\nabla^2\psi(x,t)\,, \tag{114}$$

we take the following wavefunction

$$\psi(x,t=0) \propto \psi_{32-1}(x) + \psi_{210}(x), \quad \text{and} \quad q^*_{t=0}(x) = |\psi(x,t=0)|^2, \tag{115}$$

where n, l, m are quantum numbers and ψ_{nlm} is the eigenstate of the corresponding Hamiltonian (see (Griffiths & Schroeter, 2018)). For all the details on sampling and the exact formulas for the initial state, we refer the reader to the code github.com/necludov/action-matching. We evolve the initial state for $T = 14 \cdot 10^3$ time units in the system $\hbar = 1, m_e = 1, e = 1, \varepsilon_0 = 1$ collecting the dataset of samples from q^*_t. For the time discretization, we take 10^3 steps; hence, we sample every 14 time units.

To evaluate each method, we collect all the generated samples from the distributions $q_t,\ t \in [0,T]$ comparing them with the samples from the training data. For the metric, we measure the Maximum Mean Discrepancy (Gretton et al., 2012) between the generated samples and the training data at 10 different timesteps $t = \frac{k}{10}T,\ k = 1,\dots,10$ and average the distance over the timesteps. For the Annealed Langevin Dynamics, we set the number of intermediate steps for $M = 5$, and select the step size dt by minimizing MMD using the exact scores $\nabla \log q_t(x)$.

For all methods, we use the same architecture, which is a multilayer perceptron with 5 layers 256 hidden units each. The architecture $h(t,x)$ takes $x \in \mathbb{R}^3$ and $t \in \mathbb{R}$ and outputs 3-d vector, i.e. $h(t,x) : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^3$. For the score-based models it already defines the score, while for action matching we use $s_t(x) = \|h(t,x) - x\|^2$ as the model and the vector field is defined as $\nabla s_t(x)$.

In Fig. 5 we plot the convergence of Average MMD over time for Action Matching and the baselines. Despite that both SM and SSM accurately recover the ground truth scores for the marginal distributions (see the right plot in Fig. 5), one cannot efficiently use them for the sampling from the ground truth dynamics.

Algorithm 3 Annealed Langevin Dynamics for the Schrödinger Equation

Require: score model $s_t(x) = \nabla \log q_t(x)$, step size dt, number of intermediate steps M
Require: initial samples $x_0^i \in \mathbb{R}^d$
  for time steps $t \in (0, T]$ **do**
    for intermediate steps $j \in 1, \dots, M$ **do**
      $\varepsilon^i \sim \mathcal{N}(0, \mathbf{1})$
      $x_t^i = x_t^i + \frac{dt}{2} \nabla \log q_t(x_t^i) + \sqrt{dt} \cdot \varepsilon^i$
    end for
    save samples $x_t^i \sim q_t(x)$
  end for
output samples $\{x_t^i\}_{t=0}^T$

Algorithm 4 Annealed Langevin Dynamics for the Image Generation

Require: score model $s_t(x) = \nabla \log q_t(x)$, step size dt, number of intermediate steps M
Require: initial samples $x_0^i \in \mathbb{R}^d$
  for time steps $t \in (0, 1)$ **do**
    $\alpha = \alpha_1 (1 - t)^2$
    for intermediate steps $j \in 1, \dots, M$ **do**
      $\varepsilon^i \sim \mathcal{N}(0, \mathbf{1})$
      $x_t^i = x_t^i + \frac{\alpha}{2} \nabla \log q_t(x_t^i) + \sqrt{\alpha} \cdot \varepsilon^i$
    end for
  end for
  $\alpha = \alpha_1$
  for intermediate steps $j \in 1, \dots, M$ **do**
    $\varepsilon^i \sim \mathcal{N}(0, \mathbf{1})$
    $x_1^i = x_1^i + \frac{\alpha}{2} \nabla \log q_1(x_1^i) + \sqrt{\alpha} \cdot \varepsilon^i$
  end for
output samples x_1^i

E.2. Generative Modeling

For the architecture of the neural network parameterizing s_t, we follow (Salimans & Ho, 2021) with a small modification. Namely, we parameterize $s_t(x)$ as $\langle \text{unet}(t, x), x \rangle$, where $\text{unet}(t, x)$ is the output of the U-net architecture (Ronneberger et al., 2015). For the U-net architecture, we follow (Song et al., 2020b). We consider the same U-net architecture for the baseline to parameterize $\nabla \log q_t$.

For diffusion, we take VP-SDE from (Song et al., 2020b), which corresponds to $\alpha_t = \exp(-\frac{1}{2} \int \beta(s) ds)$ and $\sigma_t = \sqrt{1 - \exp(-\int \beta(s) ds)}$, where $\beta(s) = 0.1 + 19.9t$. All images are normalized to the interval $[-1, 1]$.

For the baseline, we managed to generate the images only taking into account the noise variance of the current distribution q_t as proposed in (Song & Ermon, 2019). For propagating samples in time we select the time step $dt = 10^{-2}$ and perform 10 sampling steps for every q_t. We additionally run 100 sampling steps for the final distribution. In total we run 1000 steps to generate images. See Algorithm 4 for the pseudocode.

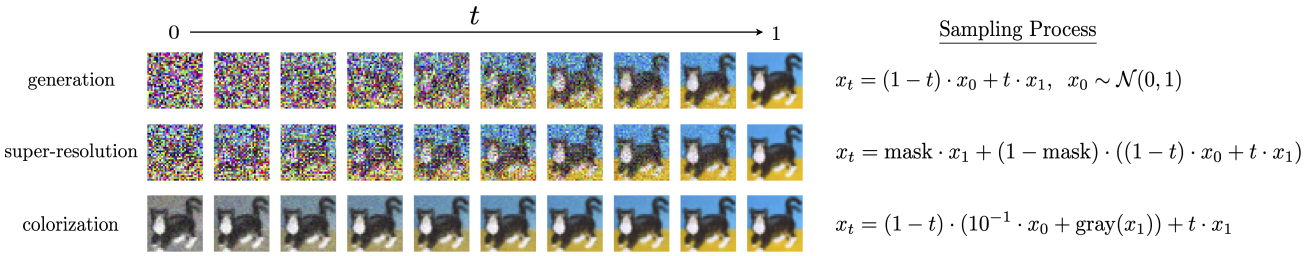

Figure 6. Examples of different noising processes used for different vision tasks. The processes interpolate between the prior distribution at $t = 0$ and the target distribution $t = 1$. For all the processes $x_0 \sim \mathcal{N}(0, 1)$.

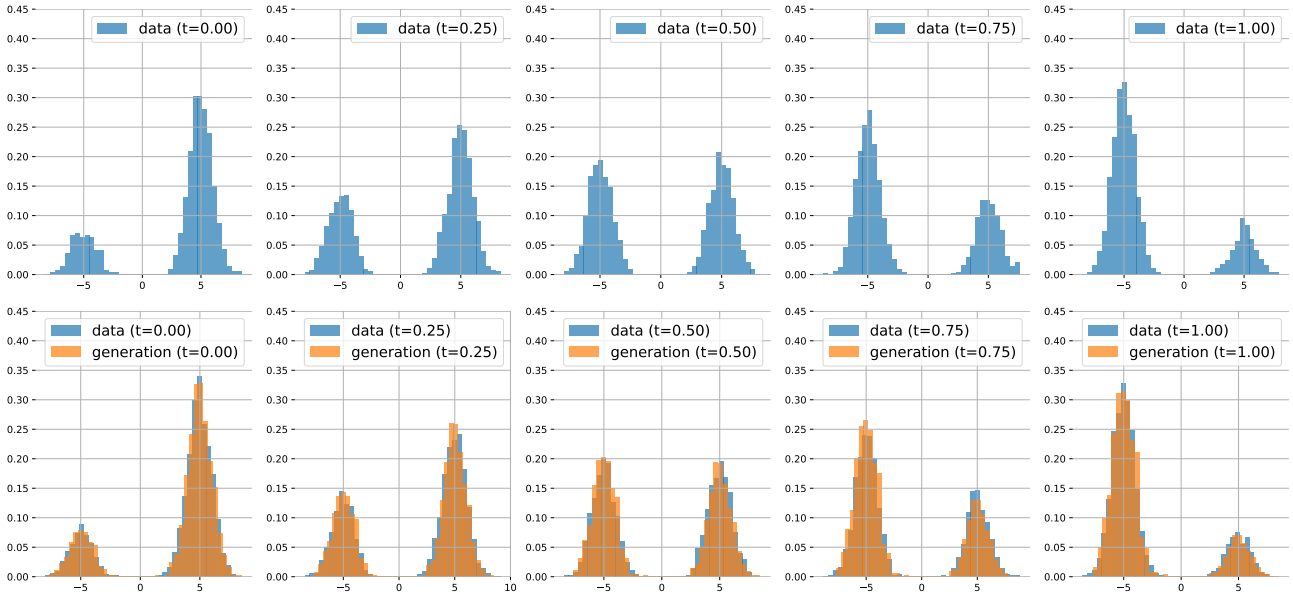

Figure 7. The histograms of the training data (top row) changing in time and the histograms of the generated samples by Unbalanced Action Matching (bottom row).

E.3. Unbalanced Action Matching

We showcase Unbalanced Action Matching on a toy data for which we consider a mixture of gaussians, i.e.,

$$q_t(x) = \alpha_t \mathcal{N}(-5, 1) + (1 - \alpha_t)\mathcal{N}(5, 1)\,, \tag{116}$$

and change α_t linearly from 0.2 to 0.8. In Fig. 7, we demonstrate the data samples and the samples generated by Unbalanced Action Matching starting from the ground truth samples at time $t = 0$ and reweighting particles according to Eq. (19).

Instead of attempting to transport particles from one mode to another, Unbalanced Action Matching is able to model this change of probability mass using the growth term $g_t(x) = s_t(x)$.

F. Generated Images

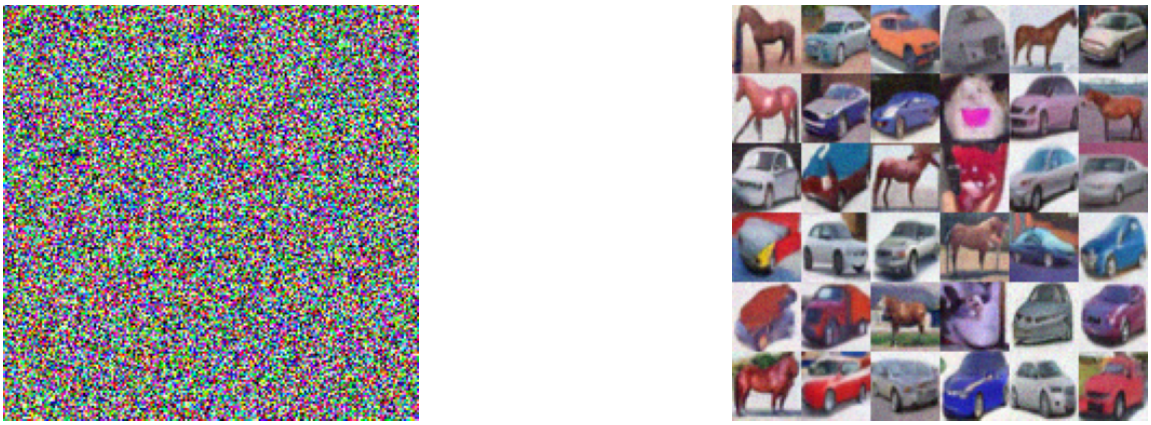

Figure 8. Images generated (right) by the baseline (ALD + SSM) from the noise (left).

Figure 9. Images generated (right) by Action Matching from the noise (left).

Figure 10. Images generated (right) by VP-SDE from the noise (left).

Figure 11. Images generated (right) by Action Matching from the lower resolution images (left).

Figure 12. Images colored (right) by Action Matching from the grayscale images (left).