Free energy Estimation on Any State Space

Jiajun He[1,†] **Zijing Ou**[2] **Francisco Vargas**[3] **Yingzhen Li**[2]
José Miguel Hernández-Lobato[1] **Carles Domingo-Enrich**[4] **Yuanqi Du**[4,†]

[1]University of Cambridge, [2]Imperial College London,
[3]Xaira Therapeutics, [4]Microsoft Research New England
[†]Corresponding to jh2383@cam.ac.uk and yuanqidu@microsoft.com.

Abstract

Free energy estimation is a fundamental yet challenging problem, from physics to statistics. Classical approaches rely on thermodynamic transformations, ranging from direct estimation, quasistatic integration, to finite-time averaging. Recent work [He and Du et al., 2025] learns neural transports to significantly accelerate the efficiency in the finite-time regime. In this paper, we generalize this framework to arbitrary state spaces. Building on this view, we develop a generalized neural transport learning approach for efficient estimation. Experiments validate the effectiveness and efficiency of the proposed method beyond continuous settings, extending to discrete and multimodal spaces as well as autoregressive settings. Beyond free energy estimation, we establish algebraic identities and reveal a group-theoretic structure linking infinitesimal time reversal and generalized Doob's h-transforms, showing that their compositions form a generalized dihedral group.

1 Introduction

Free energy estimation is a fundamental problem across physics, statistics, and machine learning. Defined as the negative logarithm of the normalization constant for an unnormalized density $\tilde{\pi}$,

$$F = -\log Z = -\log \int_{\mathcal{X}} \tilde{\pi}(x)\mathrm{d}x. \tag{1}$$

It plays a central role in characterizing probabilistic models and underlies a wide range of applications from molecular thermodynamics to Bayesian inference [Lelièvre et al., 2010, Tuckerman, 2023].

However, brute force integration of Equation (1) in high dimensions is generally intractable. A standard strategy is to recast the problem as a Monte Carlo (MC) estimation for relative free energy. Let $\tilde{P}$ and $\tilde{Q}$ be two finite, nonnegative measures on $\mathcal{X}$, with normalizing constants $Z_P = \tilde{P}(\mathcal{X}) = \int_{\mathcal{X}} \tilde{P}(\mathrm{d}x)$ and $Z_Q = \tilde{Q}(\mathcal{X}) = \int_{\mathcal{X}} \tilde{Q}(\mathrm{d}x)$, and normalized probability measures $P(\mathrm{d}x) = \tilde{P}(\mathrm{d}x)/Z_P$ and $Q(\mathrm{d}x) = \tilde{Q}(\mathrm{d}x)/Z_Q$. Assuming $\tilde{Q}(\mathrm{d}x) \ll \tilde{P}(\mathrm{d}x)$, the relative free energy is

$$\Delta F = -\log(Z_Q/Z_P) = -\log \mathbb{E}_{X\sim P}\left[\mathrm{d}\tilde{Q}/\mathrm{d}\tilde{P}(X)\right]. \tag{2}$$

Here, $\mathrm{d}\tilde{Q}/\mathrm{d}\tilde{P}$ denotes the Radon-Nikodym derivative between the two unnormalized measures. We keep this formulation measure-theoretic, since the state space $\mathcal{X}$ does not always need to have a Lebesgue measure, and P and Q may therefore not admit a canonical notion of density, while Equation (1) is only the special case where $Z_P = 1$ and Q admit density to Lebesgue measure. This perspective gives rise to a broad family of free energy estimators [Lelièvre et al., 2010], ranging from direct sample reweighting to finite-time path averaging. Representative examples include free energy perturbation (FEP) and the Jarzynski equality (JE), which we discuss in Section A.

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

Unfortunately, both FEP and JE face well-known limitations. Direct reweighting becomes unreliable in high dimensions when the overlap between the reference and target distributions is small, often necessitating a sequence of intermediate distributions to bridge the gap. Finite-time nonequilibrium estimators such as JE typically suffer from high variance unless the underlying process is close to the quasistatic limit. In practice, this favorable regime is only attained when the transformation is infinitely slow, or when special time-reversible structures can be exploited.

To address this challenge, Vaikuntanathan and Jarzynski [2008] introduced an extension of JE based on a controlled transport, now known as the escorted Jarzynski estimator (EJE), to reduce estimation variance. More recently, this idea has attracted renewed attention in machine learning, where computational methods have been developed to learn such escorts efficiently [Vargas et al., 2024, Zhong et al., 2024, Albergo and Vanden-Eijnden, 2024]. These methods seek to construct a transport through a sequence of prescribed intermediate states. Existing approaches mainly differ in how these intermediate states are defined: some specify them through their log densities, while others define them through interpolating paths in sample space. Among these methods, FEAT [He and Du et al., 2025] provides a notable example of learning transport-based finite-time estimators efficiently by interpolating in the sample space. However, these recent advances mostly focus on continuous state spaces with SDE-based transport, while the concept of free energy can appear in any state space. In this paper, we close this gap by the following contributions:

1. To the best of our knowledge, we establish the first explicit escorted Jarzynski equality and corresponding Crooks fluctuation theorem for general Markov processes on arbitrary state spaces, preserving their thermodynamic interpretation (Section 3).
2. We introduce the first general framework for learning bidirectional, time-reversal-consistent stochastic interpolants directly between two endpoint distributions on arbitrary state spaces, enabling efficient neural-transport free-energy estimation beyond Euclidean diffusions (Sections 4 and 5).
3. We empirically verify our proposed framework on different modalities, from discrete to multimodal distributions, and further broaden the notion of non-equilibrium transport to autoregressive models, substantially expanding the scope of these methods beyond traditional thermodynamic and continuous-state settings (Section 7).
4. Beyond free energy estimation, we reveal a broader algebraic structure connecting infinitesimal time reversal and generalized Doob's h-transforms. In particular, any finite composition of these maps reduces to either a single generalized Doob transform or a single infinitesimal reversal. This property exposes an underlying group-theoretic structure: the transformations generated by these two operations form a group isomorphic to a generalized dihedral group (Section C).

2 Background

In this section, we review the nonequilibrium approaches to free energy estimation based on finite-time stochastic transformations between two distributions. Without loss of generality, from now on, we assume the unnormalized density is a Boltzmann distribribution $\tilde{\pi}(x) \propto \exp(-U(x))$ to recover the thermodynamic picture of free energy. We further ignore the Boltzmann constant, i.e. $k_B T = 1$.

Jarzynski Equality and Crooks Fluctuation Theorem. The Jarzynski equality [JE, Jarzynski, 1997] expresses the free energy difference as an exponential average of work:

$$\exp(-\Delta F) = \mathbb{E}_{\overrightarrow{P}}\left[\exp\left(-W(X_{[0,1]})\right)\right], \qquad W(X_{[0,1]}) = \int_0^1 \partial_t U_t(X_t)\,\mathrm{d}t, \tag{3}$$

where $\overrightarrow{P}$ denotes the path measure of a finite-time stochastic process (annealed Langevin) connecting two systems, and W is the accumulated work along the process. The Crooks fluctuation theorem [CFT, Crooks, 1999] generalizes JE by relating the forward and backward processes with work:

$$\mathrm{d}\overleftarrow{Q}/\mathrm{d}\overrightarrow{P}(X_{[0,1]}) = \exp\left(-W(X_{[0,1]}) + \Delta F\right) \tag{4}$$

with JE recovered when taking an expectation on both sides of the equality.

Escorted JE (EJE) and Crooks. To reduce the variance of JE estimators, Vaikuntanathan and Jarzynski [2008] introduced the escorted Jarzynski estimator, which augments the dynamics with an

auxiliary transport v_t, or *escort*, that guides trajectories towards the prefixed intermediate states. The resulting estimator preserves the JE form with a modification to the work,

$$\exp(-\Delta F) = \mathbb{E}_{\vec{P}^v}\left[\exp(-\widetilde{W})\right], \quad \widetilde{W} = \int_0^1 \Big(\partial_t U_t - \nabla \cdot v_t + \nabla U_t \cdot v_t\Big)\mathrm{d}t \tag{5}$$

where $\vec{P}^v$ denotes the escorted forward path measure and $\widetilde{W}$ is a generalization of the work. An analogous Crooks fluctuation theorem also holds, $\mathrm{d}\overleftarrow{Q}^v/\mathrm{d}\vec{P}^v(X_{[0,1]}) = \exp\Big(-\widetilde{W}(X_{[0,1]}) + \Delta F\Big)$.

Generator of Markov Process. Our work mainly uses generators to characterize Markov processes, so we briefly provide their intuition here; see Section B for a more detailed discussion. For a time-inhomogeneous Markov process, the time-dependent generator is the infinitesimal object that determines the local evolution of the process. More precisely, if $P_{s,t}$ denotes the Markov transition operator from time s to t, then for a small time increment $\delta > 0$,

$$P_{t,t+\delta} f(x) = f(x) + \delta\, \mathcal{L}_t f(x) + o(\delta). \tag{6}$$

Thus, $\mathcal{L}_t f(x)$ represents the instantaneous rate of change of the observable f. By adjointness[1], the adjoint operator $\mathcal{L}_t^\dagger$ describes the corresponding infinitesimal evolution of densities.

3 Generalized Escorted Jarzynski and Crooks Fluctuation Theorem

In this section, we present the escorted Jarzynski equality and the Crooks fluctuation theorem for a general Markov process, aiming to estimate the free-energy difference between U_A and U_B. We start with a (time-inhomogeneous) continuous-time Markov process $(X_t)_{t\geq 0}$ with generator $(\mathcal{L}_t)_{t\geq 0}$. Denote the law of X_t as p_t. The following conclusion generalizes the escorted JE for the free energy difference between U_A and U_B (see Theorem B.1 for the technical conditions):

Proposition 3.1 (Generalized Escorted Jarzynski Equality)**.** *Let $(E, \mathcal{B})$ be the state space with reference measure μ, and let $(U_t)_{t\in[0,1]}$ be an energy path interpolating between $U_A =: U_0$ and $U_B =: U_1$. Define*

$$\gamma_t(x) := e^{-U_t(x)}, \qquad Z_t := \int_E \gamma_t(x)\,\mu(\mathrm{d}x), \qquad \pi_t(x) := \gamma_t(x)/Z_t. \tag{7}$$

Let $\mathcal{L}_t^F$ be the generator of a forward continuous-time Markov process, with adjoint $(\mathcal{L}_t^F)^\dagger$. Let $(X_t)_{t\in[0,1]}$ be the corresponding process whose marginal law p_t satisfies

$$\partial_t p_t = (\mathcal{L}_t^F)^\dagger p_t, \qquad p_0 = \pi_A, \qquad \pi_A \propto e^{-U_A}, \tag{8}$$

and denote its path law by $\vec{P}$. Then

$$\Delta F = -\log(Z_B/Z_A) = -\log \mathbb{E}_{\vec{P}}\left[\exp(-\widetilde{W}(X_{[0,1]}))\right], \tag{9}$$

where the "generalized" work is defined as

$$\widetilde{W}(X_{[0,1]}) := \underbrace{\int_0^1 \partial_t U_t(X_t)\,\mathrm{d}t}_{\text{work}} + \underbrace{\int_0^1 \frac{((\mathcal{L}_t^F)^\dagger \gamma_t)(X_t)}{\gamma_t(X_t)}\,\mathrm{d}t}_{\text{escorting}}. \tag{10}$$

Similarly, we can generalize the Crooks fluctuation theorem as follows:

Proposition 3.2 (Generalized Crooks Fluctuation Theorem)**.** *Define the generator and its adjoint for a backward continuous-time Markov process*

$$\mathcal{L}_t^B f := \frac{1}{\gamma_t}(\mathcal{L}_t^F)^\dagger(\gamma_t f) - \frac{1}{\gamma_t}(\mathcal{L}_t^F)^\dagger(\gamma_t)\, f, \qquad (\mathcal{L}_t^B)^\dagger p = \gamma_t\, \mathcal{L}_t^F\left(\frac{p}{\gamma_t}\right) - \frac{p}{\gamma_t}\,(\mathcal{L}_t^F)^\dagger \gamma_t \tag{11}$$

Let $\overleftarrow{Q}$ be the path law of the backward process with marginal law q_t satisfying

$$\partial_t q_t = -(\mathcal{L}_t^B)^\dagger q_t, \qquad q_1 = \pi_B, \qquad \pi_B \propto e^{-U_B}. \tag{12}$$

The RND between the forward ($\vec{P}$ in Equation (8)) and this backward path laws is

$$\frac{\mathrm{d}\overleftarrow{Q}}{\mathrm{d}\vec{P}}(X_{[0,1]}) = \exp\big(-\widetilde{W}(X_{[0,1]}) + \Delta F\big). \tag{13}$$

[1]We consider the adjoint with respect to the Lebesgue measure, $\int (\mathcal{L}_t f)(x)\,\rho(x)\mathrm{d}x = \int f(x)\,(\mathcal{L}_t^\dagger \rho)(x)\mathrm{d}x$.

Remark 3.3. *We note two special cases of Propositions 3.1 and 3.2.*

1. *When the marginal of the forward process satisfies $p_t = \pi_t$, the integrand in Equation (10) becomes constant: $\partial_t U_t(X_t) + \gamma_t^{-1}(X_t)((\mathcal{L}_t^F)^\dagger \gamma_t)(X_t) = \partial_t U_t(X_t) + p_t^{-1}(X_t)\partial_t p_t(X_t) = -\partial_t \log Z_t$. Moreover, the backward process in Equation (11) is exactly the* time-reversal *of the forward process, and its marginal satisfies $q_t = p_t = \pi_t$. This corresponds to the equilibrium setting. Also, when π_t is marginal, Equation (11) recovers Nelson's relation [Nelson, 1967] in generator form.*

2. *When the forward dynamics is stationary at π_t for each t, namely, $(\mathcal{L}_t^F)^\dagger \pi_t = 0$, the escort term vanishes, which reduces to the standard Jarzynski equality. In addition, the backward generator is also stationary at π_t, $(\mathcal{L}_t^B)^\dagger \pi_t = 0$, yielding the standard Crooks fluctuation theorem.*

Additionally, the generalized work in Equation (10) is not specific to a single forward-backward pair. In particular, there are infinitely many forward-backward pairs that give rise to the same functional:

Remark 3.4. *If $\mathcal{K}_t$ is any generator satisfying detailed balance with respect to π_t, namely $\mathcal{K}_t^\dagger(\pi_t f) = \pi_t \mathcal{K}_t f$ for all f, then for any $\alpha \in \mathbb{R}$ we can define modified dynamics by $\partial_t p_t = (\mathcal{L}_t^F + \alpha\mathcal{K}_t)^\dagger p_t$ and $\partial_t q_t = -(\mathcal{L}_t^B + \alpha\mathcal{K}_t)^\dagger q_t$, and the generalized work functional remains unchanged.*

In the following, we present two perspectives that provide intuition for the conclusion above. First, we present the corresponding discrete-time formulation, which makes clear what the escort represents and how the choice of backward kernel should be interpreted. Second, we interpret the result through the generalized Doob's h-transform, which provides a path-weight perspective to sketch the proof.

Generalized EJE and Crooks in Discrete time We now present the discrete-time EJE and Crooks. We state a compact version here and provide the full version in Proposition D.9.

Proposition 3.5 (Discrete-time Jarzynski and Crooks, short version)**.** *Let $\{U_n\}_{n=0}^N$ be an discrete-time energy path, and define $\gamma_n, Z_n, \pi_n(x)$ in the same way as in Proposition 3.1. Assuming the forward processes have transition kernels $\{P_{F,n}\}_{n=1}^N$, the discrete-time generalized work is given by*

$$\widetilde{W}(X_{0:N}) := \sum_{n=1}^{N} \left[U_n(X_n) - U_{n-1}(X_n) + \log \frac{(P_{F,n}\#\pi_{n-1})(X_n)}{\pi_{n-1}(X_n)} \right] \tag{14}$$

Assume the backward process has transition kernels $\{P_{B,n}\}_{n=1}^N$ and

$$\pi_{n-1}(x_{n-1})\, P_{F,n}(x_n \mid x_{n-1}) = (P_{F,n}\#\pi_{n-1})(x_n)\, P_{B,n}(x_{n-1} \mid x_n). \tag{15}$$

Then the Crooks Fluctuation Theorem holds:

$$\frac{\mathrm{d}\overleftarrow{Q}}{\mathrm{d}\overrightarrow{P}}(X_{0:N}) = \frac{\pi_B(x_N) \prod_{n=1}^N P_{B,n}(x_{n-1} \mid x_n)}{\pi_A(x_0) \prod_{n=1}^N P_{F,n}(x_n \mid x_{n-1})} = \exp\big(-\widetilde{W}(X_{0:N}) + \Delta F\big). \tag{16}$$

From this proposition, we see that the generalized work can be interpreted as the *unnormalized density ratio* between the forward and backward path measures, while the backward path is constructed by reversing the forward kernel at each step. We summarize the correspondence between the discrete- and continuous-time settings in Table 1. We also note that the continuous-time escorting term can be obtained by taking a first-order expansion:

$$\log \frac{(P_{F,t+\delta}\#\pi_t)(x)}{\pi_t(x)} = \log \frac{\pi_t(x) + \delta(\mathcal{L}_t^F)^\dagger \pi_t(x) + o(\delta)}{\pi_t(x)} = \delta \frac{(\mathcal{L}_t^F)^\dagger \pi_t(x)}{\pi_t(x)} + o(\delta). \tag{17}$$

The forward-backward relation can be interpreted following the same Taylor argument, which we make explicit in the proof of Proposition D.1 in the Appendix.

Generalized Doob's h-transform and Generalized EJE Generalized Doob's h-transform provides a general perspective and approach to handle path weight. We state generalized Doob's h in Theorem B.8, and connect it with the adjoint and time reversal in Section D. In the following theorem, we establish a connection between the generalized Doob's h-transform and its marginal prescription.

	Crooks Relation	Escort
Discrete time	$\pi_{n-1}P_{F,n} = (P_{F,n}\#\pi_{n-1})P_{B,n}$	$\log \frac{P_{F,n}\#\pi_{n-1}}{\pi_{n-1}}$
Generator form	$\mathcal{L}_t^B f = \gamma_t^{-1}(\mathcal{L}_t^F)^\dagger(\gamma_t f) - \gamma_t^{-1}(\mathcal{L}_t^F)^\dagger(\gamma_t) f$	$\frac{((\mathcal{L}_t^F)^\dagger\gamma_t)(X_t)}{\gamma_t(X_t)}$

Table 1: Discrete-time and generator forms of the escorted relation.

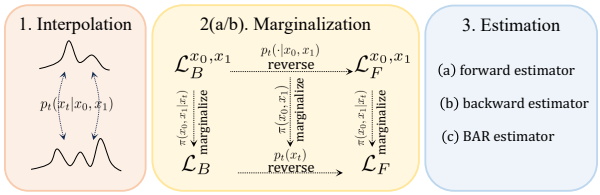


Figure 1: Pipeline of learning the transport.

Theorem 3.6 (Generalized Doob's h-transform and Marginal prescription for it)**.** *Let $(\mathcal{L}_t)_{t\in[0,1]}$ be a family of time-inhomogeneous Markov generators, with adjoints $(\mathcal{L}_t^\dagger)_{t\in[0,1]}$. Let $h_t : E \to (0,\infty)$ be sufficiently regular. The generalized Doob's h-transform generator is given by*

$$\mathcal{L}_t^h f = h_t^{-1}\mathcal{L}_t(h_t f) - h_t^{-1} f\, \mathcal{L}_t h_t. \tag{18}$$

Let π_t be a strictly positive function and define $\rho_t^h = h_t\pi_t$. Then ρ_t^h solves the forward equation of the h-transformed process, i.e.,

$$\partial_t\rho_t^h = (\mathcal{L}_t^h)^\dagger\rho_t^h, \qquad \textbf{\textit{if and only if}} \qquad h_t{}^{-1}\Big(\partial_t h_t + \mathcal{L}_t h_t\Big) = \pi_t^{-1}\Big(\mathcal{L}_t^\dagger\pi_t - \partial_t\pi_t\Big). \tag{19}$$

This result is an extension of results derived in previous work in the scenario of marginal prescription [Chetrite and Touchette, 2013, Denker et al., 2024, Heng et al., 2021]. This proposition allows us to write the unknown marginal density ρ^h in the form of the known predefined π_t. The generalized EJE and Crooks can be derived from this proposition, as we can view the backward process with the h-transformed process with known marginal. We discuss more conclusions and intuition revolving this theorem in detail in Section D.

In summary, this section generalizes the escorted Jarzynski and Crooks in both continuous and discrete time to the general state space. However, although Equation (9) holds for any energy path U_t and any dynamics, in practice, when estimating it using Monte Carlo samples, different choices lead to different statistical efficiencies. More precisely, by Equation (13), we can further define the dissipated work $W_{\text{diss}} := \widetilde{W} - \Delta F$, and $\mathbb{E}_{\vec{P}}[W_{\text{diss}}] = \mathbb{E}_{\vec{P}}\left[\log\frac{\mathrm{d}\vec{P}}{\mathrm{d}\overleftarrow{Q}}\right] = D_{\mathrm{KL}}(\vec{P}\,\|\,\overleftarrow{Q}) \geq 0$, $\mathbb{E}_{\overleftarrow{Q}}[-W_{\text{diss}}] = \mathbb{E}_{\overleftarrow{Q}}\left[\log\frac{\mathrm{d}\overleftarrow{Q}}{\mathrm{d}\vec{P}}\right] = D_{\mathrm{KL}}(\overleftarrow{Q}\,\|\,\vec{P}) \geq 0$, which is *the second law of thermodynamics* in non-equilibrium thermodynamics. In fact, the KL lower bounds the variance of the estimator:

Remark 3.7 (Variance-dissipation Relation)**.** *Denote the free energy estimate as $\widehat{Z} := \widehat{\exp(-\Delta F)} = \frac{1}{N}\sum_{i=1}^N \exp(-\widetilde{W}(X_{[0,1]}^{(i)}))$ with $X_{[0,1]}^{(i)} \sim \vec{P}$. Then*

$$\mathrm{Var}(\widehat{Z}) = \mathcal{O}\left(\tfrac{1}{N}\chi^2(\overleftarrow{Q}\,\|\,\vec{P})\right) \geq \mathcal{O}\left(\tfrac{1}{N}D_{\mathrm{KL}}(\overleftarrow{Q}\,\|\,\vec{P})\right). \tag{20}$$

Therefore, our goal is to reduce the gap between the forward and backward process, and thus improve the variance of the free-energy estimator. Ideally, we want to find dynamics that directly generate samples from $\pi_t \propto e^{-U_t}$ at each time t. Equivalently, we seek forward and backward dynamics that are time-reversals and bridges between U_A and U_B. In this ideal setup, the dissipated work is 0, and the estimate has zero variance. In Section 4, we discuss how to construct such dynamics by learning a neural transport.

Beyond free energy. Infinitesimal reversal and generalized Doob's h-transforms form a generalized dihedral group, with every finite composition reducing to a single reversal or Doob transform. We present these identities and their group-theoretic formulation in Section C.

4 Learning the Transport Models

Propositions 3.1 and 3.2 provide us with two identities for free energy estimation. We can either (1) choose an energy path U_t and a forward process, and calculate work with Equation (10), or (2) choose the forward-backward process pair, and calculate the work with the forward-backward

Radon-Nikodym derivative in Equation (13). Similar to He and Du et al., 2025, we consider the second way to estimate the free energy with a forward-backward pair. This allows us to skip defining or learning the energy path U_t explicitly, yielding a more robust and scalable estimator.

There are also two main paradigms for learning a forward-backward pair: (1) we can specify the energy path U_t first, and learn the forward and backward generators so that the resulting paths follow this energy in optimality; (2) we first obtain samples drawn from the two endpoint distributions, define an interpolating path between samples, and then learn the forward and backward generators that match this interpolant, with the corresponding energy defined only implicitly. While the former allows us to define more thermodynamic-related energy paths, the path may suffer from teleportation [Máté and Fleuret, 2023], and learning can be more challenging and expensive. The latter, on the other hand, is generally more scalable and compatible with recent advances in diffusion and flow models [Song et al., 2020, Lipman et al., 2022, Benton et al., 2024, Holderrieth et al., 2024, Albergo et al., 2025]. Therefore, in the following, we will focus our discussion on the second option.

We start by defining the interpolant as a conditional distribution $x_t \sim p_t(x_t|x_0, x_1)$ given a pair of samples from a coupling of two boundary distributions $(x_0, x_1) \sim \pi(x_0, x_1)$. Examples include

$$\text{Average: } x_t = \alpha_t x_0 + \beta_t x_1 + \gamma_t \epsilon, \quad \epsilon \perp (x_0, x_1), \quad \epsilon \sim \mathcal{N}(0, I) \tag{21}$$

$$\text{Mixture: } x_t \sim \alpha_t \delta_{x_0}(x_t) + \beta_t \delta_{x_1}(x_t), \quad \text{or} \quad x_t \sim \alpha_t \delta_{x_0}(x_t) + \beta_t \delta_{x_1}(x_t) + \gamma_t \delta_M(x_t) \tag{22}$$

We then seek to learn a forward process and a backward process such that, at optimum, (i) both generate the marginal density $p_t(x_t) = \mathbb{E}_{(x_0,x_1)\sim\pi}[p_t(x_t \mid x_0, x_1)]$ for every $t \in [0, 1]$, and (ii) the two processes are time-reversal. Note that (ii) is a stronger requirement than (i).

Step 1: *Define a coupling $\pi(x_0, x_1)$ and an interpolant path $x_t \sim p_t(x_t|x_0, x_1)$, and learn a pair of time-reversal dynamics with marginal $p_t(x_t) = \mathbb{E}_{(x_0,x_1)\sim\pi}[p_t(x_t|x_0, x_1)]$.*

We discuss two frameworks to achieve this in the following. In the first method, we directly learn the forward and backward processes, while in the second method, we learn one direction only and learn the marginal statistics to reverse it. We now explain these two frameworks respectively:

Learning forward and backward processes We extend generator matching [Holderrieth et al., 2024] to the bridge setting, where we learn the transport process directly between two endpoint distributions. Specifically, we first choose a coupling $\pi(x_0, x_1)$ between the two endpoint distributions, and then define conditional forward and backward generators, $\mathcal{L}_{F,t}^{x_0,x_1}$ and $\mathcal{L}_{B,t}^{x_0,x_1}$, associated with the interpolant path introduced in Equations (21) and (22). That is,

$$\partial_t p_t(x_t|x_0, x_1) = \left(\mathcal{L}_{F,t}^{x_0,x_1}\right)^\dagger p_t(x_t|x_0, x_1), \quad \partial_t p_t(x_t|x_0, x_1) = -\left(\mathcal{L}_{B,t}^{x_0,x_1}\right)^\dagger p_t(x_t|x_0, x_1). \tag{23}$$

We then aim to learn the marginal forward and backward generators by averaging the conditional generators over the posterior of the endpoints given the current state:

$$\mathcal{L}_t^F f(x_t) = \mathbb{E}_{x_0,x_1|x_t}\left[\mathcal{L}_{F,t}^{x_0,x_1} f(x_t)\right], \quad \mathcal{L}_t^B f(x_t) = \mathbb{E}_{x_0,x_1|x_t}\left[\mathcal{L}_{B,t}^{x_0,x_1} f(x_t)\right]. \tag{24}$$

One can verify that the marginal density $p_t(x_t) = \mathbb{E}_{(x_0,x_1)\sim\pi}[p_t(x_t|x_0, x_1)]$ evolves according to

$$\partial_t p_t(x_t) = (\mathcal{L}_t^F)^\dagger p_t(x_t), \quad \partial_t p_t(x_t) = -(\mathcal{L}_t^B)^\dagger p_t(x_t). \tag{25}$$

Following Holderrieth et al. [2024], we can parameterize our network for the marginal generator and define a Bregman divergence D as the objective.

However, this workflow has an important caveat. *Even if two generators both satisfy Equation (25), so that they induce the same marginal density at every time t, this does NOT imply that they are time reversals of one another.* Fortunately, the following proposition shows that if we define the conditional generators to be time reversals of each other, then the corresponding marginal generators are also time reversals, which we prove in Section D.8. This saves us from having to verify the marginal construction directly: it suffices to design the conditional generators to be time-reversal pairs, which is fully tractable.

Proposition 4.1 (Marginalization maintains time-reversal)**.** *Let $p_t(x_t|x_0, x_1)$ be a conditional interpolant, and let $\mathcal{L}_{F,t}^{x_0,x_1}$ and $\mathcal{L}_{B,t}^{x_0,x_1}$ be the corresponding forward and backward conditional*

generators. Assume that, for every (x_0, x_1), they are time reversals of each other, i.e.

$$\mathcal{L}_{B,t}^{x_0,x_1} f = \frac{1}{p_t(x_t|x_0,x_1)}(\mathcal{L}_{F,t}^{x_0,x_1})^\dagger(p_t(x_t|x_0,x_1)f) - \frac{1}{p_t(x_t|x_0,x_1)}(\mathcal{L}_{F,t}^{x_0,x_1})^\dagger p_t(x_t|x_0,x_1)\, f \quad (26)$$

Then $\mathcal{L}_{F,t}$ and $\mathcal{L}_{B,t}$ defined by Equation (24) are also time-reversal with respect to $p_t(x_t)$, namely

$$\mathcal{L}_{B,t} f = {p_t}^{-1}\mathcal{L}_{F,t}^\dagger(p_t f) - {p_t}^{-1}\mathcal{L}_{F,t}^\dagger p_t\, f. \quad (27)$$

This proposition leads to the following learning strategy:

Step 2(A): *Define a pair of forward and backward conditional generators which are time-reversal with $p_t(x_t|x_0,x_1)$, and learn to marginalize them.*

Learning one direction plus its marginal statistics Another way to learn the transports is to learn a forward process, plus some statistics that allows we to obtain the backward process from the forward process. From Equation (27), we can see that we only need to learn the marginal p_t[2].

Since we can directly sample from p_t, one could in principle use implicit score matching [ISM; Hyvärinen and Dayan, 2005]; see also Benton et al. [2024] for a generalization to arbitrary state spaces. However, ISM typically requires expensive computations, such as evaluating divergences or summing/integrating over the entire state space. A more practical alternative is denoising score matching [Vincent, 2011, Benton et al., 2024], adapted to our interpolant setting, as discribed next.

Assume we are given an interpolant such that the conditional distribution $p(x_t|x_0,x_1)$ is tractable. For example, one may first construct an interpolating state I_t, such as a linear interpolation $I_t = \alpha_t x_0 + \beta_t x_1$ or a mixture $I_t \sim \alpha_t\delta_{x_0} + \beta_t\delta_{x_1}$, and then apply a tractable corruption process such as Gaussian corruption or random masking, described by $\partial_\tau q_\tau(\cdot|I) = \mathcal{A}q_\tau(\cdot|I)$ with $q_0(\cdot|I) = \delta_I$. Note that the original interpolant $p(x_t|x_0,x_1)$ can then be written as $p(x_t|x_0,x_1) = \int q_{\tau=f(t)}(x_t|I)\, p_t^I(I \mid x_0,x_1)\,\mathrm{d}I$, where p_t^I denotes the law of the uncorrupted interpolant and $f(t)$ controls the amount of corruption at each time t, with $f(0) = f(1) = 0$. Then, the DSM objective is given by

$$\mathcal{J}_{\mathrm{DSM}}(\varphi_t) = \mathbb{E}_{x_t,I_t}\left[\frac{\mathcal{A}(q_\tau(x_t|I_t)/\varphi_t(x_t))}{q_\tau(x_t|I_t)/\varphi_t(x_t)} - \mathcal{A}\log(q_\tau(x_t|I_t)/\varphi_t(x_t))\Big|_{\tau=f(t)}\right] \quad (28)$$

and $\arg\min_{\varphi_t}\mathcal{J}_{\mathrm{DSM}} = p_t$, which leads to the following learning strategy:

Step 2(B): *Learn a forward generator and the marginal statistics to obtain the time-reversal.*

We note that, although both approaches are valid, their practical difficulty can vary substantially across settings. For example, in continuous state spaces, both strategies can be feasible in practice, while directly learning the score is a more natural choice [Albergo et al., 2025]. In contrast, in discrete state spaces, the second strategy of learning the marginal (in the form of the concrete score) can be much more expensive, since constructing the backward generator from the forward and the marginal concrete score generally requires summing over all states unless additional structure is imposed, e.g., local equivariance [Holderrieth et al., 2025].

5 Inference and Estimation

After training the forward and backward transports, we sample trajectories and evaluate the generalized work $\widetilde{W}$ using the forward-backward Radon-Nikodym derivative in Equation (13). As we have two dynamics, we can sample from each of them and obtain two estimators for the free energy: with $X_{[0,1]}^{F,(i)} \sim \overrightarrow{P}$ and $X_{[0,1]}^{B,(j)} \sim \overleftarrow{Q}$, the free energy can be estimated by the following estimators

$$\widehat{\Delta F}_F = -\log\frac{1}{N}\sum_{i=1}^N \exp\left(-\widetilde{W}(X_{[0,1]}^{F,(i)})\right), \quad \widehat{\Delta F}_B = \log\frac{1}{N}\sum_{j=1}^N \exp\left(\widetilde{W}(X_{[0,1]}^{B,(j)})\right). \quad (29)$$

[2]Here, the marginal does not only mean the marginal energy or density. Instead, it can be any sufficient statistics to reverse a process. For example, for Ito's SDE, we only need the score; and for discrete CTMC, we only need the concrete score, i.e., density ratio between two states.

Following the same principle by He and Du et al., 2025, we can combine both directions for a Bennett Acceptance Ratio (BAR)-style estimator [Bennett, 1976]

$$\widehat{\Delta F}_{\mathrm{BAR}}(C) = C + \log \frac{\sum_{j=1}^{N} \phi\Big(-\widetilde{W}(X^{B,(j)}_{[0,1]}) + C\Big)}{\sum_{i=1}^{N} \phi\Big(\widetilde{W}(X^{F,(i)}_{[0,1]}) - C\Big)}, \qquad \phi(z) = \frac{1}{1+\exp(z)}. \tag{30}$$

When $C = \Delta F$, this estimator minimizes the MSE error, thus often referred as *the minimum variance estimator*. In practice, we initialize C by averaging the two one-sided estimates and iterate, setting C to be the current estimate of ΔF, until convergence. This leads to the final step of the pipeline:

Step 3: *Calculate generalized work for both forward and backward trajectories, and obtain the minimal variance estimator using Equation* (29) *or Equation* (30)*.*

6 Related Work

Neural free-energy methods learn invertible maps for targeted FEP [Wirnsberger et al., 2020, Ding and Zhang, 2021], intermediate paths for thermodynamic integration [Máté et al., 2024a], or escorts for nonequilibrium estimators [Vargas et al., 2024, Albergo and Vanden-Eijnden, 2024]. FEAT learns escorts through stochastic interpolation between endpoint samples [He and Du et al., 2025]. We extend this approach to general Markov transports on arbitrary state spaces, using a bidirectional BAR-style estimator [Bennett, 1976]. A full discussion, including BAR/MBAR and complexity results, is given in Section G.

7 Examples and Experiments

In this section, we make the pipeline described above concrete. We provide examples in Euclidean space, discrete space, and autoregressive models. We also illustrate that this framework is compatible with multimodal space and can be applied to transport with momentum.

In this section, we present the main conclusions and experimental results. In Section E, we explain in detail how the general conclusions and learning framework introduced in Sections 3 and 4 lead to the dynamics and the corresponding learning objectives. In Section J, we describe the systems we work on, the experimental settings we choose, and the hyperparameters.

Euclidean State Space Proosition 3.2 in the form of SDE recovers the escorted Langevin dynamics:

$$\mathrm{d}X_t = v_t(X_t)\mathrm{d}t + \sigma_t^2 s_t(X_t)\mathrm{d}t + \sqrt{2}\sigma_t \overrightarrow{\mathrm{d}W_t}, \qquad \mathcal{L}_{F,t} f = (v_t + \sigma_t^2 s_t)\cdot\nabla f + \sigma_t^2 \Delta f \tag{31}$$

$$\mathrm{d}X_t = v_t(X_t)\mathrm{d}t - \sigma_t^2 s_t(X_t)\mathrm{d}t + \sqrt{2}\sigma_t \overleftarrow{\mathrm{d}W_t}, \qquad \mathcal{L}_{B,t} f = (-v_t + \sigma_t^2 s_t)\cdot\nabla f + \sigma_t^2 \Delta f \tag{32}$$

where the score $s = \nabla \log \pi_t$ in Theorem 3.1. Choosing interpolant $x_t \sim \alpha_t x_0 + \beta_t x_1 + \gamma_t \epsilon$, and learn the score and vector field for this path, our framework leads to the standard stochastic interpolant [Albergo et al., 2025] and FEAT [He and Du et al., 2025]. As this is the setup for FEAT and has been verified by He and Du et al., 2025 and Schebek et al. [2025], we do not repeat their experiments here.

Below, we consider cases beyond those studied in previous works. Since our method can be viewed as a generalization of FEAT [He and Du et al., 2025], we keep the same naming convention, prefixing FEAT with the corresponding modality to specify the setting.

Discrete State Space Proosition 3.2 gives us the following forward and backward process:

$$p_t^F(x,y) = \delta_x(y) + R_t^F(x,y)\Delta t + o(\Delta t), \qquad \mathcal{L}_{F,t} f(x) = \sum_y R_t^F(x,y) f(y) \tag{33}$$

$$p_t^B(x,y) = \delta_x(y) + R_t^B(x,y)\Delta t + o(\Delta t), \qquad \mathcal{L}_{B,t} f(x) = \sum_y R_t^B(x,y) f(y) \tag{34}$$

where the forward and backward rate matrices are related by the concrete score of π_t.

Choosing a mixture-style interpolant $x_t \sim \alpha_t \delta_{x_0}(x_t) + \beta_t \delta_{x_1}(x_t) + \gamma_t \delta_M(x_t)$, we can learn both the forward and backward with generator matching. As there exist many different paths that could realize

Table 2: Ising models using CTMC-FEAT. For 15×15–32×32, standard deviations are obtained from three independently trained networks.

Dim D	Method	Estimates ($\Delta F/D \times 10^3$)	Reference value ($\Delta F/D \times 10^3$)
15×15 (225)	DFM	-145.44 ±0.44	-145.41
	GM-no mask	-145.41 ±0.33	
	GM-w. mask	-145.14 ±0.57	
25×25 (625)	DFM	-145.30 ±0.28	-144.88
	GM-no mask	-144.45 ±0.91	
	GM-w. mask	-144.74 ±1.92	
32×32 (1024)	DFM	-143.89 ±0.60	-144.84
	GM-no mask	-143.93 ±0.75	
	GM-w. mask	-144.62 ±1.93	
50×50 (2500)	DFM	-144.93 ±0.23	-144.83

Table 3: Ising models w. AR-FEAT. The standard deviation is obtained by training three networks and running estimators for each.

Dim D	β	AR FEAT ($\Delta F/D \times 10^3$)	Reference value ($\Delta F/D \times 10^3$)
15×15 (225)	0.2	-734.52 ± 0.01	-734.53
	0.4	-879.90 ± 0.03	-879.94
	0.6	-1213.22 ± 0.01	-1213.21
25×25 (625)	0.2	-734.51 ± 0.01	-734.53
	0.4	-879.43 ± 0.02	-879.41
	0.6	-1211.25 ± 0.01	-1211.24
32×32 (1024)	0.2	-734.52 ± 0.04	-734.53
	0.4	-879.40 ± 0.04	-879.37
	0.6	-1210.73 ± 0.01	-1210.81

Figure 2: Estimated and reference free energy difference for Ising fluid with number of particles $N = 55$ and different attraction parameters R. The standard deviation is computed by training the network once but running the estimator 3 times using 2,000 samples.

N	R	Estimated values	Reference values
55	$0.5 \leftrightarrow 1.0$	319.72 ± 0.23	319.84
	$0.5 \leftrightarrow 1.5$	412.64 ± 0.24	412.82
	$0.5 \leftrightarrow 2.0$	455.16 ± 0.16	455.36

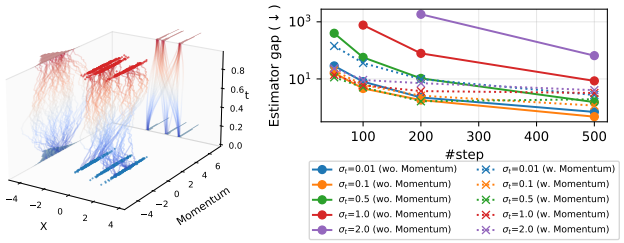


Figure 3: Illustration of Figure 4: Comparing standard SDE transport with momentum and transport with momentum for between two 1D GMMs. free energy of 40D GMMs.

the same interpolant, we have a big design space, as we discuss in Section E.2.3. Here, we consider three variants for learning the generator: (1) discrete flow matching [Gat et al., 2024] without masking (i.e., $\gamma_t = 0$), which we discussed in CTMC Case 1 with loss described in Equations (490) and (491); (2) Generator Matching without masking, which we discussed in CTMC Case 1 with loss described in Equations (495) and (496); and (3) Generator Matching with masking (i.e., $\gamma_t \neq 0, t \in (0, 1)$), which we discussed in CTMC Case 2 with loss described in Equations (516) and (517).

We evaluate these three settings by estimating free energy differences between Ising models with inverse temperature $\beta = 0.2$ and $\beta = 0.4$ across different lattice dimensions. For fair comparisons, all methods employ similar architectures with different hyperparameters (see Section J.1 for details). Reference values are computed analytically using the method in Ferdinand and Fisher [1969]. Results in Table 2 demonstrate that all three methods achieve a relatively reliable estimation of the free energy difference, with the no-mask choices slightly more robust. In Table 9, we further consider more challenging transport settings with $\beta = 0.2$ and $\beta = 0.6$, where the discrepancy between the masking and no-mask choices becomes more pronounced.

We additionally evaluate the DFM variant on a 50×50 lattice (2,500 spins). As shown in Table 2, its estimate remains close to the reference value (-144.93 ± 0.23 versus -144.83). Comparisons with classical free-energy estimators are given in Section L.1.

Autoregressive Transport As discussed in the discrete-time Crooks in Proposition 3.5 and the importance-sampling perspective in Section A, AR models can also be used for free-energy estimation. We here evaluate AR transport on Ising models across different dimensions and inverse temperatures. In contrast to the transports between two target distributions, an AR model transports an empty sequence to a full sequence. Consequently, it estimates the absolute free energy rather than the free energy difference. As shown in Table 3, AR transport achieves very strong performance. We hypothesize that this is because AR training provides a particularly strong learning signal, allowing the model to more easily capture dependencies across positions in the target distribution.

Besides these standard settings, we further demonstrate the flexibility of our framework. In particular, we show that it is compatible with multi-modal state spaces and can also be extended to momentum-augmented dynamics. These additional validations broaden the applicability of our framework.

Multi-modal State Space Another interesting and important case for free-energy estimation is when the system contains more than one modality. Here, we demonstrate the applicability of our framework in this regime. Specifically, we consider a component-wise construction that combines Langevin dynamics for the continuous variables with a CTMC for the discrete variables. The resulting Markov process has a generator as the sum of the two component generators, as detailed in Section F.

We consider an Ising fluid system [Omelyan et al., 2004] with different attraction parameters R. The continuous component represents the particles' coordinates, while the discrete part represents the spin. For simplicity, we learn the transport between a common prior, consisting of a centered Gaussian and a fully masked spin, to each system, and obtain the free energy difference by subtracting the value for each system. See Section J.3 for more details on the system, and Figures 10 and 11 for visualization of the systems. We evaluate $N = 55$ in Figure 2. The reference values are obtained with MBAR. We can see that our method provides a reliable estimate.

Transport with Momentum Besides the flexibility of modality, our framework also supports different transport. Here, we consider an "underdamped" Langevin as the Markov process. In this class of transport, we augment the state space with a momentum space and inject stochasticity into the momentum. In Figure 3, we illustrate this transport between two one-dimensional target distributions. From Proposition 3.2, we obtain the forward and backward process in the following form:

$$\mathrm{d}X_t = Y_t \mathrm{d}t, \qquad \mathrm{d}Y_t = v_t(X_t, Y_t)\mathrm{d}t + \sigma_t^2 s_t(X_t, Y_t)\mathrm{d}t + \sqrt{2}\sigma_t \overrightarrow{\mathrm{d}W_t} \tag{35}$$

$$\mathrm{d}X_t = Y_t \overleftarrow{\mathrm{d}t}, \qquad \mathrm{d}Y_t = v_t(X_t, Y_t)\mathrm{d}t - \sigma_t^2 s_t(X_t, Y_t)\mathrm{d}t + \sqrt{2}\sigma_t \overleftarrow{\mathrm{d}W_t} \tag{36}$$

To learn this transport, as we describe in Section E.2.2, we take a cubic Hermite interpolant:

$$x_t = h_{00}(t)x_0 + h_{10}(t)y_0 + h_{01}(t)x_1 + h_{11}(t)y_1, \quad y_t = \dot{x}_t. \tag{37}$$

with $h_{00}(t) = 2t^3 - 3t^2 + 1, h_{10}(t) = t^3 - 2t^2 + t, h_{01}(t) = -2t^3 + 3t^2, h_{11}(t) = t^3 - t^2$ and objective defined in Equations (429) and (431).

To verify the effectiveness of this method, we evaluate it for the free energy difference between two 40-dimensional Gaussian Mixture models and compare standard SDE transport in FEAT. Since both standard SDE transport and momentum-augmented transport can be run with different choices of $\sigma_t > 0$ using the same learned network, we study their performance across σ_t levels and different numbers of discretization steps. As both yield good estimates of the free energy, we report the gap between the forward and backward estimators in Equation (29). As shown, both achieve comparable performance, while momentum-augmented transport is slightly more robust to larger noises.

8 Conclusion

In this paper, we generalize the escorted Jarzynski equality and the Crooks fluctuation theorem to arbitrary state spaces, and develop a general framework for learning transport models in this setting. Our framework enables the use of generative models for efficient free energy estimation, significantly broadening the class of transport processes for nonequilibrium free energy computation.

Limitations and future works. While efficient and broadly applicable, our framework is still limited by the expressiveness the chosen network parameterization and may require domain-specific tuning in practice. In addition, as discussed in Section I, heavy-tailed target distributions (e.g., Cauchy) can pose difficulties for finite-time transport models. Although such heavy-tailed behavior is less typical for Boltzmann distributions, it may arise in Bayesian inference settings.

Acknowledgment

JH acknowledges support from the University of Cambridge Harding Distinguished Postgraduate Scholar Programme. ZO is supported by the Lee Family Scholarship and would like to thank Alexander Ganose for insightful discussions on potential applications in materials design. JMHL acknowledges funding from AI Hub in Generative Models, under grant EP/Y028805/1.

References

M. Albergo, N. M. Boffi, and E. Vanden-Eijnden. Stochastic interpolants: A unifying framework for flows and diffusions. *Journal of Machine Learning Research*, 26(209):1–80, 2025.

M. S. Albergo and E. Vanden-Eijnden. Nets: A non-equilibrium transport sampler. *arXiv preprint arXiv:2410.02711*, 2024.

C. H. Bennett. Efficient estimation of free energy differences from monte carlo data. *Journal of Computational Physics*, 22(2):245–268, 1976. ISSN 0021-9991. doi: https://doi.org/10.1016/0021-9991(76)90078-4. URL https://www.sciencedirect.com/science/article/pii/0021999176900784.

J. Benton, Y. Shi, V. De Bortoli, G. Deligiannidis, and A. Doucet. From denoising diffusions to denoising markov models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 86(2):286–301, 2024.

D. Blessing, J. Berner, L. Richter, and G. Neumann. Underdamped diffusion bridges with applications to sampling. *arXiv preprint arXiv:2503.01006*, 2025.

T. Chen, J. Gu, L. Dinh, E. A. Theodorou, J. Susskind, and S. Zhai. Generative modeling with phase stochastic bridges. *arXiv preprint arXiv:2310.07805*, 2023a.

T. Chen, G.-h. Liu, M. Tao, and E. A. Theodorou. Deep multi-marginal momentum schrödinger bridge. In *Proceedings of the 37th International Conference on Neural Information Processing Systems*, pages 57058–57086, 2023b.

R. Chetrite and S. Gupta. Two refreshing views of fluctuation theorems through kinematics elements and exponential martingale. *Journal of Statistical Physics*, 143(3):543–584, 2011.

R. Chetrite and H. Touchette. Nonequilibrium microcanonical and canonical ensembles and their equivalence. *Physical review letters*, 111(12):120601, 2013.

R. Chetrite and H. Touchette. Nonequilibrium markov processes conditioned on large deviations. In *Annales Henri Poincaré*, volume 16, pages 2005–2057. Springer, 2015.

G. E. Crooks. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Physical Review E*, 60(3):2721, 1999.

A. Denker, F. Vargas, S. Padhy, K. Didi, S. Mathis, V. Dutordoir, R. Barbano, E. Mathieu, U. J. Komorowska, and P. Lio. Deft: Efficient fine-tuning of diffusion models by learning the generalised h-transform. *Advances in Neural Information Processing Systems*, 37:19636–19682, 2024.

X. Ding and B. Zhang. Deepbar: A fast and exact method for binding free energy computation. *The Journal of Physical Chemistry Letters*, 12(10):2509–2515, 2021. doi: 10.1021/acs.jpclett.1c00189. URL https://doi.org/10.1021/acs.jpclett.1c00189. PMID: 33719449.

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

A. Doucet, W. Grathwohl, A. G. Matthews, and H. Strathmann. Score-based diffusion meets annealed importance sampling. *Advances in Neural Information Processing Systems*, 35:21482–21494, 2022.

W. Du, H. Zhang, T. Yang, and Y. Du. A flexible diffusion model. In *International Conference on Machine Learning*, pages 8678–8696. PMLR, 2023.

A. E. Ferdinand and M. E. Fisher. Bounded and inhomogeneous ising models. i. specific-heat anomaly of a finite lattice. *Physical Review*, 185(2):832, 1969.

I. Gat, T. Remez, N. Shaul, F. Kreuk, R. T. Chen, G. Synnaeve, Y. Adi, and Y. Lipman. Discrete flow matching. *Advances in Neural Information Processing Systems*, 37:133345–133385, 2024.

D. T. Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *The Journal of chemical physics*, 115(4):1716–1733, 2001.

W. Guo, M. Tao, and Y. Chen. Complexity analysis of normalizing constant estimation: from jarzynski equality to annealed importance sampling and beyond. In *International Conference on Learning Representations*, 2026.

A. M. Hahn and H. Then. Characteristic of bennett's acceptance ratio method. *Phys. Rev. E*, 80: 031111, Sep 2009. doi: 10.1103/PhysRevE.80.031111. URL https://link.aps.org/doi/10.1103/PhysRevE.80.031111.

J. He, Y. Du, F. Vargas, Y. Wang, C. P. Gomes, J. M. Hernández-Lobato, and E. Vanden-Eijnden. Feat: Free energy estimators with adaptive transport. *NeurIPS*, 2025.

J. Heng, V. De Bortoli, A. Doucet, and J. Thornton. Simulating diffusion bridges with score matching. *Biometrika*, 112(4):asaf048, 2021.

P. Holderrieth, M. Havasi, J. Yim, N. Shaul, I. Gat, T. Jaakkola, B. Karrer, R. T. Chen, and Y. Lipman. Generator matching: Generative modeling with arbitrary markov processes. *arXiv preprint arXiv:2410.20587*, 2024.

P. Holderrieth, M. S. Albergo, and T. Jaakkola. Leaps: A discrete neural sampler via locally equivariant networks. *arXiv preprint arXiv:2502.10843*, 2025.

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

C. Jarzynski. Nonequilibrium equality for free energy differences. *Physical Review Letters*, 78(14): 2690, 1997.

J. Jo, S. Lee, and S. J. Hwang. Score-based generative modeling of graphs via the system of stochastic differential equations. In *International conference on machine learning*, pages 10362–10383. PMLR, 2022.

T. Lelièvre, M. Rousset, and G. Stoltz. *Free Energy Computations: A Mathematical Perspective*. World Scientific, 2010.

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

A. Lou, C. Meng, and S. Ermon. Discrete diffusion modeling by estimating the ratios of the data distribution. *arXiv preprint arXiv:2310.16834*, 2023.

B. Máté and F. Fleuret. Learning interpolations between boltzmann densities. *arXiv preprint arXiv:2301.07388*, 2023.

B. Máté, F. Fleuret, and T. Bereau. Neural thermodynamic integration: Free energies from energy-based diffusion models. *The Journal of Physical Chemistry Letters*, 15(45):11395–11404, 2024a.

B. Máté, F. Fleuret, and T. Bereau. Solvation free energies from neural thermodynamic integration. *arXiv preprint arXiv:2410.15815*, 2024b.

D. D. Minh and J. D. Chodera. Optimal estimators and asymptotic variances for nonequilibrium path-ensemble averages. *The Journal of chemical physics*, 131(13), 2009.

E. Nelson. *Dynamical Theories of Brownian Motion*. Princeton University Press, 1967. ISBN 9780691079509.

I. Omelyan, I. Mryglod, R. Folk, and W. Fenz. Ising fluids in an external magnetic field: An integral equation approach. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 69(6): 061506, 2004.

Y. Ren, G. M. Rotskoff, and L. Ying. A unified approach to analysis and design of denoising markov models. *arXiv preprint arXiv:2504.01938*, 2025.

K. Rojas, Y. Zhu, S. Zhu, F. X.-F. Ye, and M. Tao. Diffuse everything: Multimodal diffusion models on arbitrary state spaces. *arXiv preprint arXiv:2506.07903*, 2025.

J. L. Rosa-Raíces and D. T. Limmer. Nonadiabatic force matching for alchemical free-energy estimation. *Journal of Chemical Theory and Computation*, 21(22):11455–11462, 2025.

M. Schebek, J. He, E. Hoffmann, Y. Du, F. Noé, and J. Rogal. Assessing generative modeling approaches for free energy estimates in condensed matter. *arXiv preprint arXiv:2512.23930*, 2025.

J. Shi, K. Han, Z. Wang, A. Doucet, and M. Titsias. Simplified and generalized masked diffusion for discrete data. *Advances in neural information processing systems*, 37:103131–103167, 2024.

M. R. Shirts and J. D. Chodera. Statistically optimal analysis of samples from multiple equilibrium states. *The Journal of chemical physics*, 129(12), 2008.

M. R. Shirts, E. Bair, G. Hooker, and V. S. Pande. Equilibrium free energies from nonequilibrium measurements using maximum-likelihood methods. *Physical review letters*, 91(14):140601, 2003.

R. Singhal, M. Goldstein, and R. Ranganath. Where to diffuse, how to diffuse, and how to get back: Automated learning for multivariate diffusions. *arXiv preprint arXiv:2302.07261*, 2023.

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

P. Theodoropoulos, A. D. Saravanos, E. A. Theodorou, and G.-H. Liu. Momentum multi-marginal schr\" odinger bridge matching. *arXiv preprint arXiv:2506.10168*, 2025.

M. E. Tuckerman. *Statistical mechanics: theory and molecular simulation*. Oxford university press, 2023.

S. Vaikuntanathan and C. Jarzynski. Escorted free energy simulations: Improving convergence by reducing dissipation. *Physical Review Letters*, 100(19):190601, 2008.

S. Vaikuntanathan and C. Jarzynski. Escorted free energy simulations. *The Journal of chemical physics*, 134(5), 2011.

A. Van den Oord, N. Kalchbrenner, L. Espeholt, O. Vinyals, A. Graves, et al. Conditional image generation with pixelcnn decoders. *Advances in neural information processing systems*, 29, 2016.

F. Vargas, S. Padhy, D. Blessing, and N. Nüsken. Transport meets variational inference: Controlled monte carlo diffusions. *The Twelfth International Conference on Learning Representations*, 2024.

P. Vincent. A connection between score matching and denoising autoencoders. *Neural computation*, 23(7):1661–1674, 2011.

P. Wirnsberger, A. J. Ballard, G. Papamakarios, S. Abercrombie, S. Racanière, A. Pritzel, D. Jimenez Rezende, and C. Blundell. Targeted free energy estimation via learned mappings. *The Journal of Chemical Physics*, 153(14), 2020.

L. Zhang, P. Potaptchik, J. He, Y. Du, A. Doucet, F. Vargas, H.-D. Dau, and S. Syed. Accelerated parallel tempering via neural transports. *arXiv preprint arXiv:2502.10328*, 2025.

L. Zhao and L. Wang. Bounding free energy difference with flow matching. *Chinese Physics Letters*, 40(12):120201, 2023.

A. Zhong, B. Kuznets-Speck, and M. R. DeWeese. Time-asymmetric fluctuation theorem and efficient free-energy estimation. *Physical Review E*, 110(3):034121, 2024.

Appendix:
Free energy Estimation on Any State Space

A Generalization of Free Energy estimation from Importance Sampling **16**

B Assumptions, Definitions, and Preliminaries **17**

C Beyond Free Energy: Infinitesimal Reversal and Generalized Doob's h **19**

C.1 Algebraic Identities 19

C.2 Group Theory of Reversals and Generalized Doob's h-Transforms 20

D Theorems and Proofs **20**

D.1 Infinitesimal reversal and its connection to adjoint 20

D.2 Infinitesimal reversal and its connection to Generalized Doob's h-transform 23

D.3 Marginal law and generalized Doob's h-transform 26

D.4 Generalized EJE, Crooks and Score Matching Objectives as Corollary of Theorem D.6 and Generalized Doob's h-transform 27

D.5 Proof of Theorem C.4 30

D.6 Discrete-time Generalized EJE and Crooks 30

D.7 Proof of Remark 3.7 (Variance-dissipation Relation) 32

D.8 Proof of Proposition 4.1 (marginalization maintains time-reversal) 33

D.9 Connection between Generator Matching and Denoising Score Matching 34

E Concrete Instantiations and Examples of the Generalized EJE, Crooks and Learning Methods **35**

E.1 Instantiate the Generalized EJE and Crooks 36

E.1.1 Deterministic Flow in Continuous Euclidean Space 36

E.1.2 Diffusion Process in Continuous Euclidean Space 37

E.1.3 Momentum-Augmented Transport in Continuous Space 37

E.1.4 Continuous-Time Markov Chains on Finite State Spaces 40

E.2 Exemplify Learning Methods for the Transport 41

E.2.1 Diffusion Process in Continuous Euclidean Space 42

E.2.2 Momentum-Augmented Transport in Continuous Space 43

E.2.3 Continuous-Time Markov Chains on Finite State Spaces 46

F Composition of Local Generators **54**

G Related Works on Neural Free Energy Estimation **55**

H Simplified Forward and Backward Ratio for Mask Diffusion **56**

I A Note on Heavy-tailed Distributions **57**

J Additional Experimental Details **58**

J.1 CTMC FEAT: Lattice Ising Models 58

J.1.1 System 58

J.1.2 CTMC FEAT Variants 59

J.1.3 Hyperparameters and Settings for CTMC FEAT 59

J.2 AR FEAT: Lattice Ising Models 60

J.2.1 System 60

J.2.2 Hyperparameters and Settings for AR FEAT 60

J.3 Multi FEAT: Ising Fluids 61

J.3.1 System 61

J.3.2 Hyperparameters and Settings 62

J.4 Momentum-Augmented FEAT: Gaussian Mixture Models 62

J.4.1 System 62

J.4.2 Hyperparameters and Settings 63

K Licenses for Existing Assets **63**

L Visualizations and Additional Results **63**

L.1 Comparison with Classical Free-Energy Estimators 63

L.2 Additional Ising Estimator Ablations 63

L.3 CTMC FEAT on Ising Models 65

L.4 AR FEAT on Ising Model 67

L.5 Ising Fluid 68

M Broader Impact **69**

N Declaration of Experiments compute resources **69**

O Declaration of LLM usage **70**

A Generalization of Free Energy estimation from Importance Sampling

We expand the general perspective for estimating free energy difference between two systems in the perspective of importance sampling from Equation (2).

$$\Delta F = -\log(Z_Q/Z_P) = -\log \mathbb{E}_{X\sim P}\left[\mathrm{d}\tilde{Q}/\mathrm{d}\tilde{P}(X)\right]. \tag{38}$$

Free energy perturbation (FEP) directly reweights the proposal samples with the target density

$$\text{FEP:} \quad \mathrm{d}\tilde{Q}/\mathrm{d}x = e^{-U_B}, \mathrm{d}\tilde{P}/\mathrm{d}x = e^{-U_A}, \mathrm{d}P/\mathrm{d}x = \pi_A, \tag{39}$$

$$\Delta F = -\log \mathbb{E}_{X\sim\pi_A}\left[\exp\big(-U_B(X) + U_A(X)\big)\right] \tag{40}$$

Targeted FEP learns an (invertible) transport map T from the proposal distribution to the target distribution, then reweight the transported samples under the target density while taking into account the work induced by the transport

$$\text{Targeted FEP:} \quad \mathrm{d}\tilde{Q}/\mathrm{d}x = e^{-U_B}, \mathrm{d}\tilde{P}/\mathrm{d}x = e^{-U_A}|\det \nabla T|^{-1}, \mathrm{d}P/\mathrm{d}x = \pi_A \tag{41}$$

$$\Delta F = -\log \mathbb{E}_{X\sim\pi_A}\left[\exp\big(-U_B(T(X)) + U_A(X)\big)\,|\det \nabla T(X)|\right] \tag{42}$$

Jarzynski equality and escorted Jarzynski equality augment the proposal with a path measure induced by a finite-time stochastic process, then reweight path samples following the weights determined by the Crooks as in Equation (4)

$$\text{JE and EJE:} \quad \mathrm{d}\tilde{Q} = e^{-U_B(X_1)}\mathrm{d}\overleftarrow{Q}(\cdot|X_1), \;\; \mathrm{d}\tilde{P} = e^{-U_A(X_0)}\mathrm{d}\overrightarrow{P}(\cdot|X_0), \;\; P = \overrightarrow{P} \tag{43}$$

$$\Delta F = -\log \mathbb{E}_{X_{[0,1]}\sim\overrightarrow{P}}\left[e^{-U_B(X_1)}\mathrm{d}\overleftarrow{Q}(\cdot|X_1)/e^{-U_A(X_0)}\mathrm{d}\overrightarrow{P}(\cdot|X_0))\right] \tag{44}$$

while the key difference between JE and EJE is that EJE uses a different stochastic process with *escort* for the transport.

Now we consider autoregressive models (AR) as transport. As AR begins with all empty tokens, we consider the absolute free energy case or the relative free energy between an unnormalized density and a normalized density.

We define the vocabulary set $\mathcal{V}$ and extended vocabulary set $\bar{\mathcal{V}} := \mathcal{V} \cup \{\varnothing\}$ including the empty token. Now we define an extended state space $\mathcal{X}^N := \bar{\mathcal{V}}^N$ with maximum of N tokens. The autoregressive models factorize as

$$p(X_N) = p(X_0)\prod_{i=1}^{N} p(X_i|X_{i-1}) \tag{45}$$

where each state grows from previous state by filling empty tokens from the vocabulary set, i.e. $X_i \in \mathcal{V}^i \times \{\varnothing\}^{N-i} \subset \bar{\mathcal{V}}^N$.

$$\text{AR transport:} \;\; \tilde{Q} = e^{-U_B(X_N)}\prod_{i=1}^{N} q(X_{i-1}|X_i), \;\; \tilde{P} = p(X_0)\prod_{i=1}^{N} p(X_i|X_{i-1}), \;\; P = p \tag{46}$$

$$\Delta F = -\log Z_B = -\log \mathbb{E}_{X_N\sim p}\left[\frac{\exp(-U_B(X_N))}{p(X_0)}\frac{\prod_{i=1}^{N} q(X_{i-1}|X_i)}{\prod_{i=1}^{N} p(X_i|X_{i-1})}\right] \tag{47}$$

Now, we need to consider the choice of the "backward kernel", i.e. $q(X_{i-1}|X_i)$, for AR. It turns out the optimal "backward kernel" is the deterministic deletion kernel, $D_i(x_{0:i}, \varnothing_{i+1:N}) = (x_{0:i-1}, \varnothing_{i:N})$.

$$q(X_{i-1}|X_i) = \mathbf{1}\{X_{i-1} = D_i(X_i)\} \tag{48}$$

This follows from the optimality of the AR model, where the free energy estimator holds not only in expectation but almost surely. It can also be noted by rewriting the free energy estimator as

$$\Delta F = -\log Z_Q = -\log \mathbb{E}_{X_N\sim p}\left[\frac{\exp(-U_B(X_N))}{p(X_N)}\right] \tag{49}$$

where with the optimal AR $Z_B = \exp(-U_B(X_N))/p^*(X_N)$.

B Assumptions, Definitions, and Preliminaries

Remark B.1 (Regularity assumptions for the generalized EJE and Crooks identities)**.** *The continuous-time identities in Propositions 3.1 and 3.2 hold under the following sufficient regularity conditions. Let $(E, \mathcal{B})$ be a measurable state space equipped with a common σ-finite reference measure μ. The intermediate measures $\gamma_t \, \mathrm{d}\mu$ have finite, strictly positive normalizing constants Z_t and densities $\gamma_t > 0$ on the relevant common support. Assume $t \mapsto \log \gamma_t(x)$ is absolutely continuous for μ-almost every x, and that the time-dependent Markov generators and their μ-adjoints are well defined on the functions used in these identities, with well-posed forward and backward evolution equations. We require the generalized-work integrand to be pathwise integrable and its exponential to have finite expectation. For the Crooks path-space identity, we additionally require mutual absolute continuity of the forward and reversed path laws, so that their Radon–Nikodym derivatives are well defined, along with regularity justifying the path-space change of measure. These conditions do not require a Euclidean reference measure and are not intended to be minimal.*

In our paper, we consider transport as a Markov process. The key concept is the Markov semigroup, which we defined below.

Definition B.2 (Markov semigroup)**.** Let $(E, \mathcal{B})$ be a measurable state space, and let $(P_t)_{t\geq 0}$ be a family of linear operators acting on a space $\mathcal{F}$ of bounded measurable functions $f : E \to \mathbb{R}$. We say that $(P_t)_{t\geq 0}$ is a Markov semigroup if, for all $s, t \geq 0$,

1. $P_0 f = f$ for all $f \in \mathcal{F}$;
2. $P_{t+s} = P_t P_s$;
3. $P_t f \geq 0$ whenever $f \geq 0$;
4. $P_t \mathbf{1} = \mathbf{1}$.

The Markov process can often be characterized through its infinitesimal generator, defined below.

Definition B.3 (Infinitesimal generator of a Markov process)**.** Let $(P_t)_{t\geq 0}$ be a Markov semigroup acting on bounded measurable functions $f : E \to \mathbb{R}$. Its infinitesimal generator $\mathcal{L}$ is defined pointwise by

$$\mathcal{L}f(x) := \lim_{t \downarrow 0} \frac{P_t f(x) - f(x)}{t}, \tag{50}$$

for all functions f such that the above limit exists for every $x \in E$.

The same definition applies to time-inhomogeneous Markov processes.

Definition B.4 (Time-inhomogeneous Markov transition operators)**.** Let $(E, \mathcal{B})$ be a measurable state space, and let $(P_{s,t})_{0\leq s\leq t}$ be a family of linear operators acting on a space $\mathcal{F}$ of bounded measurable functions $f : E \to \mathbb{R}$. We say that $(P_{s,t})_{0\leq s\leq t}$ is a time-inhomogeneous Markov transition family if, for all $0 \leq s \leq t \leq u$,

1. $P_{s,s} f = f$ for all $f \in \mathcal{F}$;
2. $P_{s,u} = P_{s,t} P_{t,u}$;
3. $P_{s,t} f \geq 0$ whenever $f \geq 0$;
4. $P_{s,t} \mathbf{1} = \mathbf{1}$.

For the Markov process $(X_t)_{t\in[0,1]}$, the transition operator is interpreted as

$$P_{s,t} f(x) = \mathbb{E}[f(X_t) \mid X_s = x], \qquad 0 \leq s \leq t. \tag{51}$$

Definition B.5 (Time-dependent infinitesimal generator)**.** Let $(P_{s,t})_{0\leq s\leq t}$ be a time-inhomogeneous Markov transition family. Its time-dependent infinitesimal generator $\mathcal{L}_t$ is defined pointwise by

$$\mathcal{L}_t f(x) := \lim_{\delta \downarrow 0} \frac{P_{t,t+\delta} f(x) - f(x)}{\delta}, \tag{52}$$

for all functions f such that the above limit exists.

The generator $\mathcal{L}$ describes how test functions, or observables, change under the Markov transition. By duality, the same Markov evolution induces an evolution on probability measures or densities. This dual evolution is described by the adjoint operator of the generator.

Definition B.6 (Adjoint of a Markov generator)**.** Let $\mathcal{L}$ be the infinitesimal generator of a Markov process on a state space $(E, \mathcal{B})$, acting on a class of test functions. Let μ be a reference measure on E. The adjoint operator $\mathcal{L}^\dagger$ is defined through the duality relation

$$\int_E (\mathcal{L}f)(x)\, g(x)\, \mu(\mathrm{d}x) = \int_E f(x)\, (\mathcal{L}^\dagger g)(x)\, \mu(\mathrm{d}x), \tag{53}$$

for suitable functions f and g for which the above expression is well-defined.

Equivalently, if $\langle f, g\rangle_\mu := \int_E f(x)g(x)\, \mu(\mathrm{d}x)$, then $\mathcal{L}^\dagger$ is characterized by

$$\langle \mathcal{L}f, g\rangle_\mu = \langle f, \mathcal{L}^\dagger g\rangle_\mu. \tag{54}$$

If ρ_t denotes the density of the law of X_t with respect to μ, then we have

$$\partial_t \rho_t = \mathcal{L}^\dagger \rho_t, \tag{55}$$

which is known as the forward Kolmogorov equation. For diffusion processes, this equation is also commonly called the Fokker-Planck equation.

For a time-inhomogeneous Markov process with time-dependent generator $\mathcal{L}_t$, we simply have

$$\partial_t \rho_t = \mathcal{L}_t^\dagger \rho_t. \tag{56}$$

Remark B.7 (Change of reference measure)**.** *The adjoint operator depends on the choice of reference measure. To make this dependence explicit, write $\mathcal{L}^{\dagger,\mu}$ for the adjoint with respect to μ. Suppose that μ is absolutely continuous with respect to the Lebesgue measure and write*

$$\mu(\mathrm{d}x) = m(x)\, \mathrm{d}x. \tag{57}$$

Then the adjoint with respect to μ and the adjoint with respect to the Lebesgue measure are formally related by

$$\mathcal{L}^{\dagger,m} g = \frac{1}{m}\, \mathcal{L}^{\dagger,\mathrm{Leb}}(mg), \tag{58}$$

whenever the expression is well-defined. In particular, $\mathcal{L}^{\dagger,\mu} g$ should be interpreted as the infinitesimal change of the density g with respect to the reference measure μ, not as a normalized density itself.

In this work, unless otherwise stated, we take μ to be the Lebesgue measure and write $\mathcal{L}^\dagger$ for $\mathcal{L}^{\dagger,\mathrm{Leb}}$.

The generalized Doob's h-transform [Chetrite and Gupta, 2011, Chetrite and Touchette, 2013, 2015] allows us to derive the path weight for the tilted path and its generator, which we will use in the following to derive the generalized EJE.

Lemma B.8 (Generalized Doob's h-transform)**.** *Let $(X_t)_{t\in[0,1]}$ be a time-inhomogeneous Markov process on a state space $(E, \mathcal{B})$ with path measure P and time-dependent generator $(\mathcal{L}_t)_{t\in[0,1]}$. Let $h_t : E \to (0, \infty)$ be strictly positive and sufficiently regular. Define*

$$\mathcal{L}_t^h f := h_t^{-1} \mathcal{L}_t(h_t f) - h_t^{-1} f\, \mathcal{L}_t h_t. \tag{59}$$

Assume that $(\mathcal{L}_t^h)_{t\in[0,1]}$ generates a time-inhomogeneous Markov The path measure P^h is absolutely continuous with respect to P, with Radon–Nikodym derivative

$$\frac{\mathrm{d}P^h}{\mathrm{d}P}(x_{(0,1]}|x_0) = \frac{h_1(x_1)}{h_0(x_0)} \exp\left(-\int_0^1 \left[h_t^{-1} \partial_t h_t + h_t^{-1} \mathcal{L}_t h_t \right](x_t)\, \mathrm{d}t \right). \tag{60}$$

The adjoint generator of the h-transformed process is

$$(\mathcal{L}_t^h)^\dagger g = h_t\, \mathcal{L}_t^\dagger \left({h_t}^{-1} g \right) - {h_t}^{-1} \mathcal{L}_t h_t\, g. \tag{61}$$

C Beyond Free Energy: Infinitesimal Reversal and Generalized Doob's h

In this section, we discuss results that go beyond the generalized EJE and Crooks fluctuation theorem, highlighting a connection between a generalized notion of time reversal and generalized Doob's h-transform. These broader results arise from unifying the two perspectives developed in Section 3 for understanding the generalized EJE: the Doob's h-transform perspective and the discrete-time perspective. This unification further admits a more general and abstract algebraic interpretation.

C.1 Algebraic Identities

We first introduce the definition of infinitesimal time reversal.

Definition C.1 (Infinitesimal π_t-reversal)**.** Let $(\mathcal{L}_t)_{t\in[0,1]}$ be a time-inhomogeneous Markov generator, and let $\mathcal{L}_t^\dagger$ denote its adjoint with respect to a reference measure μ. Let π_t be a strictly positive density with respect to μ. The infinitesimal reverse generator with respect to π_t is defined by

$$\overleftarrow{\mathcal{L}}_t^\pi f = {\pi_t}^{-1}\mathcal{L}_t^\dagger(\pi_t f) - {\pi_t}^{-1}\mathcal{L}_t^\dagger \pi_t f. \tag{62}$$

Here, we introduce Equation (62) as a definition directly for simplification. In Proposition D.1, we derive this form and show that $\overleftarrow{\mathcal{L}}_t^\pi$ is exactly the infinitesimal generator associated with the backward operator $\overleftarrow{P}^{\pi_t}_{t+\delta,t}$ which satisfies the following equation:

$$\pi_t(x)\, P_{t,t+\delta}(x,\mathrm{d}y)\,\mu(\mathrm{d}x) = (\pi_t P_{t,t+\delta})(y)\,\mu(\mathrm{d}y)\,\overleftarrow{P}^{\pi_t}_{t+\delta,t}(y,\mathrm{d}x). \tag{63}$$

where $P_{t,t+\delta}$ is the forward operator associated with $\mathcal{L}_t$.

For simplicity, we define the infinitesimal time reversal and the generalized Doob's h-transform stated in Theorem B.8 as two maps on generators. For strictly positive function h_t and π_t, define

$$R_\pi : \mathcal{L}_t \mapsto \overleftarrow{\mathcal{L}}_t^\pi, \quad H_h : \mathcal{L}_t \mapsto \mathcal{L}_t^h, \tag{64}$$

Then, we have the following conclusions:

Theorem C.2 (Algebraic identities for reversal and Doob transforms)**.** *Let p, q, h, h' be strictly positive densities/functions with respect to the reference measure μ. Then, as maps on generators,*

$$R_p \circ R_q = H_{h:=p/q}. \tag{65}$$

In particular, taking $p = q$ gives

$$R_p \circ R_p = H_{h:=1} = I, \tag{66}$$

where I denotes the identity map on generators.
Using the associativity of composition, the reversal and generalized Doob transforms satisfy

$$R_p \circ H_h = R_{p/h}, \quad H_h \circ R_q = R_{qh}, \quad H_h \circ H_{h'} = H_{hh'}. \tag{67}$$

These identities can be abstracted by the following diagram. We drop the subscripts for simplicity.



Here, the nodes represent generators, and the edges represent maps between generators. Since we suppress the subscripts in the diagram, the same symbol denotes the same family of maps rather than the same specific map. By composing maps, we obtain the following general conclusions.

Corollary C.3 (Reduction of finite compositions)**.** *Any finite composition of reversal maps R and generalized Doob transforms H reduces to either a single generalized Doob transform or a single*

reversal map. More precisely, let

$$T = A_n \circ A_{n-1} \circ \cdots \circ A_1, \tag{68}$$

where each A_i is either of the form R_{p_i} or H_{h_i}. Then T can be written as

$$T = \begin{cases} H_g, & \text{if the number of reversal maps appearing in } T \text{ is even}, \\ R_g, & \text{if the number of reversal maps appearing in } T \text{ is odd}, \end{cases} \tag{69}$$

for some strictly positive function g.

C.2 Group Theory of Reversals and Generalized Doob's h-Transforms

The identities above show that the collection of reversal maps and generalized Doob transforms forms a group under composition. Let

$$\mathcal{G} = \{H_h : h > 0\} \cup \{R_p : p > 0\}, \tag{70}$$

Then $(\mathcal{G}, \circ)$ is a group, since:

- **Closure.** Theorem C.2 shows that any composition of two elements of $\mathcal{G}$ is again in $\mathcal{G}$.
- **Associativity.** The operation is a composition of maps on generators, which is associative.
- **Identity.** The identity element is $H_1 = I$.
- **Inverses.** Each element has an inverse: $H_h^{-1} = H_{1/h}, R_p^{-1} = R_p$, as $H_h \circ H_{1/h} = R_p \circ R_p = I$.

The generalized Doob transforms $\{H_h : h > 0\}$ form a subgroup of $\mathcal{G}$, while the reversals $\{R_p : p > 0\}$ do not. Additionally, as proved in Section D.5, $\mathcal{G}$ is isomorphic to the generalized dihedral group, where the generalized h-transform is analogous to rotation and reversal is analogous to reflection:

Proposition C.4 (Generalized dihedral structure)**.** *Let $\mathcal{F}_+$ denote the multiplicative group of strictly positive functions, and let*

$$A = \mathcal{F}_+ / \mathbb{R}_{>0} \tag{71}$$

be the quotient by positive constants. Then $\mathcal{G}$ defined in Equation (70) is isomorphic to the generalized dihedral group

$$\mathcal{G} \cong \mathrm{Dih}(A) = A \rtimes \mathbb{Z}_2, \tag{72}$$

where the nontrivial element of $\mathbb{Z}_2$ acts on A by inversion.

The isomorphism lets us import properties from the generalized dihedral group directly into $\mathcal{G}$.

In summary, this section introduces a general structural view of infinitesimal reversal and generalized Doob's h-transform. By choosing one of the reversal marginals to coincide with the energy path, the abstract identities above recover the generalized EJE and Crooks fluctuation theorem introduced in Section 3. Detailed proofs are provided in Section D.

D Theorems and Proofs

D.1 Infinitesimal reversal and its connection to adjoint

Proposition D.1 (Infinitesimal π_t-reversal)**.** *Let $(\mathcal{L}_t)_{t\in[0,1]}$ be a time-inhomogeneous Markov generator, and let $\mathcal{L}_t^\dagger$ denote its adjoint with respect to a reference measure μ. Let π_t be a strictly positive density with respect to μ. The infinitesimal reverse generator with respect to π_t is*

$$\overleftarrow{\mathcal{L}}_t^{\pi} f = \frac{1}{\pi_t} \mathcal{L}_t^\dagger(\pi_t f) - \frac{\mathcal{L}_t^\dagger \pi_t}{\pi_t} f. \tag{73}$$

Before proving the formula, we clarify what "reversal with respect to π_t" means. Fix a time t and a positive density π_t. We initialize the forward transition from $X_t \sim \pi_t$, evolve one infinitesimal step

using $P_{t,t+\delta}$, and then reverse this one-step joint law by Bayes' rule. That is, we define the reverse kernel $\overleftarrow{P}^{\pi_t}_{t+\delta,t}$ through

$$\pi_t(x)\, P_{t,t+\delta}(x, \mathrm{d}y)\mathrm{d}x = (\pi_t P_{t,t+\delta})(y)\mathrm{d}y \overleftarrow{P}^{\pi_t}_{t+\delta,t}(y, \mathrm{d}x). \tag{74}$$

The infinitesimal π_t-reversal generator is the generator of this reverse kernel, defined by the expansion

$$\overleftarrow{P}^{\pi_t}_{t+\delta,t} f(y) = f(y) + \delta\, \overleftarrow{\mathcal{L}}^{\pi}_t f(y) + o(\delta), \qquad \delta \downarrow 0. \tag{75}$$

When π_t is the actual marginal density of the forward process, this is the "time-reversal" generator. For a general choice of π_t, it is a local "reversal" rather than the time-reversal of the full process.

Proof. Let $P_{t,t+\delta}$ denote the forward transition operator generated by $\mathcal{L}_t$. For small $\delta > 0$, we have the infinitesimal expansion

$$P_{t,t+\delta} g = g + \delta\, \mathcal{L}_t g + o(\delta), \qquad \delta \downarrow 0. \tag{76}$$

Moreover, in the weak sense,

$$\pi_t P_{t,t+\delta} = \pi_t + \delta\, \mathcal{L}^{\dagger}_t \pi_t + o(\delta). \tag{77}$$

By definition, the π_t-reversal kernel $\overleftarrow{P}^{\pi_t}_{t+\delta,t}$ is the reverse kernel of the one-step joint law initialized from $X_t \sim \pi_t$:

$$\pi_t(x) P_{t,t+\delta}(x, \mathrm{d}y)\, \mathrm{d}x = (\pi_t P_{t,t+\delta})(y)\, \mathrm{d}y\, \overleftarrow{P}^{\pi_t}_{t+\delta,t}(y, \mathrm{d}x). \tag{78}$$

Its infinitesimal generator is defined by

$$\overleftarrow{P}^{\pi_t}_{t+\delta,t} f = f + \delta\, \overleftarrow{\mathcal{L}}^{\pi}_t f + o(\delta), \qquad \delta \downarrow 0. \tag{79}$$

We now derive the integral characterization of $\overleftarrow{\mathcal{L}}^{\pi}_t$. Let f and g be test functions. Using the forward representation of the one-step joint law, we have

$$\mathbb{E}[f(X_t) g(X_{t+\delta})] = \int_E f(x)\, P_{t,t+\delta} g(x)\, \pi_t(x)\mathrm{d}x. \tag{80}$$

By (76),

$$\mathbb{E}[f(X_t) g(X_{t+\delta})] = \int_E f\, (g + \delta \mathcal{L}_t g + o(\delta))\, \pi_t \mathrm{d}x \tag{81}$$

$$= \int_E fg\, \pi_t \mathrm{d}x + \delta \int_E f\, \mathcal{L}_t g\, \pi_t \mathrm{d}x + o(\delta). \tag{82}$$

On the other hand, using the reversed representation of the same one-step joint law, we obtain

$$\mathbb{E}[f(X_t) g(X_{t+\delta})] = \int_E g(y)\, \overleftarrow{P}^{\pi_t}_{t+\delta,t} f(y)\, (\pi_t P_{t,t+\delta})(y) \mathrm{d}y. \tag{83}$$

Using (79) and (77), this becomes

$$\mathbb{E}[f(X_t) g(X_{t+\delta})] = \int_E g \left(f + \delta \overleftarrow{\mathcal{L}}^{\pi}_t f + o(\delta)\right) \left(\pi_t + \delta \mathcal{L}^{\dagger}_t \pi_t + o(\delta)\right) \mathrm{d}x \tag{84}$$

$$= \int_E fg\, \pi_t \mathrm{d}x + \delta \int_E g\, \overleftarrow{\mathcal{L}}^{\pi}_t f\, \pi_t \mathrm{d}x + \delta \int_E fg\, \mathcal{L}^{\dagger}_t \pi_t \mathrm{d}x + o(\delta). \tag{85}$$

Since (82) and (85) are expansions of the same quantity, we obtain

$$\int_E g\, \overleftarrow{\mathcal{L}}^{\pi}_t f\, \pi_t \mathrm{d}x = \int_E f\, \mathcal{L}_t g\, \pi_t \mathrm{d}x - \int_E fg\, \mathcal{L}^{\dagger}_t \pi_t \mathrm{d}x. \tag{86}$$

This is the integral characterization of the infinitesimal π_t-reversal.

It remains to identify the operator explicitly. By adjointness with respect to Lesbague measure,

$$\int_E f\,\mathcal{L}_t g\,\pi_t \mathrm{d}x = \int_E \mathcal{L}_t g\,(\pi_t f)\mathrm{d}x = \int_E g\,\mathcal{L}_t^\dagger(\pi_t f)\mathrm{d}x. \tag{87}$$

Substituting this into (86), we get

$$\int_E g\,\overleftarrow{\mathcal{L}}_t^\pi f\,\pi_t \mathrm{d}x = \int_E g\left[\mathcal{L}_t^\dagger(\pi_t f) - f\,\mathcal{L}_t^\dagger \pi_t\right]\mathrm{d}x. \tag{88}$$

Equivalently,

$$\int_E g\,\overleftarrow{\mathcal{L}}_t^\pi f\,\pi_t \mathrm{d}x = \int_E g\left[\frac{1}{\pi_t}\mathcal{L}_t^\dagger(\pi_t f) - \frac{\mathcal{L}_t^\dagger \pi_t}{\pi_t} f\right]\pi_t \mathrm{d}x. \tag{89}$$

Since this holds for all test functions g, we identify

$$\overleftarrow{\mathcal{L}}_t^\pi f = \frac{1}{\pi_t}\mathcal{L}_t^\dagger(\pi_t f) - \frac{\mathcal{L}_t^\dagger \pi_t}{\pi_t} f. \tag{90}$$

This proves the formula. □

Remark D.2. *We note that this infinitesimal reversal has a very similar form to the adjoint. They only differ in an extra term. In fact, this is because they have a deeper connection: The infinitesimal reversal is simply a normalized version of π-weighted adjoint. The first term in Equation (73), i.e.,*

$$\frac{1}{\pi_t}\mathcal{L}_t^\dagger(\pi_t f) \tag{91}$$

comes directly from the π_t-adjoint of the forward transition. It is not yet a Markov generator because it does not necessarily send $\mathbf{1}$ to zero. The second term is exactly the normalization correction coming from importance-normalizing the reverse kernel:

$$\frac{1}{\pi_t}\mathcal{L}_t^\dagger(\pi_t) f = \frac{1}{\pi_t}\mathcal{L}_t^\dagger(\pi_t \mathbf{1}) f \tag{92}$$

We can also obtain a derivation using this normalization perspective to prove Proposition D.1:

Proof. Fix a small $\delta > 0$, we start from:

$$\pi_t(x)\,\mathrm{d}x\,P_{t,t+\delta}(x,\mathrm{d}y) = (\pi_t P_{t,t+\delta})(y)\,\mathrm{d}y\,\overleftarrow{P}^{\pi_t}_{t+\delta,t}(y,\mathrm{d}x). \tag{93}$$

Therefore, for test functions f and g,

$$\int_E f(x) P_{t,t+\delta} g(x)\,\pi_t(x)\,\mathrm{d}x = \int_E g(y)\,\overleftarrow{P}^{\pi_t}_{t+\delta,t} f(y)\,(\pi_t P_{t,t+\delta})(y)\,\mathrm{d}y. \tag{94}$$

Applying importance sampling to the RHS with proposal $\pi_t(y)\,\mathrm{d}y$, we obtain

$$\int_E f(x) P_{t,t+\delta} g(x)\,\pi_t(x)\,\mathrm{d}x = \int_E g(y)\,Q^{\pi_t}_{t,\delta} f(y)\,\pi_t(y)\,\mathrm{d}y, \tag{95}$$

where

$$Q^{\pi_t}_{t,\delta} f(y) := \overleftarrow{P}^{\pi_t}_{t+\delta,t} f(y)\frac{(\pi_t P_{t,t+\delta})(y)}{\pi_t(y)}. \tag{96}$$

From Equation (95), we can see that $Q^{\pi_t}_{t,\delta}$ is the adjoint of the forward transition with respect to the measure $\pi_t(x)\,\mathrm{d}x$. This becomes explicit by using the infinitesimal expansions

$$P_{t,t+\delta} g = g + \delta\,\mathcal{L}_t g + o(\delta), \tag{97}$$

and

$$Q^{\pi_t}_{t,\delta} f = f + \delta\,\mathcal{L}_t^Q f + o(\delta). \tag{98}$$

Substituting these expansions into Equation (95) gives

$$\int_E f(x)\mathcal{L}_t g(x)\pi_t(x)\,\mathrm{d}x = \int_E g(x)\mathcal{L}_t^Q f(x)\pi_t(x)\,\mathrm{d}x. \tag{99}$$

Equivalently,

$$\langle f, \mathcal{L}_t g\rangle_{\pi_t} = \langle \mathcal{L}_t^Q f, g\rangle_{\pi_t}. \tag{100}$$

Thus $\mathcal{L}_t^Q$ is the adjoint of $\mathcal{L}_t$ with respect to the reference measure $\pi_t(x)\,\mathrm{d}x$. Using the change-of-reference formula in Theorem B.7, we obtain

$$\mathcal{L}_t^Q f = \mathcal{L}_t^{\dagger,\pi_t} f = \frac{1}{\pi_t}\mathcal{L}_t^\dagger(\pi_t f), \tag{101}$$

where $\mathcal{L}_t^\dagger$ denotes the adjoint with respect to Lebesgue measure.

Recall that

$$Q_{t,\delta}^{\pi_t} f(y) := \overleftarrow{P}_{t+\delta,t}^{\pi_t} f(y)\frac{(\pi_t P_{t,t+\delta})(y)}{\pi_t(y)}, \tag{102}$$

i.e., $Q_{t,\delta}^{\pi_t}$ is an importance-weighted version of the reverse kernel. Since the backward kernel is normalized, i.e., $\overleftarrow{P}_{t+\delta,t}^{\pi_t}\mathbf{1} = \mathbf{1}$, the weight

$$\frac{(\pi_t P_{t,t+\delta})(y)}{\pi_t(y)} = Q_{t,\delta}^{\pi_t}\mathbf{1}(y) \tag{103}$$

Hence, the actual reverse kernel is recovered by dividing out this normalizing factor:

$$\overleftarrow{P}_{t+\delta,t}^{\pi_t} f = \frac{Q_{t,\delta}^{\pi_t} f}{Q_{t,\delta}^{\pi_t}\mathbf{1}}. \tag{104}$$

Therefore,

$$\overleftarrow{P}_{t+\delta,t}^{\pi_t} f = \frac{f + \delta\,\mathcal{L}_t^Q f + o(\delta)}{1 + \delta\,\mathcal{L}_t^Q\mathbf{1} + o(\delta)} \tag{105}$$

$$= f + \delta\left(\mathcal{L}_t^Q f - f\,\mathcal{L}_t^Q\mathbf{1}\right) + o(\delta). \tag{106}$$

The second equation is due to the Taylor expension of $1/(1+x)$ at $x = 0$:

$$\frac{1}{1 + \delta a + o(\delta)} = 1 - \delta a + o(\delta) \tag{107}$$

and then drop higher-order terms in δ. Hence the infinitesimal generator of the normalized reverse kernel is

$$\overleftarrow{\mathcal{L}}_t^\pi f = \mathcal{L}_t^Q f - f\,\mathcal{L}_t^Q\mathbf{1} \tag{108}$$

$$= \frac{1}{\pi_t}\mathcal{L}_t^\dagger(\pi_t f) - \frac{\mathcal{L}_t^\dagger\pi_t}{\pi_t} f. \tag{109}$$

□

D.2 Infinitesimal reversal and its connection to Generalized Doob's h-transform

The infinitesimal reversal has a deep connection to the generalized Doob's h-transform, as we stated in Section C. Here, we state a more detailed version.

Theorem D.3 (Double infinitesimal reversal and generalized Doob's h-transform)**.** *Let $(\mathcal{L}_t)_{t\in[0,1]}$ be a family of time-inhomogeneous Markov generators, and let p_t and q_t be strictly positive densities. Apply the infinitesimal reversal formula in Equation (73): first define the infinitesimal*

reversal of $\mathcal{L}_t$ with respect to q_t by

$$\widetilde{\mathcal{L}}_t f := \overleftarrow{\mathcal{L}}_t^{q_t} f = \frac{1}{q_t}\mathcal{L}_t^\dagger(q_t f) - \frac{\mathcal{L}_t^\dagger q_t}{q_t} f. \tag{110}$$

Then define the infinitesimal reversal of $\widetilde{\mathcal{L}}_t$ with respect to p_t by

$$\widehat{\mathcal{L}}_t f := \overleftarrow{\widetilde{\mathcal{L}}}_t^{p_t} f = \frac{1}{p_t}\widetilde{\mathcal{L}}_t^\dagger(p_t f) - \frac{\widetilde{\mathcal{L}}_t^\dagger p_t}{p_t} f. \tag{111}$$

Then $\widehat{\mathcal{L}}_t$ is the generator of the Doob's h-transform of $\mathcal{L}_t$ with

$$h_t = \frac{p_t}{q_t}. \tag{112}$$

That is,

$$\widehat{\mathcal{L}}_t f = \mathcal{L}_t^h f = h_t^{-1}\mathcal{L}_t(h_t f) - h_t^{-1} f\, \mathcal{L}_t h_t. \tag{113}$$

A corollary of this, which provides a different perspective to understand this result, is stated below:

Corollary D.4 (Reversal of an h-transform)**.** *Let $(\mathcal{L}_t)_{t\in[0,1]}$ be a family of time-inhomogeneous Markov generators, and let $h_t : E \to (0,\infty)$ be strictly positive. Define the generalized Doob's h-transform generator by*

$$\mathcal{L}_t^h f := h_t^{-1}\mathcal{L}_t(h_t f) - h_t^{-1} f\, \mathcal{L}_t h_t. \tag{114}$$

Let p_t be a strictly positive density, and define

$$q_t = \frac{p_t}{h_t}. \tag{115}$$

Then the infinitesimal reversal of $\mathcal{L}_t^h$ with respect to p_t is equal to the infinitesimal reversal of $\mathcal{L}_t$ with respect to q_t:

$$\overleftarrow{(\mathcal{L}_t^h)}^{p_t} f = \overleftarrow{\mathcal{L}}_t^{q_t} f. \tag{116}$$

We first prove Theorem D.3.

Proof. The proof follows directly from algebra. Fix $t \in [0,1]$ and suppress the subscript t to simplify notation. Write

$$\widetilde{\mathcal{L}} f = \frac{1}{q}\mathcal{L}^\dagger(qf) - \frac{\mathcal{L}^\dagger q}{q} f. \tag{117}$$

We first compute the adjoint of $\widetilde{\mathcal{L}}$. For suitable test functions f and g,

$$\int_E (\widetilde{\mathcal{L}} f)(x) g(x)\,\mathrm{d}x = \int_E \frac{1}{q(x)}\mathcal{L}^\dagger(qf)(x) g(x)\,\mathrm{d}x - \int_E \frac{\mathcal{L}^\dagger q(x)}{q(x)} f(x) g(x)\,\mathrm{d}x \tag{118}$$

$$= \int_E f(x)\left[q\,\mathcal{L}\left(\frac{g}{q}\right) - \frac{\mathcal{L}^\dagger q}{q} g\right](x)\,\mathrm{d}x. \tag{119}$$

Therefore,

$$\widetilde{\mathcal{L}}^\dagger g = q\,\mathcal{L}\left(\frac{g}{q}\right) - \frac{\mathcal{L}^\dagger q}{q} g. \tag{120}$$

Now apply the second reversal with respect to p:

$$\widehat{\mathcal{L}} f = \frac{1}{p}\widetilde{\mathcal{L}}^\dagger(pf) - \frac{\widetilde{\mathcal{L}}^\dagger p}{p} f. \tag{121}$$

Using Equation (120), we have

$$\frac{1}{p}\widetilde{\mathcal{L}}^{\dagger}(pf) = \frac{1}{p}\left[q\,\mathcal{L}\left(\frac{pf}{q}\right) - \frac{\mathcal{L}^{\dagger}q}{q}pf\right] \tag{122}$$

$$= \frac{q}{p}\mathcal{L}\left(\frac{p}{q}f\right) - \frac{\mathcal{L}^{\dagger}q}{q}f, \tag{123}$$

and

$$\frac{\widetilde{\mathcal{L}}^{\dagger}p}{p}f = \frac{1}{p}\left[q\,\mathcal{L}\left(\frac{p}{q}\right) - \frac{\mathcal{L}^{\dagger}q}{q}p\right]f \tag{124}$$

$$= \frac{q}{p}\mathcal{L}\left(\frac{p}{q}\right)f - \frac{\mathcal{L}^{\dagger}q}{q}f. \tag{125}$$

Therefore, we have

$$\widehat{\mathcal{L}}f = \frac{q}{p}\mathcal{L}\left(\frac{p}{q}f\right) - \frac{\mathcal{L}^{\dagger}q}{q}f - \left[\frac{q}{p}\mathcal{L}\left(\frac{p}{q}\right)f - \frac{\mathcal{L}^{\dagger}q}{q}f\right] \tag{126}$$

$$= \frac{q}{p}\mathcal{L}\left(\frac{p}{q}f\right) - \frac{q}{p}f\,\mathcal{L}\left(\frac{p}{q}\right). \tag{127}$$

Finally, using

$$h = \frac{p}{q} \tag{128}$$

we obtain

$$\widehat{\mathcal{L}}_t f = h_t^{-1}\mathcal{L}_t(h_t f) - h_t^{-1} f\,\mathcal{L}_t h_t = \mathcal{L}_t^h f. \tag{129}$$

□

This provides us with a very powerful tool to analyze infinitesimal reversal and derive path weight using generalized Doob's h weight in Theorem B.8. We also have the following corollary.

We then derive the Corollary:

Proof. The proof is twofold. First, applying infinitesimal reversal twice with respect to the same density gives back the original generator. This follows from Theorem D.3 by taking $p_t = q_t$, so that

$$h_t = \frac{p_t}{q_t} = 1. \tag{130}$$

By Theorem D.3, applying reversal first with respect to q_t and then with respect to p_t gives the h-transformed generator $\mathcal{L}_t^h$. Therefore, reversing $\mathcal{L}_t^h$ once more with respect to p_t cancels the second reversal and leaves the reversal of $\mathcal{L}_t$ with respect to q_t. Hence

$$\overleftarrow{(\mathcal{L}_t^h)}^{p_t} f = \overleftarrow{\mathcal{L}}_t^{q_t} f. \tag{131}$$

□

Another corollary is stated below:

Corollary D.5 (Doob's relation between two reversals)**.** *Let $(\mathcal{L}_t)_{t\in[0,1]}$ be a family of time-inhomogeneous Markov generators, and p_t and q_t be a strictly positive density. Then the infinitesimal reversal of $\mathcal{L}_t$ with respect to p_t is equal to the infinitesimal reversal of $\mathcal{L}_t$ with respect to q_t, followed by a Doob's h transform, i.e.,*

$$\overleftarrow{\mathcal{L}_t}^{p_t} f = (\overleftarrow{\mathcal{L}}_t^{q_t})^h f. \tag{132}$$

with

$$h_t = p_t/q_t \tag{133}$$

Proof. To prove this, we can note that when $h = p_t/q_t$, we can view it as applying time reversal with q_t first, and then applying time reversal with p_t. Therefore, the RHS for Equation (132) is applying time reversal with q_t, q_t and p_t sequentially. Applying q_t twice yields the identity map, and hence it is equivalent to applying the reversal with p_t. □

Alternatively, we can also directly prove it with algebra:

Proof.

$$\overleftarrow{\mathcal{L}}^{p_t} f = p_t^{-1}\mathcal{L}^\dagger(p_t f) - p_t^{-1}\mathcal{L}^\dagger(p_t)f, \quad \overleftarrow{\mathcal{L}}^{q_t} f = q_t^{-1}\mathcal{L}^\dagger(q_t f) - q_t^{-1}\mathcal{L}^\dagger(q_t)f \tag{134}$$

Then,

$$(\overleftarrow{\mathcal{L}}^{q_t})^h f = h^{-1}\overleftarrow{\mathcal{L}}^{q_t}(hf) - h^{-1}\overleftarrow{\mathcal{L}}^{q_t}(h)f \tag{135}$$

$$= h^{-1}q_t^{-1}\mathcal{L}^\dagger(q_t hf) - h^{-1}q_t^{-1}\mathcal{L}^\dagger(q_t h)f - h^{-1}q_t^{-1}\mathcal{L}^\dagger(q_t h)f + h^{-1}q_t^{-1}\mathcal{L}^\dagger(q_t h)f \tag{136}$$

$$= h^{-1}q_t^{-1}\mathcal{L}^\dagger(q_t hf) - h^{-1}q_t^{-1}\mathcal{L}^\dagger(q_t h)f \tag{137}$$

when we choose $h_t = p_t/q_t$, we can see this is exactly $\overleftarrow{\mathcal{L}}^{q_t} f = q_t^{-1}\mathcal{L}^\dagger(q_t f) - q_t^{-1}\mathcal{L}^\dagger(q_t)f$. □

D.3 Marginal law and generalized Doob's h-transform

The conclusions above are algebraic: so far, we have not imposed that either p_t or q_t is the actual marginal law of a Markov process. In practice, however, at least one of these density-like objects will be pinned down. For example, suppose pin one of the densities to be the marginal law. The following result gives the necessary and sufficient condition for this to hold.

Theorem D.6 (Marginal prescription for generalized Doob's h-transform)**.** *Let $(\mathcal{L}_t)_{t\in[0,1]}$ be a family of time-inhomogeneous Markov generators, with adjoints $(\mathcal{L}_t^\dagger)_{t\in[0,1]}$. Let $h_t : E \to (0,\infty)$ be strictly positive and sufficiently regular, and define the Doob's h-transform generator by*

$$\mathcal{L}_t^h f = h_t^{-1}\mathcal{L}_t(h_t f) - h_t^{-1} f\, \mathcal{L}_t h_t. \tag{138}$$

Let q_t be a strictly positive density-like function and define

$$\rho_t^h = h_t q_t. \tag{139}$$

Then ρ_t^h solves the forward equation of the h-transformed process,

$$\partial_t \rho_t^h = (\mathcal{L}_t^h)^\dagger \rho_t^h, \tag{140}$$

if and only if

$$\frac{(\partial_t + \mathcal{L}_t)h_t}{h_t} = \frac{\mathcal{L}_t^\dagger q_t - \partial_t q_t}{q_t}. \tag{141}$$

Proof. Let's prove it in both directions.

(1) Assume that $\rho_t^h = h_t q_t$ solves the forward equation

$$\partial_t \rho_t^h = (\mathcal{L}_t^h)^\dagger \rho_t^h. \tag{142}$$

Using the adjoint formula for the h-transformed generator,

$$(\mathcal{L}_t^h)^\dagger g = h_t\, \mathcal{L}_t^\dagger(h_t^{-1} g) - h_t^{-1}(\mathcal{L}_t h_t) g, \tag{143}$$

we obtain

$$(\mathcal{L}_t^h)^\dagger \rho_t^h = h_t\, \mathcal{L}_t^\dagger q_t - q_t\, \mathcal{L}_t h_t. \tag{144}$$

On the other hand,

$$\partial_t \rho_t^h = \partial_t(h_t q_t) = q_t\, \partial_t h_t + h_t\, \partial_t q_t. \tag{145}$$

Therefore,

$$q_t\,\partial_t h_t + h_t\,\partial_t q_t = h_t\,\mathcal{L}_t^\dagger q_t - q_t\,\mathcal{L}_t h_t. \tag{146}$$

Therefore,

$$q_t(\partial_t + \mathcal{L}_t)h_t = h_t(\mathcal{L}_t^\dagger q_t - \partial_t q_t). \tag{147}$$

Therefore,

$$\frac{(\partial_t + \mathcal{L}_t)h_t}{h_t} = \frac{\mathcal{L}_t^\dagger q_t - \partial_t q_t}{q_t}. \tag{148}$$

(2) Now, let's start from:

$$\frac{(\partial_t + \mathcal{L}_t)h_t}{h_t} = \frac{\mathcal{L}_t^\dagger q_t - \partial_t q_t}{q_t}. \tag{149}$$

Multiplying both sides by $h_t q_t$

$$q_t\,\partial_t h_t + q_t\,\mathcal{L}_t h_t = h_t\,\mathcal{L}_t^\dagger q_t - h_t\,\partial_t q_t. \tag{150}$$

Hence,

$$q_t\,\partial_t h_t + h_t\,\partial_t q_t = h_t\,\mathcal{L}_t^\dagger q_t - q_t\,\mathcal{L}_t h_t. \tag{151}$$

LHS is

$$q_t\,\partial_t h_t + h_t\,\partial_t q_t = \partial_t(h_t q_t) = \partial_t \rho_t^h. \tag{152}$$

Using $\rho_t^h = h_t q_t$, RHS is

$$h_t\,\mathcal{L}_t^\dagger q_t - q_t\,\mathcal{L}_t h_t = h_t\,\mathcal{L}_t^\dagger(h_t^{-1}\rho_t^h) - h_t^{-1}(\mathcal{L}_t h_t)\rho_t^h \tag{153}$$

$$= (\mathcal{L}_t^h)^\dagger \rho_t^h. \tag{154}$$

Therefore,

$$\partial_t \rho_t^h = (\mathcal{L}_t^h)^\dagger \rho_t^h. \tag{155}$$

□

Remark D.7 (Marginal prescription for generalized Doob's h-transform)**.** *By Theorem D.6, we note two cases when ρ_t^h solves the forward equation of h-transformed process:*

1. *For a general density function q_t, the generalized Doob's h-weight becomes $\frac{\mathcal{L}_t^\dagger q_t}{q_t} - \partial_t \log q_t$.*
2. *When q_t further solves the forward equation of the base process, i.e. $\mathcal{L}_t^\dagger q_t = \partial_t q_t$, the generalized Doob's h-weight is zero.*

D.4 Generalized EJE, Crooks and Score Matching Objectives as Corollary of Theorem D.6 and Generalized Doob's h-transform

Theorem D.6 plays a central role in our conclusion. In this section, we show that the generalized EJE, Crooks can be directly derived from Theorem D.6 together with generalized Doob's h-transform. Theorem D.6 also provides a new perspective for interpreting score matching objectives.

Generalized Escorted Jarzynski and Crooks fluctuation theorem We first prove Proposition 3.2, and Proposition 3.1 follows by directly taking expectation.

Proposition D.8 (Generalized Crooks fluctuation theorem, restated)**.** *Let $(\mathcal{L}_t^F)_{t\in[0,1]}$ be the forward time-inhomogeneous Markov generator, and let $(\mathcal{L}_t^F)^\dagger$ denote its adjoint. Let γ_t be a strictly positive density-like function. Define the backward generator by the infinitesimal γ_t-reversal of $\mathcal{L}_t^F$:*

$$\mathcal{L}_t^B f := \frac{1}{\gamma_t}(\mathcal{L}_t^F)^\dagger(\gamma_t f) - \frac{1}{\gamma_t}(\mathcal{L}_t^F)^\dagger(\gamma_t)\, f. \tag{156}$$

Let $\overleftarrow{Q}$ be the path law of the backward process whose marginal density q_t evolves backward in time according to

$$\partial_t q_t = -(\mathcal{L}_t^B)^\dagger q_t, \qquad q_1 = \pi_1 := \pi_B, \qquad \pi_B \propto \exp(-U_B). \tag{157}$$

Let $\overrightarrow{P}$ be the path law of the backward process whose marginal density p_t evolves backward in time according to

$$\partial_t p_t = (\mathcal{L}_t^F)^\dagger p_t, \qquad p_0 = \pi_0 := \pi_A, \qquad \pi_A \propto e^{-U_A}, \tag{158}$$

Then the Radon-Nikodym derivative between the backward path law $\overleftarrow{Q}$ and the forward path law $\overrightarrow{P}$ is

$$\frac{\mathrm{d}\overleftarrow{Q}}{\mathrm{d}\overrightarrow{P}}(X_{[0,1]}) = \exp\bigl(-\widetilde{W}(X_{[0,1]}) + \Delta F\bigr). \tag{159}$$

where

$$\widetilde{W}(X_{[0,1]}) = \int_0^1 \partial_t U_t(X_t)\,\mathrm{d}t + \int_0^1 \frac{((\mathcal{L}_t^F)^\dagger \gamma_t)(X_t)}{\gamma_t(X_t)}\,\mathrm{d}t. \tag{160}$$

Proof. We start from

$$\mathcal{L}_t^B f = \frac{1}{\gamma_t}(\mathcal{L}_t^F)^\dagger(\gamma_t f) - \frac{1}{\gamma_t}(\mathcal{L}_t^F)^\dagger(\gamma_t)\, f, \qquad (\mathcal{L}_t^B)^\dagger p = \gamma_t\, \mathcal{L}_t^F\left(\frac{p}{\gamma_t}\right) - \frac{p}{\gamma_t}\,(\mathcal{L}_t^F)^\dagger \gamma_t \tag{161}$$

We assume its marginal density is q_t:

$$\partial_t q_t = -(\mathcal{L}_t^B)^\dagger q_t, \qquad q_1 = \pi_1 := \pi_B, \qquad \pi_B \propto e^{-U_B}. \tag{162}$$

From Theorem D.6, the time-reversal of this backward process (denoted as $\overrightarrow{Q}$) has the same generator as Doob's h-transform of $\mathcal{L}_t^F$, with $h_t = q_t/\pi_t$. By generalized Doob's h in Theorem B.8, we have

$$\frac{\mathrm{d}\overrightarrow{Q}}{\mathrm{d}\overrightarrow{P}}(X_{(0,1]}|X_0) = \frac{h_1}{h_0}\exp(-\int \left[h_t^{-1}\partial_t h_t + h_t^{-1}\mathcal{L}_t h_t\right]\mathrm{d}t) \tag{163}$$

Also, $h_1 = q_1/\pi_1 = 1$ and $p_0 = \pi_0$, hence

$$\frac{\mathrm{d}\overleftarrow{Q}}{\mathrm{d}\overrightarrow{P}}(X_{[0,1]}) = \frac{\mathrm{d}\overrightarrow{Q}}{\mathrm{d}\overleftarrow{Q}}(X_{[0,1]})\frac{\mathrm{d}\overleftarrow{Q}}{\mathrm{d}\overrightarrow{P}}(X_{(0,1]}|X_0)\frac{q_0(X_0)}{p_0(X_0)} \tag{164}$$

$$= \frac{\mathrm{d}\overrightarrow{Q}}{\mathrm{d}\overrightarrow{P}}(X_{(0,1]}|X_0)h_0(X_0) \tag{165}$$

$$= \exp(-\int \left[h_t^{-1}\partial_t h_t + h_t^{-1}\mathcal{L}_t h_t\right]\mathrm{d}t) \tag{166}$$

By Theorem D.6, we have

$$h_t^{-1}\partial_t h_t + h_t^{-1}\mathcal{L}_t h_t = \pi_t^{-1}\mathcal{L}_t^\dagger \pi_t - \pi_t^{-1}\partial_t \pi_t \tag{167}$$

which proves the Crooks fluctuation theorem. □

Score matching objectives Theorem D.6 also connects with score-matching objectives. Different from the Jarzynskis discussed above, where we fix the path π_t, we aim to learn a path of marginal densities φ_t (in the form of score or concrete score) to align with the marginal law of some pre-defined forward process. Assume the diffusion model has a forward generator $\mathcal{L}$, corresponding to a known forward corruption process with marginal p_t, and we aim to learn $\varphi_t = p_t$.

Explicit score matching Let's consider the generator obtained by taking infinitesimal reversal w.r.t p_t and φ_t:

$$\overleftarrow{\mathcal{L}}^{p_t} f = p_t^{-1}\mathcal{L}^\dagger(p_t f) - p_t^{-1}\mathcal{L}^\dagger(p_t)f, \quad \overleftarrow{\mathcal{L}}^{\varphi_t} f = \varphi_t^{-1}\mathcal{L}^\dagger(\varphi_t f) - \varphi_t^{-1}\mathcal{L}^\dagger(\varphi_t)f \tag{168}$$

Using results in Proposition D.5, we have that $\overleftarrow{\mathcal{L}}^{p_t}$ is the Doob's-$h = p_t/\varphi_t$ transform of $\overleftarrow{\mathcal{L}}^{\varphi_t}$.

Additionally, as we require p_t to be the true marginal, using Theorem D.6, we have

$$\frac{(-\partial_t + \overleftarrow{\mathcal{L}}^{\varphi_t})h_t}{h_t} = \frac{(\overleftarrow{\mathcal{L}}^{\varphi_t})^\dagger \varphi_t + \partial_t \varphi_t}{\varphi_t}. \tag{169}$$

Note the sign flip as the process is running backward in time. If we want φ_t to be the marginal, i.e., $(\overleftarrow{\mathcal{L}}^{\varphi_t})^\dagger \varphi_t + \partial_t \varphi_t = 0$, it suffices for us to impose $\frac{(-\partial_t + \overleftarrow{\mathcal{L}}^{\varphi_t})h_t}{h_t} = 0$.

Taking expectations over time and state space:

$$\int_0^1 \mathbb{E}_{p_t}\left[\frac{(\partial_t - \overleftarrow{\mathcal{L}}^{\varphi_t})h_t}{h_t}\right] \mathrm{d}t, \tag{170}$$

Note that

$$\begin{aligned}
\int p_t(x) h_t^{-1} \overleftarrow{\mathcal{L}}^{\varphi_t}(h_t) \mathrm{d}x &= \int p_t h_t^{-1} \varphi_t^{-1} \mathcal{L}^\dagger(\varphi_t h_t) \mathrm{d}x - \int p_t h_t^{-1} \varphi_t^{-1} \mathcal{L}^\dagger(\varphi_t) h_t \mathrm{d}x && (171)\\
&= \int p_t h_t^{-1} \varphi_t^{-1} \partial_t p_t \mathrm{d}x - \int p_t h_t^{-1} \varphi_t^{-1} \mathcal{L}^\dagger(\varphi_t) h_t \mathrm{d}x && (172)\\
&= \cancel{\int \partial_t p_t \mathrm{d}x} - \int p_t h_t^{-1} \varphi_t^{-1} \mathcal{L}^\dagger(\varphi_t) h_t \mathrm{d}x && (173)\\
&= -\int h_t \mathcal{L}^\dagger(\varphi_t) \mathrm{d}x && (174)\\
&= -\int \varphi_t \mathcal{L}(h_t) \mathrm{d}x && (175)\\
&= -\int p_t h_t^{-1} \mathcal{L}(h_t) \mathrm{d}x && (176)
\end{aligned}$$

Hence,

$$\begin{aligned}
&\int_0^1 \mathbb{E}_{p_t}\left[\frac{(\partial_t - \overleftarrow{\mathcal{L}}^{\varphi_t})h_t}{h_t}\right] \mathrm{d}t && (177)\\
&= \int_0^1 \mathbb{E}_{p_t}\left[\frac{(\partial_t + \mathcal{L}_t)h_t}{h_t}\right] \mathrm{d}t && (178)\\
&= \int_0^1 \mathbb{E}_{p_t}\left[\partial_t \log h_t + h_t^{-1} \mathcal{L}_t h_t\right] \mathrm{d}t && (179)
\end{aligned}$$

With Dynkin's formula, we have $\mathrm{d}\log h_t = \partial_t \log h_t \mathrm{d}t + \mathcal{L}_t \log h_t \mathrm{d}t + \mathrm{d}M$, and hence

$$\begin{aligned}
&\int_0^1 \mathbb{E}_{p_t}\left[\partial_t \log h_t + h_t^{-1} \mathcal{L}_t h_t\right] \mathrm{d}t && (180)\\
&= \int_0^1 \mathbb{E}_{p_t}\left[\mathrm{d}\log h_t + h_t^{-1} \mathcal{L}_t h_t \mathrm{d}t - \mathcal{L}_t \log h_t \mathrm{d}t\right] && (181)\\
&= \underbrace{\int_0^1 \mathbb{E}_{p_t}\left[h_t^{-1} \mathcal{L}_t h_t \mathrm{d}t - \mathcal{L}_t \log h_t \mathrm{d}t\right]}_{\text{ESM}} + \mathbb{E}_{p_1}[\log h_1] - \mathbb{E}_{p_0}[\log h_0] && (182)
\end{aligned}$$

Denoising score matching Consider the generator obtained by taking infinitesimal reversal w.r.t $p_t(x_t|x_0)$ and φ_t:

$$\begin{aligned}
\overleftarrow{\mathcal{L}}^{p_t(\cdot|x_0)} f &= p_t(\cdot|x_0)^{-1} \mathcal{L}^\dagger(p_t(\cdot|x_0) f) - p_t(\cdot|x_0)^{-1} \mathcal{L}^\dagger(p_t(\cdot|x_0)) f, && (183)\\
\overleftarrow{\mathcal{L}}^{\varphi_t} f &= \varphi_t^{-1} \mathcal{L}^\dagger(\varphi_t f) - \varphi_t^{-1} \mathcal{L}^\dagger(\varphi_t) f && (184)
\end{aligned}$$

Using results in Proposition D.5, we have that $\overleftarrow{\mathcal{L}}^{p_t(\cdot|x_0)}$ is the Doob's-$h = p_t(\cdot|x_0)/\varphi_t$ transform of $\overleftarrow{\mathcal{L}}^{\varphi_t}$. As $p_t(\cdot|x_0)$ is the true marginal for the backward process, using Theorem D.6, we have

$$\frac{(-\partial_t + \overleftarrow{\mathcal{L}}^{\varphi_t})h_t}{h_t} = \frac{(\overleftarrow{\mathcal{L}}^{\varphi_t})^\dagger \varphi_t + \partial_t \varphi_t}{\varphi_t}. \tag{185}$$

If we want φ_t to be the marginal, i.e., $(\overleftarrow{\mathcal{L}}^{\varphi_t})^\dagger \varphi_t + \partial_t \varphi_t = 0$, it suffices for us to impose $\frac{(-\partial_t + \overleftarrow{\mathcal{L}}^{\varphi_t}) h_t}{h_t} = 0$. Taking expectations over time and state space:

$$\int_0^1 \mathbb{E}_{x_0, x_t} \left[\frac{(\partial_t - \overleftarrow{\mathcal{L}}^{\varphi_t}) h_t}{h_t} \right] \mathrm{d}t, \tag{186}$$

Similar to the ESM case, we have

$$\int p_0(x) p_t(x_t|x_0) h_t^{-1} \overleftarrow{\mathcal{L}}^{\varphi_t}(h_t) \mathrm{d}x \tag{187}$$

$$= \int p_0(x) p_t(x_t|x_0) h_t^{-1} \varphi_t^{-1} \mathcal{L}^\dagger(\varphi_t h_t) \mathrm{d}x - \int p_0(x) p_t(x_t|x_0) h_t^{-1} \varphi_t^{-1} \mathcal{L}^\dagger(\varphi_t) h_t \mathrm{d}x \tag{188}$$

$$= - \int p_0 p_t(x_t|x_0) h_t^{-1} \mathcal{L}(h_t) \mathrm{d}x \tag{189}$$

We then apply Dynkin's formula to obtain the DSM formulation with boundary conditions.

D.5 Proof of Theorem C.4

Proof. The algebraic identities

$$H_h H_{h'} = H_{hh'}, \qquad H_h R_q = R_{hq}, \qquad R_p H_h = R_{p/h}, \qquad R_p R_q = H_{p/q} \tag{190}$$

show that $\mathcal{G}$ is closed under composition. Associativity follows from associativity of composition. The identity element is $H_1 = I$, and the inverses are

$$H_h^{-1} = H_{1/h}, \qquad R_p^{-1} = R_p. \tag{191}$$

Thus $\mathcal{G}$ is a group.

Now define

$$\Phi : \mathcal{G} \to A \rtimes \mathbb{Z}_2 \tag{192}$$

by

$$\Phi(H_h) = ([h], 0), \qquad \Phi(R_p) = ([p], 1), \tag{193}$$

where $[h]$ denotes the equivalence class of h modulo positive constants. The multiplication in $A \rtimes \mathbb{Z}_2$ is

$$(a, z_1)(b, z_2) = \big(a b^{(-1)^{z_1}}, z_1 + z_2 \bmod 2\big). \tag{194}$$

Using this rule, the four products above are mapped respectively to

$$([h], 0)([h'], 0) = ([hh'], 0), \tag{195}$$

$$([h], 0)([q], 1) = ([hq], 1), \tag{196}$$

$$([p], 1)([h], 0) = ([p/h], 1), \tag{197}$$

$$([p], 1)([q], 1) = ([p/q], 0). \tag{198}$$

Hence Φ preserves multiplication. It is also bijective, since every element of $A \rtimes \mathbb{Z}_2$ is uniquely of the form $([h], 0)$ or $([p], 1)$. Therefore,

$$\mathcal{G} \cong A \rtimes \mathbb{Z}_2 = \mathrm{Dih}(A). \tag{199}$$

□

D.6 Discrete-time Generalized EJE and Crooks

Proposition D.9 (Discrete-time Jarzynski equality and Crooks relation, restate)**.** *Let $\{U_n\}_{n=0}^N$ be an energy path with $U_0 = U_A$ and $U_N = U_B$, and define*

$$\gamma_n(x) := e^{-U_n(x)}, \qquad Z_n := \int \gamma_n(x)\, \mathrm{d}x, \qquad \pi_n(x) := \frac{\gamma_n(x)}{Z_n}. \tag{200}$$

Let $\{P_{F,n}\}_{n=1}^N$ be forward transition kernels, and let $\overrightarrow{P}$ be the forward path law

$$X_0 \sim \pi_A \propto e^{-U_A}, \qquad X_n \mid X_{n-1} \sim P_{F,n}(\cdot \mid X_{n-1}), \qquad n = 1, \dots, N, \tag{201}$$

Define the pushforward density

$$(P_{F,n}\#\pi_{n-1})(x) := \int P_{F,n}(x \mid x')\, \pi_{n-1}(x')\, \mathrm{d}x'. \tag{202}$$

Then the discrete-time generalized work is

$$W(X_{0:N}) = \sum_{n=1}^{N} \left[U_n(X_n) - U_{n-1}(X_n) + \log \frac{(P_{F,n}\#\pi_{n-1})(X_n)}{\pi_{n-1}(X_n)} \right], \tag{203}$$

The discrete-time Jarzynski equality is

$$\Delta F = -\log \frac{Z_B}{Z_A} = -\log \mathbb{E}_{\overrightarrow{P}} \left[\exp\bigl(-W(X_{0:N})\bigr) \right]. \tag{204}$$

Moreover, assume that there exist backward kernels $\{P_{B,n}\}_{n=1}^N$ satisfying, for each $n = 1, \dots, N$,

$$\pi_{n-1}(x_{n-1})\, P_{F,n}(x_n \mid x_{n-1}) = (P_{F,n}\#\pi_{n-1})(x_n)\, P_{B,n}(x_{n-1} \mid x_n). \tag{205}$$

Let $\overleftarrow{Q}$ be the corresponding backward path law

$$X_N \sim \pi_B \propto e^{-U_B}, \qquad X_{n-1} \mid X_n \sim P_{B,n}(\cdot \mid X_n), \qquad n = N, \dots, 1, \tag{206}$$

Then the discrete-time Crooks relation is

$$\frac{\mathrm{d}\overleftarrow{Q}}{\mathrm{d}\overrightarrow{P}}(X_{0:N}) = \frac{\pi_B(X_N) \prod_{n=1}^N P_{B,n}(X_{n-1} \mid X_n)}{\pi_A(X_0) \prod_{n=1}^N P_{F,n}(X_n \mid X_{n-1})} \tag{207}$$

$$= \exp\bigl(-W(X_{0:N}) + \Delta F\bigr). \tag{208}$$

Proof.

$$\frac{\pi_B(X_N) \prod_{n=1}^N P_{B,n}(X_{n-1} \mid X_n)}{\pi_A(X_0) \prod_{n=1}^N P_{F,n}(X_n \mid X_{n-1})} \tag{209}$$

$$= \frac{P_{B,1}(x_0|x_1) \cdots P_{B,N}(x_{N-1}|x_N) \pi_N(x_N)}{\pi_0(x_0) P_{F,1}(x_1|x_0) \cdots P_{F,N}(x_N|x_{N-1})} \tag{210}$$

$$= \frac{1}{\pi_0(x_0)} \frac{P_{B,1}(x_0|x_1)}{P_{F,1}(x_1|x_0)} \cdots \frac{P_{B,N}(x_{N-1}|x_N)}{P_{F,N}(x_N|x_{N-1})} \pi_N(x_N) \tag{211}$$

$$= \frac{1}{\pi_0(x_0)} \frac{P_{B,1}(x_0|x_1)}{P_{F,1}(x_1|x_0)} \frac{\pi_0(x_0) P_{F,1}(x_1|x_0)}{P_{F,1}\#\pi_0(x_1) P_{B,1}(x_0|x_1)} \cdots$$

$$\cdots \frac{P_{B,N}(x_{N-1}|x_N)}{P_{F,N}(x_N|x_{N-1})} \frac{\pi_{N-1}(x_{N-1}) P_{F,N}(x_N|x_{N-1})}{P_{F,N}\#\pi_{N-1}(x_N) P_{B,N}(x_{N-1}|x_N)} \pi_N(x_N) \tag{212}$$

$$= \frac{1}{\pi_0(x_0)} \frac{\pi_0(x_0)}{P_{F,1}\#\pi_0(x_1)} \frac{\pi_1(x_1)}{P_{F,2}\#\pi_1(x_2)} \cdots \frac{\pi_{N-1}(x_{N-1})}{P_{F,N}\#\pi_{N-1}(x_N)} \pi_N(x_N) \tag{213}$$

$$= \frac{1}{\pi_0(x_0)} \frac{\pi_0(x_0)\pi_0(x_1)}{P_{F,1}\#\pi_0(x_1)\pi_0(x_1)} \frac{\pi_1(x_1)\pi_1(x_2)}{P_{F,2}\#\pi_1(x_2)\pi_1(x_2)} \cdots \frac{\pi_{N-1}(x_{N-1})\pi_{N-1}(x_N)}{P_{F,N}\#\pi_{N-1}(x_N)\pi_{N-1}(x_N)} \pi_N(x_N) \tag{214}$$

$$= \frac{\pi_0(x_1)}{P_{F,1}\#\pi_0(x_1)} \frac{\pi_1(x_1)}{\pi_0(x_1)} \frac{\pi_1(x_2)}{P_{F,2}\#\pi_1(x_2)} \frac{\pi_2(x_2)}{\pi_1(x_2)} \cdots \frac{\pi_{N-1}(x_N)}{P_{F,N}\#\pi_{N-1}(x_N)} \frac{\pi_N(x_N)}{\pi_{N-1}(x_N)} \tag{215}$$

Therefore, taking log-space and canceling normalization factors for π-s, we have

$$W = (U_n(x_n) - U_{n-1}(x_n)) + \log \frac{P_{F,n}\#\pi_{n-1}(x_n)}{\pi_{n-1}(x_n)} \tag{216}$$

Then, taking expectation over the work, we obtain the discrete-time Jarzynski. □

D.7 Proof of Remark 3.7 (Variance-dissipation Relation)

Remark D.10 (Variance-dissipation Relation, restate)**.** *Denote the free energy estimate as* $\widehat{Z} := \widehat{\exp(-\Delta F)} = \frac{1}{N}\sum_{i=1}^{N}\exp(-\widetilde{W}(X_{[0,1]}^{(i)}))$ *with* $X_{[0,1]}^{(i)} \sim \overrightarrow{P}$*. Then*

$$\mathrm{Var}(\widehat{Z}) = \mathcal{O}\left(\tfrac{1}{N}\chi^2(\overleftarrow{Q}\,\|\,\overrightarrow{P})\right) \geq \mathcal{O}\left(\tfrac{1}{N}D_{\mathrm{KL}}(\overleftarrow{Q}\,\|\,\overrightarrow{P})\right) \tag{217}$$

Proof.

$$\mathrm{Var}(\widehat{Z}) = \mathbb{E}\left[(\frac{1}{N}\sum_{i=1}^{N}\exp(-W^{(i)}))^2\right] + \mathrm{const.} \tag{218}$$

$$= \mathbb{E}\left[(\frac{1}{N}\sum_{i=1}^{N}\frac{\mathrm{d}\overleftarrow{Q}}{\mathrm{d}\overrightarrow{P}})^2\right] + \mathrm{const.} \tag{219}$$

$$= \frac{1}{N}\mathrm{Var}(\frac{\mathrm{d}\overleftarrow{Q}}{\mathrm{d}\overrightarrow{P}}) + \mathrm{const.} \tag{220}$$

$$= \frac{1}{N}\int\left(\frac{\mathrm{d}\overleftarrow{Q}}{\mathrm{d}\overrightarrow{P}} - 1\right)^2 \mathrm{d}\overrightarrow{P} + \mathrm{const.} \tag{221}$$

$$= \mathcal{O}\left(\frac{1}{N}\chi^2(\overleftarrow{Q}\,\|\,\overrightarrow{P})\right) \tag{222}$$

We then prove the lower bound of this. Let

$$r = \frac{\mathrm{d}\overleftarrow{Q}}{\mathrm{d}\overrightarrow{P}}. \tag{223}$$

and we have

$$D_{\mathrm{KL}}(\overleftarrow{Q}\|\overrightarrow{P}) = \int r\log r\,\mathrm{d}\overrightarrow{P}, \tag{224}$$

$$\chi^2(\overleftarrow{Q}\|\overrightarrow{P}) = \int (r-1)^2\,\mathrm{d}\overrightarrow{P} = \int r^2\,\mathrm{d}\overrightarrow{P} - 1. \tag{225}$$

We note that

$$\log u \leq u - 1, \qquad u > 0. \tag{226}$$

and

$$u\log u \leq u(u-1) = u^2 - u. \tag{227}$$

Set u to r and take expectation,

$$D_{\mathrm{KL}}(\overleftarrow{Q}\|\overrightarrow{P}) = \int r\log r\,\mathrm{d}\overrightarrow{P} \tag{228}$$

$$\leq \int (r^2 - r)\,\mathrm{d}\overrightarrow{P} \tag{229}$$

$$= \int r^2\,\mathrm{d}\overrightarrow{P} - \int r\,\mathrm{d}\overrightarrow{P} \tag{230}$$

$$= \int r^2\,\mathrm{d}\overrightarrow{P} - 1 \tag{231}$$

$$= \chi^2(\overleftarrow{Q}\|\overrightarrow{P}), \tag{232}$$

□

D.8 Proof of Proposition 4.1 (marginalization maintains time-reversal)

Proposition D.11 (Marginalization maintains time-reversal, restate)**.** *Let $p_t(x_t|x_0, x_1)$ be a conditional interpolant, and let $\mathcal{L}_{F,t}^{x_0,x_1}$ and $\mathcal{L}_{B,t}^{x_0,x_1}$ be the corresponding forward and backward conditional generators. Assume that, for every (x_0, x_1), they are time reversals of each other, i.e.*

$$\mathcal{L}_{B,t}^{x_0,x_1} f = \frac{1}{p_t(x_t|x_0,x_1)}(\mathcal{L}_{F,t}^{x_0,x_1})^\dagger(p_t(x_t|x_0,x_1)f) - \frac{1}{p_t(x_t|x_0,x_1)}(\mathcal{L}_{F,t}^{x_0,x_1})^\dagger p_t(x_t|x_0,x_1)\, f \tag{233}$$

Then $\mathcal{L}_{F,t}$ and $\mathcal{L}_{B,t}$ defined by Equation (24) are also time-reversal with respect to $p_t(x_t)$, namely

$$\mathcal{L}_{B,t} f = \frac{1}{p_t}\mathcal{L}_{F,t}^\dagger(p_t f) - \frac{1}{p_t}\mathcal{L}_{F,t}^\dagger p_t\, f. \tag{234}$$

Proof. Write

$$z := (x_0, x_1), \qquad p_t^z(x) := p_t(x \mid x_0, x_1), \qquad \pi(z) := \pi(x_0, x_1). \tag{235}$$

The marginal density is

$$p_t(x) = \int p_t^z(x)\, \pi(z)\, \mathrm{d}z, \tag{236}$$

and the posterior over endpoints given the current state is

$$\pi_t(z \mid x) = \frac{\pi(z) p_t^z(x)}{p_t(x)}. \tag{237}$$

By the definition of the marginal generators in Equation (24),

$$\mathcal{L}_{F,t} f(x) = \int \mathcal{L}_{F,t}^z f(x)\, \pi_t(z \mid x)\, \mathrm{d}z, \qquad \mathcal{L}_{B,t} f(x) = \int \mathcal{L}_{B,t}^z f(x)\, \pi_t(z \mid x)\, \mathrm{d}z. \tag{238}$$

We first compute the adjoint of the marginal forward generator. For any test function f and density r, we have

$$\int r(x)\, \mathcal{L}_{F,t} f(x)\, \mathrm{d}x = \int r(x) \left[\int \mathcal{L}_{F,t}^z f(x)\, \pi_t(z \mid x)\, \mathrm{d}z \right] \mathrm{d}x \tag{239}$$

$$= \int \int r(x) \pi_t(z \mid x)\, \mathcal{L}_{F,t}^z f(x)\, \mathrm{d}x\, \mathrm{d}z \tag{240}$$

$$= \int \int f(x)\, (\mathcal{L}_{F,t}^z)^\dagger (r(\cdot)\pi_t(z \mid \cdot))\, (x)\, \mathrm{d}x\, \mathrm{d}z \tag{241}$$

$$= \int f(x) \left[\int (\mathcal{L}_{F,t}^z)^\dagger (r(\cdot)\pi_t(z \mid \cdot))\, (x)\, \mathrm{d}z \right] \mathrm{d}x. \tag{242}$$

Therefore,

$$\mathcal{L}_{F,t}^\dagger r(x) = \int (\mathcal{L}_{F,t}^z)^\dagger (r(\cdot)\pi_t(z \mid \cdot))\, (x)\, \mathrm{d}z. \tag{243}$$

Now take $r = p_t f$. Since

$$p_t(x)\pi_t(z \mid x) = \pi(z) p_t^z(x), \tag{244}$$

we get

$$\mathcal{L}_{F,t}^\dagger (p_t f)(x) = \int \pi(z) (\mathcal{L}_{F,t}^z)^\dagger (p_t^z f)\, (x)\, \mathrm{d}z. \tag{245}$$

We can move $\pi(z)$ outside the integral as the generator operates on x. Similarly,

$$\mathcal{L}_{F,t}^\dagger p_t(x) = \int \pi(z) (\mathcal{L}_{F,t}^z)^\dagger p_t^z(x)\, \mathrm{d}z. \tag{246}$$

Therefore, the time-reversal of the marginal forward generator is

$$\frac{1}{p_t(x)}\mathcal{L}_{F,t}^\dagger(p_t f)(x) - \frac{1}{p_t(x)}\mathcal{L}_{F,t}^\dagger p_t(x)\, f(x) \tag{247}$$

$$= \frac{1}{p_t(x)}\int \pi(z)\left[(\mathcal{L}_{F,t}^z)^\dagger(p_t^z f)(x) - (\mathcal{L}_{F,t}^z)^\dagger p_t^z(x)\, f(x)\right] \mathrm{d}z. \tag{248}$$

On the other hand, we assume

$$\mathcal{L}_{B,t}^z f(x) = \frac{1}{p_t^z(x)}(\mathcal{L}_{F,t}^z)^\dagger(p_t^z f)(x) - \frac{1}{p_t^z(x)}(\mathcal{L}_{F,t}^z)^\dagger p_t^z(x)\, f(x). \tag{249}$$

Thus the marginal backward generator satisfies

$$\mathcal{L}_{B,t} f(x) = \int \mathcal{L}_{B,t}^z f(x)\, \pi_t(z \mid x)\, \mathrm{d}z \tag{250}$$

$$= \int \left[\frac{1}{p_t^z(x)}(\mathcal{L}_{F,t}^z)^\dagger(p_t^z f)(x) - \frac{1}{p_t^z(x)}(\mathcal{L}_{F,t}^z)^\dagger p_t^z(x)\, f(x)\right] \frac{\pi(z) p_t^z(x)}{p_t(x)}\, \mathrm{d}z \tag{251}$$

$$= \frac{1}{p_t(x)}\int \pi(z)\left[(\mathcal{L}_{F,t}^z)^\dagger(p_t^z f)(x) - (\mathcal{L}_{F,t}^z)^\dagger p_t^z(x)\, f(x)\right] \mathrm{d}z. \tag{252}$$

Comparing this with Equation (248), we obtain

$$\mathcal{L}_{B,t} f(x) = \frac{1}{p_t(x)}\mathcal{L}_{F,t}^\dagger(p_t f)(x) - \frac{1}{p_t(x)}\mathcal{L}_{F,t}^\dagger p_t(x)\, f(x). \tag{253}$$

□

D.9 Connection between Generator Matching and Denoising Score Matching

We consider two approaches in the main text: we either learn a backward process directly, or learn the marginal to reverse the forward process. We will show that learning the backward process with generator matching is the same as learning the marginal.

For simplicity, we consider the case of a denoising Markov model. For the bridge model we discussed above, we could simply choose $x_0 = I_t$, and define a new axis of time τ from I to x. Assume we have a forward generator $\mathcal{L}$, corresponding to a known forward process. Assume the law of forward process, conditional on the initial sample x_0, is $p_t^{x_0}$. The backward conditional generator is given by

$$\tilde{\mathcal{L}}_t^{x_0} f = (p_t^{x_0})^{-1}(\mathcal{L}_t)^\dagger(p_t^{x_0} f) - (p_t^{x_0})^{-1}(\mathcal{L}_t)^\dagger(p_t^{x_0}) f, \tag{254}$$

We define another backward process as

$$\tilde{\mathcal{L}}_t f = \varphi_t^{-1}(\mathcal{L}_t)^\dagger(\varphi_t f) - \varphi_t^{-1}(\mathcal{L}_t)^\dagger(\varphi_t) f, \tag{255}$$

Recall the DSM objective:

$$\mathcal{J}_{\text{DSM}}(\varphi_t) = \mathbb{E}_{x_0, x_t}\left[\frac{\mathcal{L}_t(p(x_t|x_0)/\varphi_t(x_t))}{p(x_t|x_0)/\varphi_t(x_t)} - \mathcal{L}_t \log(p(x_t|x_0)/\varphi_t(x_t))\right] \tag{256}$$

We now assume an additive perturbation in log-space of φ:

$$\varphi_t' = \varphi_t \exp(-\epsilon\phi) \tag{257}$$

Then, we have

$$\mathcal{J}_{\text{DSM}}(\varphi_t') = \mathbb{E}_{x_0|x_t}\left[\frac{\mathcal{L}_t(\exp(\epsilon\phi(x_t))p(x_t|x_0)/\varphi_t(x_t))}{\exp(\epsilon\phi(x_t))p(x_t|x_0)/\varphi_t(x_t)} - \mathcal{L}_t \log(\exp(\epsilon\phi(x_t))p(x_t|x_0)/\varphi_t(x_t))\right] \tag{258}$$

Take the gradient w.r.t. ϵ, we have

$$\frac{\partial}{\partial\epsilon}\mathcal{J}_{\text{DSM}}(\varphi_t') \tag{259}$$

$$= \mathbb{E}_{x_0,x_t}\Bigg[\frac{\mathcal{L}_t(\phi(x_t)\exp(\epsilon\phi(x_t))\,p(x_t|x_0)/\varphi_t(x_t))}{\exp(\epsilon\phi(x_t))\,p(x_t|x_0)/\varphi_t(x_t)} \tag{260}$$

$$- \frac{\mathcal{L}_t(\exp(\epsilon\phi(x_t))\,p(x_t|x_0)/\varphi_t(x_t))}{(\exp(\epsilon\phi(x_t))\,p(x_t|x_0)/\varphi_t(x_t))^2}\phi(x_t)\exp(\epsilon\phi(x_t))\,p(x_t|x_0)/\varphi_t(x_t) \tag{261}$$

$$- \mathcal{L}_t\phi(x_t)\Bigg]. \tag{262}$$

Hence,

$$\frac{\partial}{\partial\epsilon}\mathcal{J}_{\text{DSM}}(\varphi_t') = \mathbb{E}_{x_0,x_t}\Bigg[\frac{\mathcal{L}_t(\phi(x_t)\exp(\epsilon\phi(x_t))\,p(x_t|x_0)/\varphi_t(x_t))}{\exp(\epsilon\phi(x_t))\,p(x_t|x_0)/\varphi_t(x_t)} \tag{263}$$

$$- \frac{\mathcal{L}_t(\exp(\epsilon\phi(x_t))\,p(x_t|x_0)/\varphi_t(x_t))}{\exp(\epsilon\phi(x_t))\,p(x_t|x_0)/\varphi_t(x_t)}\,\phi(x_t) - \mathcal{L}_t\phi(x_t)\Bigg]. \tag{264}$$

Therefore, the Gâteaux derivative of DSM objective is given by

$$\frac{d}{d\epsilon}\mathcal{J}_{\text{DSM}}(\varphi_t')\Big|_{\epsilon=0} \tag{265}$$

$$=\mathbb{E}_{x_0,x_t}\left[\frac{\mathcal{L}_t(\phi(x_t)p(x_t|x_0)/\varphi_t(x_t))}{p(x_t|x_0)/\varphi_t(x_t)} - \frac{\mathcal{L}_t(p(x_t|x_0)/\varphi_t(x_t))}{p(x_t|x_0)/\varphi_t(x_t)}\,\phi(x_t) - \mathcal{L}_t\phi(x_t)\right] \tag{266}$$

$$=\mathbb{E}_{x_0}\left[\int \frac{\mathcal{L}_t(\phi(x_t)p(x_t|x_0)/\varphi_t(x_t))}{1/\varphi_t(x_t)} - \frac{\mathcal{L}_t(p(x_t|x_0)/\varphi_t(x_t))}{1/\varphi_t(x_t)}\,\phi(x_t) - p(x_t|x_0)\mathcal{L}_t\phi(x_t)\mathrm{d}x_t\right] \tag{267}$$

$$=\mathbb{E}_{x_0}\Bigg[\int \frac{\mathcal{L}_t(\phi(x_t)p(x_t|x_0)/\varphi_t(x_t))}{1/\varphi_t(x_t)} - \frac{\mathcal{L}_t(p(x_t|x_0)/\varphi_t(x_t))}{1/\varphi_t(x_t)}\,\phi(x_t) \tag{268}$$

$$- p(x_t|x_0)\mathcal{L}_t\phi(x_t) + p(x_t|x_0)\phi(x_t)\mathcal{L}_t 1\mathrm{d}x_t\Bigg] \tag{269}$$

$$=\mathbb{E}_{x_0}\Bigg[\int \phi(x_t)p(x_t|x_0)/\varphi_t(x_t)\mathcal{L}_t^\dagger(\varphi_t(x_t)) - p(x_t|x_0)/\varphi_t(x_t)\mathcal{L}_t^\dagger\left(\varphi_t(x_t)\phi(x_t)\right) \tag{270}$$

$$- \phi(x_t)\mathcal{L}_t^\dagger p(x_t|x_0) + \mathcal{L}_t^\dagger(p(x_t|x_0)\phi(x_t))\mathrm{d}x_t\Bigg] \tag{271}$$

$$=\mathbb{E}_{x_0,x_t}\Bigg[\phi(x_t)/\varphi_t(x_t)\mathcal{L}_t^\dagger(\varphi_t(x_t)) - 1/\varphi_t(x_t)\mathcal{L}_t^\dagger\left(\varphi_t(x_t)\phi(x_t)\right) \tag{272}$$

$$- \phi(x_t)/p(x_t|x_0)\mathcal{L}_t^\dagger p(x_t|x_0) + 1/p(x_t|x_0)\mathcal{L}_t^\dagger(p(x_t|x_0)\phi(x_t))\Bigg] \tag{273}$$

$$=\mathbb{E}_{x_0,x_t}[-\tilde{\mathcal{L}}_t\phi + \tilde{\mathcal{L}}_t^{x_0}\phi] \tag{274}$$

Therefore, optimizing the DSM objective is weakly equivalent to matching the generators.

E Concrete Instantiations and Examples of the Generalized EJE, Crooks and Learning Methods

In the main text, we state the generalized EJE and Crooks fluctuation theorem in an abstract Markov process form, and discuss learning strategies at the same level of generality. Here, we instantiate

these results in several representative settings. We consider continuous Euclidean dynamics based on deterministic flows, overdamped SDEs, and momentum-augmented SDEs, as well as finite-state Continuous-Time Markov Chains (CTMC) dynamics.

E.1 Instantiate the Generalized EJE and Crooks

E.1.1 Deterministic Flow in Continuous Euclidean Space

We first instantiate Propositions 3.1 and 3.2 for flow-based transport in Euclidean space. Let $X_t \in \mathbb{R}^d$ follow the time-dependent ODE

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = v_t(X_t), \qquad X_0 \sim \pi_A. \tag{275}$$

The corresponding generator is

$$\mathcal{L}_t^F f = v_t \cdot \nabla f, \qquad (\mathcal{L}_t^F)^\dagger p = -\nabla \cdot (v_t p). \tag{276}$$

Therefore,

$$\frac{((\mathcal{L}_t^F)^\dagger \gamma_t)(x)}{\gamma_t(x)} = -\frac{\nabla \cdot (v_t(x)\gamma_t(x))}{\gamma_t(x)} \tag{277}$$

$$= -\nabla \cdot v_t(x) - v_t(x) \cdot \nabla \log \gamma_t(x) \tag{278}$$

$$= -\nabla \cdot v_t(x) + v_t(x) \cdot \nabla U_t(x). \tag{279}$$

Plugging this into Equation (10), the generalized work becomes

$$\widetilde{W}(X_{[0,1]}) = \int_0^1 \left[\partial_t U_t(X_t) + v_t(X_t) \cdot \nabla U_t(X_t) - \nabla \cdot v_t(X_t)\right] \mathrm{d}t. \tag{280}$$

By the definition of total derivative, we have

$$\frac{\mathrm{d}}{\mathrm{d}t} U_t(X_t) = \partial_t U_t(X_t) + v_t(X_t) \cdot \nabla U_t(X_t), \tag{281}$$

and hence we have

$$\widetilde{W}(X_{[0,1]}) = U_1(X_1) - U_0(X_0) - \int_0^1 \nabla \cdot v_t(X_t)\, \mathrm{d}t. \tag{282}$$

Equivalently, if T denotes the deterministic map from X_0 to X_1, then

$$\widetilde{W}(X_{[0,1]}) = U_1(T(X_0)) - U_0(X_0) - \log \left|\det \nabla T(X_0)\right|. \tag{283}$$

This recovers the standard formulation of target FEP [Wirnsberger et al., 2020, Zhao and Wang, 2023, Ding and Zhang, 2021].

The corresponding backward generator from Equation (11) can be derived similarly:

$$(\mathcal{L}_t^F)^\dagger(\gamma_t f) = -\nabla \cdot (v_t \gamma_t f), \qquad (\mathcal{L}_t^F)^\dagger \gamma_t = -\nabla \cdot (v_t \gamma_t). \tag{284}$$

Hence

$$\mathcal{L}_t^B f = -\frac{1}{\gamma_t} \nabla \cdot (v_t \gamma_t f) + \frac{f}{\gamma_t} \nabla \cdot (v_t \gamma_t) \tag{285}$$

$$= -\frac{1}{\gamma_t} \left[f \nabla \cdot (v_t \gamma_t) + \gamma_t v_t \cdot \nabla f\right] + \frac{f}{\gamma_t} \nabla \cdot (v_t \gamma_t) \tag{286}$$

$$= -v_t \cdot \nabla f. \tag{287}$$

Therefore,

$$(\mathcal{L}_t^B)^\dagger q = \nabla \cdot (v_t q). \tag{288}$$

The backward marginal equation in Proposition 3.2 is then

$$\partial_t q_t = -(\mathcal{L}_t^B)^\dagger q_t = -\nabla \cdot (v_t q_t), \qquad q_1 = \pi_B. \tag{289}$$

Equivalently, the backward path is the same as the forward path:

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = v_t(X_t), \qquad X_1 \sim \pi_B, \tag{290}$$

but running from $t = 1$ to $t = 0$.

However, in the deterministic-flow setting, the forward and backward transition kernels are Dirac measures. Therefore, the path-space RND cannot be computed through a ratio of transition densities, and we can only calculate the work with Equation (282). Therefore, this strategy is typically computationally expensive due to the divergence term in Equation (282).

E.1.2 Diffusion Process in Continuous Euclidean Space

We next instantiate Propositions 3.1 and 3.2 for diffusion-based transport, which recover the stochastic interpolant [Albergo et al., 2025]. Precisely, we use the following process:

$$\mathrm{d}X_t = b_t(X_t)\,\mathrm{d}t + \sqrt{2}\sigma_t\,\overrightarrow{\mathrm{d}W_t}, \qquad X_0 \sim \pi_A. \tag{291}$$

The corresponding generator and adjoint are

$$\mathcal{L}_t^F f = b_t \cdot \nabla f + \sigma_t^2 \Delta f, \qquad (\mathcal{L}_t^F)^\dagger p = -\nabla \cdot (b_t p) + \sigma_t^2 \Delta p. \tag{292}$$

Plugging this generator into Equation (11), we obtain

$$\mathcal{L}_t^B f = \frac{1}{\gamma_t}\left[-\nabla\cdot(b_t\gamma_t f) + \sigma_t^2\Delta(\gamma_t f)\right] - \frac{f}{\gamma_t}\left[-\nabla\cdot(b_t\gamma_t) + \sigma_t^2\Delta\gamma_t\right] \tag{293}$$

$$= -b_t\cdot\nabla f + \sigma_t^2\Delta f + 2\sigma_t^2\nabla\log\gamma_t\cdot\nabla f \tag{294}$$

$$= \left(-b_t - 2\sigma_t^2\nabla U_t\right)\cdot\nabla f + \sigma_t^2\Delta f. \tag{295}$$

Thus, if the forward drift is b_t, the backward generator has drift

$$-b_t - 2\sigma_t^2\nabla U_t. \tag{296}$$

i.e., the diffusion running backward in time is

$$\mathrm{d}X_t = \left(b_t(X_t) + 2\sigma_t^2\nabla U_t\right)\mathrm{d}t + \sqrt{2}\sigma_t\,\overleftarrow{\mathrm{d}W_t}, \qquad X_1 \sim \pi_B. \tag{297}$$

A more symmetric parametrization is to decompose b_t:

$$\mathrm{d}X_t = \left(v_t(X_t) - \sigma_t^2\nabla U_t(X_t)\right)\mathrm{d}t + \sqrt{2}\sigma_t\,\overrightarrow{\mathrm{d}W_t}, \qquad X_0 \sim \pi_A. \tag{298}$$

By the calculation above, the corresponding backward dynamics is

$$\mathrm{d}X_t = \left(v_t(X_t) + \sigma_t^2\nabla U_t(X_t)\right)\mathrm{d}t + \sqrt{2}\sigma_t\,\overleftarrow{\mathrm{d}W_t}, \qquad X_1 \sim \pi_B, \tag{299}$$

We now derive the generalized work for this symmetric choice. The forward generator is

$$\mathcal{L}_t^F f = \left(v_t - \sigma_t^2\nabla U_t\right)\cdot\nabla f + \sigma_t^2\Delta f. \tag{300}$$

and hence

$$\frac{((\mathcal{L}_t^F)^\dagger\gamma_t)(x)}{\gamma_t(x)} \tag{301}$$

$$= -\nabla\cdot\left(v_t(x) - \sigma_t^2\nabla U_t(x)\right) + \left(v_t(x) - \sigma_t^2\nabla U_t(x)\right)\cdot\nabla U_t(x) + \sigma_t^2\left(\|\nabla U_t(x)\|^2 - \Delta U_t(x)\right) \tag{302}$$

$$= -\nabla\cdot v_t(x) + v_t(x)\cdot\nabla U_t(x). \tag{303}$$

and hence the generalized work simplifies to

$$\widetilde{W}(X_{[0,1]}) = \int_0^1 \left[\partial_t U_t(X_t) + v_t(X_t)\cdot\nabla U_t(X_t) - \nabla\cdot v_t(X_t)\right]\mathrm{d}t. \tag{304}$$

Thus, under the symmetric diffusion parametrization, we recover the standard formula of escorted Jarzynski [Vaikuntanathan and Jarzynski, 2008].

E.1.3 Momentum-Augmented Transport in Continuous Space

We next instantiate Propositions 3.1 and 3.2 for momentum-augmented transport. Let $Z_t = (X_t, Y_t) \in \mathbb{R}^d \times \mathbb{R}^d$, and define an augmented energy $H_t(x, y)$ on the joint space, as well as the unnormalized density

$$\Gamma_t(x,y) = e^{-H_t(x,y)}, \quad \Pi_t(x,y) = \frac{e^{-H_t(x,y)}}{\int e^{-H_t(x,y)}\mathrm{d}xy} \tag{305}$$

For two endpoints, we can set $Y_0 \perp X_0, Y_0 \sim \mathcal{N}$ and $Y_0 \perp X_1, Y_1 \sim \mathcal{N}$, so that the free energy between H_0 and H_1 remains the same as the free energy between U_0 and U_1.

Consider the forward momentum-augmented diffusion

$$\mathrm{d}X_t = Y_t\,\mathrm{d}t, \tag{306}$$
$$\mathrm{d}Y_t = B_t(X_t, Y_t)\,\mathrm{d}t + \sqrt{2}\sigma_t\,\overrightarrow{\mathrm{d}W_t}, \qquad (X_0, Y_0) \sim \Pi_A. \tag{307}$$

The corresponding generator and adjoint are

$$\mathcal{L}_t^F f = y \cdot \nabla_x f + B_t \cdot \nabla_y f + \sigma_t^2 \Delta_y f, \tag{308}$$
$$(\mathcal{L}_t^F)^\dagger p = -\nabla_x \cdot (yp) - \nabla_y \cdot (B_t p) + \sigma_t^2 \Delta_y p. \tag{309}$$

Plugging this generator into Equation (11), we obtain

$$\begin{aligned}\mathcal{L}_t^B f = &\frac{1}{\Gamma_t}\left[-\nabla_x \cdot (y\Gamma_t f) - \nabla_y \cdot (B_t \Gamma_t f) + \sigma_t^2 \Delta_y(\Gamma_t f)\right] \\ &- \frac{f}{\Gamma_t}\left[-\nabla_x \cdot (y\Gamma_t) - \nabla_y \cdot (B_t\Gamma_t) + \sigma_t^2 \Delta_y \Gamma_t\right] &(310)\\ = &-y \cdot \nabla_x f - B_t \cdot \nabla_y f + \sigma_t^2 \Delta_y f + 2\sigma_t^2 \nabla_y \log \Gamma_t \cdot \nabla_y f &(311)\\ = &-y \cdot \nabla_x f + \left(-B_t - 2\sigma_t^2 \nabla_y H_t\right) \cdot \nabla_y f + \sigma_t^2 \Delta_y f. &(312)\end{aligned}$$

where $\nabla_y H_t = -\nabla_y \log \Gamma_t$. Therefore, the diffusion running backward in time is

$$\mathrm{d}X_t = Y_t\,\overleftarrow{\mathrm{d}t}, \tag{313}$$
$$\mathrm{d}Y_t = \left(B_t(X_t, Y_t) + 2\sigma_t^2 \nabla_y H_t(X_t, Y_t)\right)\mathrm{d}t + \sqrt{2}\sigma_t\,\overleftarrow{\mathrm{d}W_t}, \qquad (X_1, Y_1) \sim \Pi_B. \tag{314}$$

Similar to the diffusion case, a more symmetric parametrization is to decompose the momentum drift:

$$B_t(x, y) = v_t(x, y) - \sigma_t^2 \nabla_y H_t(x, y). \tag{315}$$

Then the forward dynamics becomes

$$\mathrm{d}X_t = Y_t\,\mathrm{d}t, \tag{316}$$
$$\mathrm{d}Y_t = \left(v_t(X_t, Y_t) - \sigma_t^2 \nabla_y H_t(X_t, Y_t)\right)\mathrm{d}t + \sqrt{2}\sigma_t\,\overrightarrow{\mathrm{d}W_t}, \qquad (X_0, Y_0) \sim \Pi_A. \tag{317}$$

By the calculation above, the corresponding backward dynamics is

$$\mathrm{d}X_t = Y_t\,\overleftarrow{\mathrm{d}t}, \tag{318}$$
$$\mathrm{d}Y_t = \left(v_t(X_t, Y_t) + \sigma_t^2 \nabla_y H_t(X_t, Y_t)\right)\mathrm{d}t + \sqrt{2}\sigma_t\,\overleftarrow{\mathrm{d}W_t}, \qquad (X_1, Y_1) \sim \Pi_B. \tag{319}$$

We now derive the generalized work for this symmetric choice. The forward generator is

$$\mathcal{L}_t^F f = y \cdot \nabla_x f + \left(v_t - \sigma_t^2 \nabla_y H_t\right) \cdot \nabla_y f + \sigma_t^2 \Delta_y f. \tag{320}$$
$$(\mathcal{L}_t^F)^\dagger p = -\nabla_x \cdot (yp) - \nabla_y \cdot \left(\left(v_t - \sigma_t^2 \nabla_y H_t\right) p\right) + \sigma_t^2 \Delta_y p. \tag{321}$$

Therefore, using $\nabla_x \cdot y = 0$, we have

$$\begin{aligned}&\frac{((\mathcal{L}_t^F)^\dagger \Gamma_t)(x, y)}{\Gamma_t(x, y)} &(322)\\ = &-\nabla_x \cdot y + y \cdot \nabla_x H_t(x, y) - \nabla_y \cdot \left(v_t(x, y) - \sigma_t^2 \nabla_y H_t(x, y)\right) \\ &+ \left(v_t(x, y) - \sigma_t^2 \nabla_y H_t(x, y)\right) \cdot \nabla_y H_t(x, y) + \sigma_t^2 \left(\|\nabla_y H_t(x, y)\|^2 - \Delta_y H_t(x, y)\right) &(323)\\ = &\, y \cdot \nabla_x H_t(x, y) - \nabla_y \cdot v_t(x, y) + v_t(x, y) \cdot \nabla_y H_t(x, y), &(324)\end{aligned}$$

Hence the generalized work simplifies to

$$\widetilde{W}(Z_{[0,1]}) = \int_0^1 \left[\partial_t H_t(Z_t) + Y_t \cdot \nabla_x H_t(Z_t) + v_t(Z_t) \cdot \nabla_y H_t(Z_t) - \nabla_y \cdot v_t(Z_t)\right] \mathrm{d}t. \tag{325}$$

Integrator of Momentum-Augmented Transport Unlike SDE without momentum, we now have two parts: one is the ODE for sample position, and the other is the SDE for momentum. We now describe how to discretize these dynamics and how to calculate the forward-backward path RND. Following the notation of Blessing et al. [2025], the dynamics can be split into only two parts:

$$\begin{bmatrix} \mathrm{d}X_t \\ \mathrm{d}Y_t \end{bmatrix} = \underbrace{\begin{bmatrix} Y_t \\ 0 \end{bmatrix} \mathrm{d}t}_{A} + \underbrace{\begin{bmatrix} 0 \\ \big(v_t(X_t, Y_t) - \sigma_t^2 \nabla_y H_t(X_t, Y_t)\big)\,\mathrm{d}t + \sqrt{2}\sigma_t \overrightarrow{\mathrm{d}W_t} \end{bmatrix}}_{O}. \tag{326}$$

A is the deterministic position update, and O is the stochastic momentum update. For compactness, define

$$G_t^F(x, y) := v_t(x, y) - \sigma_t^2 \nabla_y H_t(x, y), \tag{327}$$

Similarly, from the backward dynamics, we can also split it into O and A steps:

$$\begin{bmatrix} \mathrm{d}X_t \\ \mathrm{d}Y_t \end{bmatrix} = \underbrace{\begin{bmatrix} Y_t \\ 0 \end{bmatrix} \overleftarrow{\mathrm{d}t}}_{A} + \underbrace{\begin{bmatrix} 0 \\ \big(v_t(X_t, Y_t) + \sigma_t^2 \nabla_y H_t(X_t, Y_t)\big)\,\mathrm{d}t + \sqrt{2}\sigma_t \overleftarrow{\mathrm{d}W_t} \end{bmatrix}}_{O}. \tag{328}$$

and we also write

$$G_t^B(x, y) := v_t(x, y) + \sigma_t^2 \nabla_y H_t(x, y). \tag{329}$$

Now, let's assume we discretize the time horizon with N boundaries:

$$0 = t_0 < t_1 < \cdots < t_N = 1, \qquad \delta_t = t_{i+1} - t_i, \qquad \sigma_i = \sigma_{t_i}. \tag{330}$$

A forward transition goes from $Z_i = (x_i, y_i)$ to $Z_{i+1} = (x_{i+1}, y_{i+1})$. A backward transition goes from Z_{i+1} back to Z_i. Instead of directly doing the transition for the entire Z space, we split the transition into two parts, one for A and one for O. Depending on the order of the A and O steps, we have the AO or OA integrators:

(1) AO INTEGRATOR. The AO integrator first applies A, then O.

Starting from (x_i, y_i), the A-step is deterministic:

$$x_i' = x_i + \delta_i y_i, \qquad y_i' = y_i. \tag{331}$$

Then the O-step updates only the momentum:

$$y_{i+1} \sim \mathcal{N}\left(y_i' + \delta_i G_{t_i}^F(x_i', y_i'), 2\sigma_i^2 \delta_i I\right), \tag{332}$$

$$x_{i+1} = x_i'. \tag{333}$$

Therefore, the forward AO transition kernel is

$$\overrightarrow{p}_{i+1|i}^{\mathrm{AO}}(Z_{i+1}|Z_i) = \boldsymbol{\delta}_{x_i + \delta_i y_i}(x_{i+1})\,\mathcal{N}\left(y_{i+1} \,\middle|\, y_i' + \delta_i G_{t_i}^F(x_i', y_i'), 2\sigma_i^2 \delta_i I\right). \tag{334}$$

where $\boldsymbol{\delta}$ is the Delta-measure.

(2) OA INTEGRATOR. The OA integrator first applies O, then A.

Starting from (x_i, y_i), the O-step updates only momentum:

$$y_i' \sim \mathcal{N}\left(y_i + \delta_i G_{t_i}^F(x_i, y_i), 2\sigma_i^2 \delta_i I\right), \tag{335}$$

$$x_i' = x_i. \tag{336}$$

Then the A-step updates only position:

$$x_{i+1} = x_i' + \delta_i y_i', \qquad y_{i+1} = y_i'. \tag{337}$$

Therefore, the OA transition kernel is

$$\overrightarrow{p}_{i+1|i}^{\mathrm{OA}}(Z_{i+1}|Z_i) = \boldsymbol{\delta}_{x_i + \delta_i y_{i+1}}(x_{i+1})\,\mathcal{N}\left(y_i + \delta_i G_{t_i}^F(x_i, y_i), 2\sigma_i^2 \delta_i I\right). \tag{338}$$

(3) BACKWARD INTEGRATOR OF AO. The reverse of a forward AO step is a backward OA step. Starting from $Z_{i+1} = (x_{i+1}, y_{i+1})$, we first apply the backward O-step at fixed x_{i+1}:

$$y_i \sim \mathcal{N}\left(y_{i+1} - \delta_i G_{t_{i+1}}^B(x_{i+1}, y_{i+1}), 2\sigma_{i+1}^2 \delta_i I\right). \tag{339}$$

We then invert the deterministic A-step:

$$x_i = x_{i+1} - \delta_i y_i. \tag{340}$$

Therefore, the backward transition kernel is

$$\overleftarrow{p}^{\mathrm{OA}}_{i|i+1}(Z_i|Z_{i+1}) = \boldsymbol{\delta}_{x_{i+1}-\delta_i y_i}(x_i)\,\mathcal{N}\left(y_i \,\middle|\, y_{i+1} - \delta_i G^B_{t_{i+1}}(x_{i+1}, y_{i+1}), 2\sigma^2_{i+1}\delta_i I\right). \tag{341}$$

(4) BACKWARD INTEGRATOR OF OA. The reverse of a forward OA step is a backward AO step. Starting from $Z_{i+1} = (x_{i+1}, y_{i+1})$, we first invert the deterministic A-step:

$$x_i = x_{i+1} - \delta_i y_{i+1}. \tag{342}$$

We then apply the backward O-step at fixed x_i:

$$y_i \sim \mathcal{N}\left(y_{i+1} - \delta_i G^B_{t_{i+1}}(x_i, y_{i+1}), 2\sigma^2_{i+1}\delta_i I\right). \tag{343}$$

Therefore, the backward transition kernel is

$$\overleftarrow{p}^{\mathrm{AO}}_{i|i+1}(Z_i|Z_{i+1}) = \boldsymbol{\delta}_{x_{i+1}-\delta_i y_{i+1}}(x_i)\,\mathcal{N}\left(y_i \,\middle|\, y_{i+1} - \delta_i G^B_{t_{i+1}}(x_i, y_{i+1}), 2\sigma^2_{i+1}\delta_i I\right). \tag{344}$$

DENSITY RATIO BETWEEN AO AND OA. We now consider the density ratio between a forward path discretized by OA and a backward path discretized by AO. The key point is that the deterministic A-step is volume-preserving:

$$(x, y) \mapsto (x + \delta_i y, y) \tag{345}$$

has Jacobian determinant equal to one. Therefore, the A-step contributes no Jacobian correction to the transition-density ratio. All nontrivial density contributions come from the Gaussian O-steps.

For one forward OA step and its backward AO reverse, we have

$$\log \frac{\overrightarrow{p}^{\mathrm{OA}}_{i+1|i}(Z_{i+1}|Z_i)}{\overleftarrow{p}^{\mathrm{AO}}_{i|i+1}(Z_i|Z_{i+1})} = \log \mathcal{N}\left(y_{i+1} \,\middle|\, \mu_i^{F,\mathrm{OA}}, 2\sigma_i^2\delta_i I\right) - \log \mathcal{N}\left(y_i \,\middle|\, \mu_i^{B,\mathrm{AO}}, 2\sigma^2_{i+1}\delta_i I\right), \tag{346}$$

where

$$\mu_i^{F,\mathrm{OA}} = y_i + \delta_i G^F_{t_i}(x_i, y_i) \tag{347}$$

$$\mu_i^{B,\mathrm{AO}} = y_{i+1} - \delta_i G^B_{t_{i+1}}(x_i, y_{i+1}) \tag{348}$$

The discrete path ratio for work is then obtained by summing only the Gaussian log-density differences, together with the endpoint density ratio:

$$\begin{aligned}
&\log \frac{\mathrm{d}\overrightarrow{P}}{\mathrm{d}\overleftarrow{Q}}(Z_{0:N}) \\
&= \log \frac{\Pi_A(Z_0)}{\Pi_B(Z_N)} + \sum_{i=0}^{N-1} \left[\log \mathcal{N}\left(y_{i+1} \,\middle|\, \mu_i^{F,\mathrm{OA}}, 2\sigma_i^2\delta_i I\right) - \log \mathcal{N}\left(y_i \,\middle|\, \mu_i^{B,\mathrm{AO}}, 2\sigma^2_{i+1}\delta_i I\right)\right].
\end{aligned} \tag{349}$$

Momentum, or generally, auxiliary variables have been introduced to build generative models with the goal to introduce more flexible parameterization and faster inference [Dockhorn et al., 2021, Du et al., 2023, Singhal et al., 2023, Blessing et al., 2025, Chen et al., 2023a]. They have also been baked into bridge models for smooth trajectory inference [Chen et al., 2023b, Theodoropoulos et al., 2025].

E.1.4 Continuous-Time Markov Chains on Finite State Spaces

We finally instantiate Propositions 3.1 and 3.2 for finite discrete spaces. Let's define the transition rate matrix $R^F_t(i, j) \geq 0$, for $i \neq j$. We also define the diagonal entries by

$$R^F_t(i, i) = -\sum_{j \neq i} R^F_t(i, j). \tag{350}$$

Then the corresponding generator can be written compactly as

$$\mathcal{L}_t^F f(i) = \sum_{j \in \mathcal{X}} R_t^F(i,j) f(j). \tag{351}$$

The adjoint with respect to counting measure is

$$(\mathcal{L}_t^F)^\dagger p(i) = \sum_{j \in \mathcal{X}} R_t^F(j,i) p(j). \tag{352}$$

Therefore,

$$\frac{((\mathcal{L}_t^F)^\dagger \gamma_t)(i)}{\gamma_t(i)} = \sum_{j \in \mathcal{X}} R_t^F(j,i) \frac{\gamma_t(j)}{\gamma_t(i)} \tag{353}$$

$$= \sum_{j \in \mathcal{X}} R_t^F(j,i) \exp\left(U_t(i) - U_t(j)\right). \tag{354}$$

and the term $\exp\left(U_t(i) - U_t(j)\right)$ are known as the concrete score or discrete score [Lou et al., 2023]. Plugging this into Equation (10), the generalized work becomes

$$\widetilde{W}(X_{[0,1]}) = \int_0^1 \left[\partial_t U_t(X_t) + \sum_{j \in \mathcal{X}} R_t^F(j, X_t) \exp\left(U_t(X_t) - U_t(j)\right) \right] \mathrm{d}t. \tag{355}$$

The corresponding backward generator from Equation (11) can be derived similarly:

$$(\mathcal{L}_t^F)^\dagger(\gamma_t f)(i) = \sum_{j \in \mathcal{X}} R_t^F(j,i) \gamma_t(j) f(j), \qquad (\mathcal{L}_t^F)^\dagger \gamma_t(i) = \sum_{j \in \mathcal{X}} R_t^F(j,i) \gamma_t(j). \tag{356}$$

Hence

$$\mathcal{L}_t^B f(i) = \frac{1}{\gamma_t(i)} \sum_{j \in \mathcal{X}} R_t^F(j,i) \gamma_t(j) f(j) - \frac{f(i)}{\gamma_t(i)} \sum_{j \in \mathcal{X}} R_t^F(j,i) \gamma_t(j) \tag{357}$$

$$= \sum_{j \in \mathcal{X}} R_t^F(j,i) \frac{\gamma_t(j)}{\gamma_t(i)} \left[f(j) - f(i)\right] \tag{358}$$

$$= \sum_{j \in \mathcal{X}} R_t^F(j,i) \exp\left(U_t(i) - U_t(j)\right) \left[f(j) - f(i)\right]. \tag{359}$$

Therefore, the backward process 's rate matrix is given by

$$R_t^B(i,j) = R_t^F(j,i) \frac{\gamma_t(j)}{\gamma_t(i)} = R_t^F(j,i) \exp\left(U_t(i) - U_t(j)\right), i \neq j \tag{360}$$

and

$$R_t^B(i,i) = -\sum_{j \neq i} R_t^B(i,j). \tag{361}$$

The backward marginal equation in Proposition 3.2 is then

$$\partial_t q_t = -(\mathcal{L}_t^B)^\dagger q_t, \qquad q_1 = \pi_B. \tag{362}$$

E.2 Exemplify Learning Methods for the Transport

In this section, we instantiate the learning strategy for several concrete cases. For completeness, we also include the diffusion-process setting. However, since this case has already been studied in [He and Du et al., 2025], we do not repeat the corresponding experiments in this work.

We also emphasize that the learning design is not unique. For example, we can design many different interpolant paths bridging between two distributions. Moreover, even for a fixed interpolant, the generator realizing the same marginal path is generally non-unique. Our goal here is therefore not to provide a systematic study of all possible choices, but rather to illustrate the general recipe through representative examples. Regardless of the specific design, the learning principle discussed in Section 4 remains unchanged and applies generally.

E.2.1 Diffusion Process in Continuous Euclidean Space

We first consider the diffusion process setting in continuous Euclidean space. Following Albergo et al. [2025], we define a stochastic interpolant between endpoint samples $(x_0, x_1) \sim \pi(x_0, x_1)$, with $x_0 \sim \pi_A$ and $x_1 \sim \pi_B$:

$$x_t = \alpha_t x_0 + \beta_t x_1 + \gamma_t \epsilon, \qquad \epsilon \sim \mathcal{N}(0, I), \tag{363}$$

where $\alpha_0 = 1, \beta_0 = 0, \gamma_0 = 0$ and $\alpha_1 = 0, \beta_1 = 1, \gamma_1 = 0$. For fixed (x_0, x_1), this gives the conditional density path

$$p_t(x|x_0, x_1) = \mathcal{N}\left(x \,\middle|\, I_t(x_0, x_1), \gamma_t^2 I\right), \qquad I_t(x_0, x_1) = \alpha_t x_0 + \beta_t x_1. \tag{364}$$

The corresponding conditional velocity and conditional score realizing this path are

$$v_t^{x_0,x_1}(x) = \dot{\alpha}_t x_0 + \dot{\beta}_t x_1 + \frac{\dot{\gamma}_t}{\gamma_t}\left(x - I_t(x_0, x_1)\right), \tag{365}$$

$$s_t^{x_0,x_1}(x) := \nabla_x \log p_t(x|x_0, x_1) = -\frac{x - I_t(x_0, x_1)}{\gamma_t^2}. \tag{366}$$

Now, we will use the two strategies discussed in Section 4 (step 2(A) and step 2(B)) to obtain the learning objective.

Strategy 1: directly learn the forward and backward generators. For each fixed pair (x_0, x_1), the conditional forward and backward generators are given by

$$\mathcal{L}_{F,t}^{x_0,x_1} f = \left(v_t^{x_0,x_1} + \sigma_t^2 s_t^{x_0,x_1}\right) \cdot \nabla f + \sigma_t^2 \Delta f, \tag{367}$$

$$\mathcal{L}_{B,t}^{x_0,x_1} f = \left(-v_t^{x_0,x_1} + \sigma_t^2 s_t^{x_0,x_1}\right) \cdot \nabla f + \sigma_t^2 \Delta f. \tag{368}$$

By construction, these two conditional generators are time reversals with respect to $p_t(x|x_0, x_1)$. Therefore, by Proposition 4.1, their marginalization will also be time reversals.

Therefore, we may parameterize the forward and backward drifts $b_{\psi,t}^F(x)$ and $b_{\phi,t}^B(x)$. The marginal generators are then given by

$$\mathcal{L}_{F,t} f = b_{\psi,t}^F \cdot \nabla f + \sigma_t^2 \Delta f, \tag{369}$$

$$\mathcal{L}_{B,t} f = b_{\phi,t}^B \cdot \nabla f + \sigma_t^2 \Delta f. \tag{370}$$

Matching them to the conditional generators, we obtain the objectives

$$\mathcal{J}_F(\psi) = \mathbb{E}_{t,x_0,x_1,\epsilon}\left[\left\| b_{\psi,t}^F(x_t) - \left(v_t^{x_0,x_1}(x_t) + \sigma_t^2 s_t^{x_0,x_1}(x_t)\right)\right\|^2\right], \tag{371}$$

$$\mathcal{J}_B(\phi) = \mathbb{E}_{t,x_0,x_1,\epsilon}\left[\left\| b_{\phi,t}^B(x_t) - \left(-v_t^{x_0,x_1}(x_t) + \sigma_t^2 s_t^{x_0,x_1}(x_t)\right)\right\|^2\right]. \tag{372}$$

In practice, we can also use the symmetric diffusion parametrization, i.e., parameterize the forward and backward drifts through a shared vector field $v_{\psi,t}$ and a shared score $s_{\theta,t}$. Then the forward and backward generators for the marginal path p_t are

$$\mathcal{L}_t^F f = \left(v_{\psi,t} + \sigma_t^2 s_{\theta,t}\right) \cdot \nabla f + \sigma_t^2 \Delta f, \tag{373}$$

$$\mathcal{L}_t^B f = \left(-v_{\psi,t} + \sigma_t^2 s_{\theta,t}\right) \cdot \nabla f + \sigma_t^2 \Delta f. \tag{374}$$

Matching these two generators to the conditional forward and backward generators gives

$$\begin{aligned}\mathcal{J}_{\text{sym}}(\psi, \theta) = \mathbb{E}_{t,x_0,x_1,\epsilon}\Bigg[&\left\| v_{\psi,t}(x_t) + \sigma_t^2 s_{\theta,t}(x_t) - \left(v_t^{x_0,x_1}(x_t) + \sigma_t^2 s_t^{x_0,x_1}(x_t)\right)\right\|^2 \\ &+ \left\| -v_{\psi,t}(x_t) + \sigma_t^2 s_{\theta,t}(x_t) - \left(-v_t^{x_0,x_1}(x_t) + \sigma_t^2 s_t^{x_0,x_1}(x_t)\right)\right\|^2\Bigg].\end{aligned} \tag{375}$$

Equivalently, this can be written as directly matching the marginal vector field and score:

$$\mathcal{J}_v(\psi) = \mathbb{E}_{t,x_0,x_1,\epsilon}\left[\left\| v_{\psi,t}(x_t) - v_t^{x_0,x_1}(x_t)\right\|^2\right], \tag{376}$$

$$\mathcal{J}_s(\theta) = \mathbb{E}_{t,x_0,x_1,\epsilon}\left[\left\| s_{\theta,t}(x_t) - s_t^{x_0,x_1}(x_t)\right\|^2\right], \tag{377}$$

which recovers the objective by Albergo et al. [2025] and used in [He and Du et al., 2025]. In practice, we may also reweight the objective with time-dependent weights to balance the learning at different times.

Strategy 2: learn one direction and the marginal score. A second strategy is to learn one direction together with the marginal statistic needed to reverse it.

Recall that the Gaussian interpolant can be viewed as first defining the uncorrupted interpolating state

$$I_t = \alpha_t x_0 + \beta_t x_1, \tag{378}$$

and then applying a Gaussian corruption

$$x_t = I_t + \gamma_t \epsilon, \qquad \epsilon \sim \mathcal{N}(0, I). \tag{379}$$

We can view this corruption as a new stochastic process, with a time-independent corruption generator

$$\mathcal{A} = \frac{1}{2}\Delta. \tag{380}$$

For each time t, we will calculate the marginal of this new corruption process with

$$q_\tau(x|I_t) = \mathcal{N}(x|I_t, \gamma_t^2 I). \tag{381}$$

Therefore, the marginal interpolant can be written as

$$p_t(x) = \int \mathcal{N}(x|I_t, \gamma_t^2 I)\, p_t^I(I)\, \mathrm{d}I. \tag{382}$$

Applying the general DSM objective from Section 4, and using $\mathcal{A} = \frac{1}{2}\Delta$, we obtain

$$\frac{\mathcal{A}\left(q_\tau(x_t|I_t)/\varphi_t(x_t)\right)}{q_\tau(x_t|I_t)/\varphi_t(x_t)} - \mathcal{A}\log\left(q_\tau(x_t|I_t)/\varphi_t(x_t)\right) = \frac{1}{2}\left\|\nabla_x \log q_\tau(x_t|I_t) - \nabla_x \log \varphi_t(x_t)\right\|^2. \tag{383}$$

Thus, if we parameterize

$$s_{\theta,t}(x) \approx \nabla_x \log \varphi_t(x), \tag{384}$$

Then the DSM objective reduces to the usual denoising score-matching objective

$$\mathcal{J}_s(\theta) = \mathbb{E}_{t,x_0,x_1,\epsilon}\left[\left\|s_{\theta,t}(x_t) - \nabla_x \log \mathcal{N}(x|I_t, \gamma_t^2 I)\right\|^2\right]. \tag{385}$$

In practice, we may also reweight the objective with time-dependent weights to balance the learning at different times. At the optimum, this gives the marginal score

$$s_{\theta,t}(x) = \nabla_x \log p_t(x). \tag{386}$$

With the score being learned, we then only need to learn one direction, either the forward drift or the backward drift. Similar to the first strategy, we can either parameterize the entire forward drift b_t^F and obtain the backward drift as

$$b_t^B(x) = -b_t^F(x) + 2\sigma_t^2 s_{\theta,t}(x), \tag{387}$$

or parameterize a vector field v_t and recover the forward and backward drifts by

$$b_t^F(x) = v_t(x) + \sigma_t^2 s_{\theta,t}(x), \tag{388}$$

$$b_t^B(x) = -v_t(x) + \sigma_t^2 s_{\theta,t}(x), \tag{389}$$

where $s_{\theta,t}$ denotes the learned marginal score.

Additionally, note that the vector field and score do not depend on σ_t, and the above discussion holds for any $\sigma_t > 0$. Therefore, after learning v_t and $s_{\theta,t}$, we can tune the value of σ_t, which can yield different performance.

E.2.2 Momentum-Augmented Transport in Continuous Space

We now explain the training strategy for momentum-augmented transport. Recall from Section E.1.3: Let $Z_t = (X_t, Y_t) \in \mathbb{R}^d \times \mathbb{R}^d$, and define an augmented energy $H_t(x, y)$ on the joint space, as well as the unnormalized density

$$\Gamma_t(x, y) = e^{-H_t(x,y)}, \quad \Pi_t(x, y) = \frac{e^{-H_t(x,y)}}{\int e^{-H_t(x,y)}\mathrm{d}xy} \tag{390}$$

Also recall, that under the symmetric parametrization, the forward and backward dynamics is

$$\mathrm{d}X_t = Y_t\,\mathrm{d}t, \tag{391}$$

$$\mathrm{d}Y_t = \big(v_t(X_t, Y_t) - \sigma_t^2 \nabla_y H_t(X_t, Y_t)\big)\,\mathrm{d}t + \sqrt{2}\sigma_t\,\overrightarrow{\mathrm{d}W_t}, \qquad (X_0, Y_0) \sim \Pi_A. \tag{392}$$

and

$$\mathrm{d}X_t = Y_t\,\overleftarrow{\mathrm{d}t}, \tag{393}$$

$$\mathrm{d}Y_t = \big(v_t(X_t, Y_t) + \sigma_t^2 \nabla_y H_t(X_t, Y_t)\big)\,\mathrm{d}t + \sqrt{2}\sigma_t\,\overleftarrow{\mathrm{d}W_t}, \qquad (X_1, Y_1) \sim \Pi_B. \tag{394}$$

Similar to the overdamped diffusion process case, we learn the score $s_t^y(x, y) = \nabla_y \log \Gamma_t(x, y) = -\nabla_y H_t(x, y)$, and the vector field v_t to obtain the forward and backward processes. The difference is that the vector field is now an acceleration field in the momentum equation, and the required score is only the momentum score $\nabla_y \log \Gamma_t(x, y)$, rather than the full phase-space score.

Define the Stochastic Interpolant with Momentum To achieve this, we still consider constructing an interpolant between samples x_0 and x_1. Different from the standard stochastic interpolant, we augment the space with two endpoint momenta:

$$x_0 \sim \pi_A, \qquad x_1 \sim \pi_B, \qquad y_0 \sim \mathcal{N}(0, I), \qquad y_1 \sim \mathcal{N}(0, I), \tag{395}$$

with y_0, y_1 independent of the positions. This gives the endpoint laws as

$$\Gamma_A(x, y) = \pi_A(x)\mathcal{N}(y|0, I), \qquad \Gamma_B(x, y) = \pi_B(x)\mathcal{N}(y|0, I). \tag{396}$$

Different from the standard stochastic interpolant, we need an interpolant for both x_t and y_t, so that $\dot{x}_t = y_t$. This constrains our design space to a large extent.

To this end, we define a cubic Hermite interpolant. Let

$$h_{00}(t) = 2t^3 - 3t^2 + 1, \quad h_{10}(t) = t^3 - 2t^2 + t, \quad h_{01}(t) = -2t^3 + 3t^2, \quad h_{11}(t) = t^3 - t^2. \tag{397}$$

We then define

$$x_t = h_{00}(t)x_0 + h_{10}(t)y_0 + h_{01}(t)x_1 + h_{11}(t)y_1, \quad y_t = \dot{x}_t. \tag{398}$$

Additionally, to learn the vector field for y, we also need to calculate the time derivative for y, leading to the acceleration:

$$a_t = \dot{y}_t = \ddot{x}_t. \tag{399}$$

More explicitly, we have

$$y_t = (6t^2 - 6t)x_0 + (3t^2 - 4t + 1)y_0 + (-6t^2 + 6t)x_1 + (3t^2 - 2t)y_1, \tag{400}$$

$$a_t = (12t - 6)x_0 + (6t - 4)y_0 + (-12t + 6)x_1 + (6t - 2)y_1. \tag{401}$$

One can verify that:

$$x_{t=0} = x_0, \qquad y_{t=0} = y_0, \qquad x_{t=1} = x_1, \qquad y_{t=1} = y_1, \tag{402}$$

which ensures that, at both end, the momentum and position are independent, following the law defined in Equation (396).

Learning Momentum-Augmented Transport with Stochastic Interpolant We now derive how to learn the two quantities needed for the momentum-augmented dynamics: the acceleration field v_t and the momentum score s_t^y. For simplicity, we use strategy 1, where we first write down the generator condition on the samples and marginalize them.

More precisely, for a fixed endpoint pair (x_0, x_1), let $\rho_t(z|x_0, x_1)$ denote the conditional law of $Z_t = (X_t, Y_t)$, where the randomness comes from the endpoint momenta y_0, y_1. Then, we write down the forward and backward dynamics (and their generators) that realize this path law and are time reversals of each other.

We define the conditional acceleration field and the conditional momentum score as follows:

$$v_t^{x_0,x_1}(z) = a(x_0, x_1, z), \tag{403}$$

$$s_t^{y,x_0,x_1}(z) = \nabla_y \log \rho_t(z|x_0, x_1). \tag{404}$$

where $a(x_0, x_1, z)$ denote the function to solve the acceleration from x_0, x_1 and $z = (x_t, y_t)$. This is possible: for fixed (x_0, x_1), the map from (y_0, y_1) to (x_t, y_t) is linear and invertible. After solving (y_0, y_1), we can obtain $a = \ddot{h}_{00}(t)x_0 + \ddot{h}_{10}(t)y_0 + \ddot{h}_{01}(t)x_1 + \ddot{h}_{11}(t)y_1$. This is similar to the previous stochastic case in Equation (365).

For simplicity, we will later write y_0, y_1 solved from $z = (x_t, y_t)$ as $y_0(z), y_1(z)$. With this notation, we have

$$v_t^{x_0,x_1}(z) = \ddot{h}_{00}(t)x_0 + \ddot{h}_{10}(t)y_0(z) + \ddot{h}_{01}(t)x_1 + \ddot{h}_{11}(t)y_1(z) \tag{405}$$

$$s_t^{y,x_0,x_1}(z) = \nabla_y \log \rho_t(z|x_0, x_1) \tag{406}$$

$$= \nabla_y \log \mathcal{N}(y_0(z)) + \nabla_y \log \mathcal{N}(y_1(z)) \tag{407}$$

$$= \nabla_y \left(-\frac{1}{2}\|y_0(z)\|^2 - \frac{1}{2}\|y_1(z)\|^2 \right) \tag{408}$$

$$= -\nabla_y y_0(z) \cdot y_0(z) - \nabla_y y_1(z) \cdot y_1(z). \tag{409}$$

For $z = (x, y)$, the endpoint momenta y_0, y_1 can be solved from

$$\begin{bmatrix} x - h_{00}(t)x_0 - h_{01}(t)x_1 \\ y - \dot{h}_{00}(t)x_0 - \dot{h}_{01}(t)x_1 \end{bmatrix} = \begin{bmatrix} h_{10}(t) & h_{11}(t) \\ \dot{h}_{10}(t) & \dot{h}_{11}(t) \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \end{bmatrix}. \tag{410}$$

Since

$$h_{10}(t)\dot{h}_{11}(t) - h_{11}(t)\dot{h}_{10}(t) = -t^2(1-t)^2, \tag{411}$$

we obtain

$$y_0(z) = \frac{\dot{h}_{11}(t)\left[x - h_{00}(t)x_0 - h_{01}(t)x_1\right] - h_{11}(t)\left[y - \dot{h}_{00}(t)x_0 - \dot{h}_{01}(t)x_1\right]}{-t^2(1-t)^2}, \tag{412}$$

$$y_1(z) = \frac{-\dot{h}_{10}(t)\left[x - h_{00}(t)x_0 - h_{01}(t)x_1\right] + h_{10}(t)\left[y - \dot{h}_{00}(t)x_0 - \dot{h}_{01}(t)x_1\right]}{-t^2(1-t)^2}. \tag{413}$$

and hence

$$\nabla_y y_0(z) = \frac{h_{11}(t)}{t^2(1-t)^2} I = \frac{1}{t-1} \tag{414}$$

$$\nabla_y y_1(z) = -\frac{h_{10}(t)}{t^2(1-t)^2} I = -\frac{1}{t} \tag{415}$$

Hence

$$s_t^{y,x_0,x_1}(z) = -\nabla_y y_0(z) \cdot y_0(z) - \nabla_y y_1(z) \cdot y_1(z) = \frac{1}{1-t} y_0(z) + \frac{1}{t} y_1(z) \tag{416}$$

Similarly, since $y_0, y_1 \sim \mathcal{N}(0, I)$, the conditional density $\rho_t(z|x_0, x_1)$ is obtained by change of variables from the Gaussian density of (y_0, y_1). Differentiating this density with respect to y, we obtain the conditional momentum score

$$s_t^{y,x_0,x_1}(z) := \nabla_y \log \rho_t(z|x_0, x_1) \tag{417}$$

$$= \frac{y_0(z)}{1-t} + \frac{y_1(z)}{t}. \tag{418}$$

The conditional forward dynamics is defined by

$$\mathrm{d}X_t = Y_t \,\mathrm{d}t, \tag{419}$$

$$\mathrm{d}Y_t = \left(v_t^{x_0,x_1}(X_t, Y_t) + \sigma_t^2 s_t^{y,x_0,x_1}(X_t, Y_t)\right)\mathrm{d}t + \sqrt{2}\sigma_t \overrightarrow{\mathrm{d}W_t}. \tag{420}$$

The corresponding conditional generator is

$$\mathcal{L}_{F,t}^{x_0,x_1} f = y \cdot \nabla_x f + \left(v_t^{x_0,x_1} + \sigma_t^2 s_t^{y,x_0,x_1}\right) \cdot \nabla_y f + \sigma_t^2 \Delta_y f. \tag{421}$$

Also, the conditional backward dynamics, which is the time-reversal of the forward process, is

$$\mathrm{d}X_t = Y_t \overleftarrow{\mathrm{d}t}, \tag{422}$$

$$\mathrm{d}Y_t = \left(v_t^{x_0,x_1}(X_t,Y_t) - \sigma_t^2 s_t^{y,x_0,x_1}(X_t,Y_t)\right)\mathrm{d}t + \sqrt{2}\sigma_t \overleftarrow{\mathrm{d}W_t}. \tag{423}$$

with generator

$$\mathcal{L}_{B,t}^{x_0,x_1} f = -y\cdot\nabla_x f + \left(-v_t^{x_0,x_1} + \sigma_t^2 s_t^{y,x_0,x_1}\right)\cdot\nabla_y f + \sigma_t^2 \Delta_y f. \tag{424}$$

Similar to the stochastic interpolant case, we learn to marginalize these generators. We parameterize the acceleration field $v_{\psi,t}$ and the momentum score $s_{\theta,t}^y$. Then the learned generators are

$$\mathcal{L}_{\psi,\theta,t}^F f = y\cdot\nabla_x f + \left(v_{\psi,t} + \sigma_t^2 s_{\theta,t}^y\right)\cdot\nabla_y f + \sigma_t^2\Delta_y f, \tag{425}$$

$$\mathcal{L}_{\psi,\theta,t}^B f = -y\cdot\nabla_x f + \left(-v_{\psi,t} + \sigma_t^2 s_{\theta,t}^y\right)\cdot\nabla_y f + \sigma_t^2\Delta_y f. \tag{426}$$

Matching these to the conditional forward and backward generators gives

$$\mathcal{J}_{\mathrm{sym}}(\psi,\theta) \tag{427}$$

$$\begin{aligned}=&\mathbb{E}_{t,x_0,x_1,y_0,y_1}\Bigg[\left\|v_{\psi,t}(x_t,y_t) + \sigma_t^2 s_{\theta,t}^y(x_t,y_t) - \left(v_t^{x_0,x_1}(x_t,y_t) + \sigma_t^2 s_t^{y,x_0,x_1}(x_t,y_t)\right)\right\|^2 \\ &+ \left\|-v_{\psi,t}(x_t,y_t) + \sigma_t^2 s_{\theta,t}^y(x_t,y_t) - \left(-v_t^{x_0,x_1}(x_t,y_t) + \sigma_t^2 s_t^{y,x_0,x_1}(x_t,y_t)\right)\right\|^2\Bigg].\end{aligned} \tag{428}$$

Equivalently:

$$\mathcal{J}_v(\psi) = \mathbb{E}_{t,x_0,x_1,y_0,y_1}\left[\left\|v_{\psi,t}(x_t,y_t) - v_t^{x_0,x_1}(x_t,y_t)\right\|^2\right], \tag{429}$$

$$\mathcal{J}_s(\theta) = \mathbb{E}_{t,x_0,x_1,y_0,y_1}\left[\left\|s_{\theta,t}^y(x_t,y_t) - s_t^{y,x_0,x_1}(x_t,y_t)\right\|^2\right]. \tag{430}$$

Since the conditional score target blows up near $t=0$ and $t=1$ (see Equation (417)), we scale the target and the score network both by $t(1-t)$:

$$\mathcal{J}_s(\theta) = \mathbb{E}_{t,x_0,x_1,y_0,y_1}\left[\left\|t(1-t)s_{\theta,t}^y(x_t,y_t) - (ty_0 + (1-t)y_1)\right\|^2\right]. \tag{431}$$

Finally, for free energy estimation, since $N(0,I)$ are normalized, we directly estimate the free energy between the augmented distributions:

$$\Pi_A(x,y) = \pi_A(x)\mathcal{N}(y|0,I), \qquad \Pi_B(x,y) = \pi_B(x)\mathcal{N}(y|0,I). \tag{432}$$

The work is calculated by the forward-backward RND between OA and AO discretized path, as we discussed in Equation (349).

Also, similar to the diffusion case without momentum, the vector field and score do not depend on σ_t, and the above discussion holds for any $\sigma_t > 0$. Therefore, we can also tune different σ_t to improve performance further. This is studied in Figure 4.

E.2.3 Continuous-Time Markov Chains on Finite State Spaces

We now discuss how to learn transport in finite discrete spaces. As in the previous examples, we first define a conditional interpolant between two endpoint samples $(x_0,x_1)\sim\pi(x_0,x_1)$, and then construct conditional generators whose marginalization gives a generator for the marginal path.

Interpolant Design Different from the continuous cases, we design the interpolant by a mixture, as defined by Equation (22). We restate here for easier reference:

$$x_t \sim \alpha_t\delta_{x_0}(x_t) + \beta_t\delta_{x_1}(x_t), \tag{433}$$

Another opinion is to add "noise", similar to the continuous case. The "noise" can be a mask token, or a uniform distribution over the entire vocabulary. Here, we consider mask token for simplicity:

$$x_t \sim \alpha_t\delta_{x_0}(x_t) + \beta_t\delta_{x_1}(x_t) + \gamma_t\delta_M(x_t) \tag{434}$$

Interestingly, in CTMC, both choices are valid and allow us to calculate the forward backward path RND. This is different from the continuous case, where we need to add noise to the path, otherwise, we only learn a flow, which involves Delta measure over the path and cannot be handled properly with a transition kernel. In CTMC, even the interpolant without adding mask still involves the Categorical distribution as the transition kernel, and hence the path always has stochasticity.

We now discuss three design choices for the dynamics realizing the interpolants above. The first design is based on the path without adding mask, while the second and third designs involve mask.

For easier reference, we restate the rate matrix and generator definitation: throughout this section, we use the convention

$$\mathcal{L}_t f(x) = \sum_{y \in \mathcal{X}} R_t(x, y) f(y), \tag{435}$$

where $R_t(x, y)$ denotes the rate of jumping from x to y, and

$$R_t(x, x) = -\sum_{y \neq x} R_t(x, y). \tag{436}$$

"Latent" Generator and Marginalization Before discussing the individual cases, we first state the following proposition that allows us to marginalize a conditional CTMC defined on a "latent" state space.

In particular, the mixture interpolant induces a "latent" CTMC on a finite state space $\mathcal{S}$. This "latent" CTMC is not defined directly on the original data space $\mathcal{X}$. Rather, each state $s \in \mathcal{S}$ specifies the source (e.g., x_0 or x_1 or M) of the current observation x_t: for example, whether x_t is copied from x_0, copied from x_1, or replaced by a mask state. The observed process is therefore a projection of this latent CTMC onto the original space. This latent generator allows us to simplify the conditional generator. The following proposition shows that, by averaging the conditional generator of the latent CTMC with respect to the posterior distribution over latent states, one obtains a marginalized generator acting directly on the observed space.

Proposition E.1 (Marginalization of latent conditional CTMC)**.** *We define a latent CTMC $S_t \in \mathcal{S}$ with generator*

$$\mathcal{A}_t g(s) = \sum_{s' \in \mathcal{S}} A_t(s, s') g(s'), \tag{437}$$

where

$$A_t(s, s) = -\sum_{s' \neq s} A_t(s, s'). \tag{438}$$

Let the observed interpolant be obtained by a deterministic decoding map

$$X_t = \phi(x_0, x_1, S_t). \tag{439}$$

Define the marginalized generator as

$$\mathcal{L}_t f(x) = \mathbb{E}_{s_t, x_0, x_1 | X_t = x} \left[\mathcal{A}_t \left(f \circ \phi(x_0, x_1, \cdot) \right)(s_t) \right]. \tag{440}$$

and we have

$$\partial_t p_t = \mathcal{L}_t^\dagger p_t, \quad p_t = \sum_{x_0, x_1} \pi(x_0, x_1) p_t(x_t | x_0, x_1) \tag{441}$$

Equivalently, we have the rate matrix form:

$$R_t(x, y) = \mathbb{E}_{s_t, x_0, x_1 | X_t = x} \left[\sum_{s' \in \mathcal{S}} \mathbf{1}\{\phi(x_0, x_1, s') = y\} A_t(s_t, s') \right]. \tag{442}$$

and

$$\partial_t p_t(y) = \sum_{x \in \mathcal{X}} R_t(x, y) p_t(x). \tag{443}$$

Proof.

$$\sum_{x\in\mathcal{X}} R_t(x,y)p_t(x) = \sum_{x\in\mathcal{X}} \mathbb{E}_{s_t,x_0,x_1|X_t=x}\left[\sum_{s'\in\mathcal{S}} \mathbf{1}\{\phi(x_0,x_1,s')=y\}A_t(s_t,s')\right] p_t(x) \tag{444}$$

$$= \mathbb{E}_{s_t,x_0,x_1,X_t=x}\left[\sum_{s'\in\mathcal{S}} \mathbf{1}\{\phi(x_0,x_1,s')=y\}A_t(s_t,s')\right] \tag{445}$$

$$= \mathbb{E}_{s_t,x_0,x_1}\left[\sum_{s'\in\mathcal{S}} \mathbf{1}\{\phi(x_0,x_1,s')=y\}A_t(s_t,s')\right] \tag{446}$$

$$= \sum_{s'\in\mathcal{S}} \mathbb{E}_{s_t,x_0,x_1}\left[\mathbf{1}\{\phi(x_0,x_1,s')=y\}A_t(s_t,s')\right] \tag{447}$$

$$= \sum_{s'}\sum_{s_t} \mathbb{E}_{x_0,x_1|s_t}\left[p_t(s_t)\mathbf{1}\{\phi(x_0,x_1,s')=y\}A_t(s_t,s')\right] \tag{448}$$

$$\overset{x_0,\,x_1 \perp s_t}{=} \sum_{s'}\sum_{s_t} \mathbb{E}_{x_0,x_1}\left[p_t(s_t)\mathbf{1}\{\phi(x_0,x_1,s')=y\}A_t(s_t,s')\right] \tag{449}$$

$$= \sum_{s'} \mathbb{E}_{x_0,x_1}\left[\sum_{s_t} p_t(s_t)\mathbf{1}\{\phi(x_0,x_1,s')=y\}A_t(s_t,s')\right] \tag{450}$$

$$= \sum_{s'} \mathbb{E}_{x_0,x_1}\left[\sum_{s_t} p_t(s_t)A_t(s_t,s')\mathbf{1}\{\phi(x_0,x_1,s')=y\}\right] \tag{451}$$

$$= \sum_{s'} \mathbb{E}_{x_0,x_1}\left[\partial_t p_t(s')\mathbf{1}\{\phi(x_0,x_1,s')=y\}\right] \tag{452}$$

$$= \partial_t \sum_{s'} \mathbb{E}_{x_0,x_1}\left[p_t(s')\mathbf{1}\{\phi(x_0,x_1,s')=y\}\right] \tag{453}$$

$$= \partial_t p_t(y) \tag{454}$$

□

We also have a similar result as Proposition 4.1 for the latent CTMC case.

Proposition E.2 (Marginalization in latent space maintains time-reversal)**.** *Let $S_t \in \mathcal{S}$ be a latent CTMC with marginal law $p_t(s)$. Suppose A_t and B_t are time reversals with respect to $p_t(s)$, i.e. for $s \neq s'$,*

$$B_t(s,s') = A_t(s',s)\frac{p_t(s')}{p_t(s)}. \tag{455}$$

Let the observed interpolant be obtained by a deterministic decoding map

$$X_t = \phi(x_0,x_1,S_t), \tag{456}$$

and denote the marginal law of X_t by $p_t(x)$.
Define the marginalized forward and backward rate matrices by

$$R_t^F(x,y) = \mathbb{E}_{s_t,x_0,x_1|X_t=x}\left[\sum_{s'\in\mathcal{S}} \mathbf{1}\{\phi(x_0,x_1,s')=y\}A_t(s_t,s')\right], \tag{457}$$

$$R_t^B(x,y) = \mathbb{E}_{s_t,x_0,x_1|X_t=x}\left[\sum_{s'\in\mathcal{S}} \mathbf{1}\{\phi(x_0,x_1,s')=y\}B_t(s_t,s')\right]. \tag{458}$$

Then R_t^F and R_t^B are time reversals with respect to $p_t(x)$. That is, for $x \neq y$,

$$R_t^B(x,y) = R_t^F(y,x)\frac{p_t(y)}{p_t(x)}. \tag{459}$$

Proof.

$$R_t^F(y,x)p_t(y) = \mathbb{E}_{s_t,x_0,x_1|X_t=y}\left[\sum_{s'\in\mathcal{S}} \mathbf{1}\{\phi(x_0,x_1,s')=x\}A_t(s_t,s')\right]p_t(y) \tag{460}$$

$$= \sum_{s'\in\mathcal{S}} \mathbb{E}_{x_0,x_1}\left[\sum_{s_t\in\mathcal{S}} p_t(s_t)\mathbf{1}\{\phi(x_0,x_1,s_t)=y\}\mathbf{1}\{\phi(x_0,x_1,s')=x\}A_t(s_t,s')\right]. \tag{461}$$

For backward kernel:

$$R_t^B(x,y)p_t(x) = \mathbb{E}_{s_t,x_0,x_1|X_t=x}\left[\sum_{s'\in\mathcal{S}} \mathbf{1}\{\phi(x_0,x_1,s')=y\}B_t(s_t,s')\right]p_t(x) \tag{462}$$

$$= \sum_{s'\in\mathcal{S}} \mathbb{E}_{x_0,x_1}\left[\sum_{s_t\in\mathcal{S}} p_t(s_t)\mathbf{1}\{\phi(x_0,x_1,s_t)=x\}\mathbf{1}\{\phi(x_0,x_1,s')=y\}B_t(s_t,s')\right] \tag{463}$$

$$= \sum_{s_t\in\mathcal{S}} \mathbb{E}_{x_0,x_1}\left[\sum_{s'\in\mathcal{S}} p_t(s')\mathbf{1}\{\phi(x_0,x_1,s')=x\}\mathbf{1}\{\phi(x_0,x_1,s_t)=y\}B_t(s,s_t)\right]. \tag{464}$$

The last equality is just relabeling s and s'.

Since A_t and B_t are time reversals :

$$p_t(s)A_t(s,s') = p_t(s')B_t(s',s). \tag{465}$$

$$R_t^B(x,y)p_t(x) = R_t^F(y,x)p_t(y) \tag{466}$$

□

With these two propositions, we can simply design the generator in this latent space and then marginalize it to obtain the generator in the original data space.

We now consider three cases, from simple to slightly more complicated designs.

Case 1: no mask state and one-way jumps from x_0 to x_1. We first consider the simplest interpolant, with no mask state.

$$x_t \sim \alpha_t\delta_{x_0}(x_t) + \beta_t\delta_{x_1}(x_t). \tag{467}$$

We choose $\alpha_0 = 1, \beta_0 = 0$ and $\alpha_1 = 0, \beta_1 = 1$, with α_t decreasing and β_t increasing.

Let the latent state space be

$$\mathcal{S} = \{0,1\}, \tag{468}$$

with

$$\mathbb{P}(S_t=0) = \alpha_t, \qquad \mathbb{P}(S_t=1) = \beta_t, \qquad \alpha_t+\beta_t = 1. \tag{469}$$

and the decoding map is

$$\phi(x_0,x_1,0) = x_0, \qquad \phi(x_0,x_1,1) = x_1. \tag{470}$$

A conditional forward latent generator in matrix form is as follows. Under our convention, rows are source states and columns are target states,

$$A_t = \begin{bmatrix} \dot\alpha_t/\alpha_t & -\dot\alpha_t/\alpha_t \\ 0 & 0 \end{bmatrix}. \tag{471}$$

We can verify the forward Kolmogorov equation in latent space $\mathcal{S}$:

$$\partial_t \begin{bmatrix} \alpha_t \\ \beta_t \end{bmatrix} = \begin{bmatrix} \dot{\alpha}_t/\alpha_t & -\dot{\alpha}_t/\alpha_t \\ 0 & 0 \end{bmatrix}^\top \begin{bmatrix} \alpha_t \\ \beta_t \end{bmatrix} \tag{472}$$

Similarly, we have the backward generator:

$$B_t = \begin{bmatrix} 0 & 0 \\ \dot{\beta}_t/\beta_t & -\dot{\beta}_t/\beta_t \end{bmatrix}. \tag{473}$$

Using Equations (440) and (442), we can derive the marginal generator:

$$R_t^F(x,y) = \mathbb{E}_{s_t,x_0,x_1|X_t=x}\left[\sum_{s'\in\mathcal{S}} \mathbf{1}\{\phi(x_0,x_1,s')=y\} A_t(s_t,s')\right] \tag{474}$$

For $y \neq x$, the only transition that has non-zero entry in A is from x_0-state to the x_1-state, with rate

$$A_t(0,1) = -\frac{\dot{\alpha}_t}{\alpha_t}. \tag{475}$$

Therefore, for $y \neq x$,

$$\begin{aligned}
R_t^F(x,y) &= \mathbb{E}_{s_t,x_0,x_1|X_t=x}\left[\mathbf{1}\{\phi(x_0,x_1,1)=y\} A_t(s_t,1)\right] && (476)\\
&= \mathbb{E}_{s_t,x_0,x_1|X_t=x}\left[\mathbf{1}\{x_1=y\}\mathbf{1}\{s_t=0\}\left(-\frac{\dot{\alpha}_t}{\alpha_t}\right)\right] && (477)\\
&= \mathbb{E}_{s_t,x_0,x_1|X_t=x}\left[\mathbf{1}\{x_1=y\}\mathbf{1}\{s_t=0\}\left(-\frac{\dot{\alpha}_t}{\alpha_t}\right)\right] && (478)\\
&\quad + \underbrace{\mathbb{E}_{s_t,x_0,x_1|X_t=x}\left[\mathbf{1}\{x_1=y\}\mathbf{1}\{s_t=1\}\left(-\frac{\dot{\alpha}_t}{\alpha_t}\right)\right]}_{=0 \text{ due to } y\neq x} && (479)\\
&= \mathbb{E}_{s_t,x_0,x_1|X_t=x}\left[\mathbf{1}\{x_1=y\}\left(-\frac{\dot{\alpha}_t}{\alpha_t}\right)\right] && (480)\\
&= \mathbb{E}_{x_1|X_t=x}\left[\mathbf{1}\{x_1=y\}\left(-\frac{\dot{\alpha}_t}{\alpha_t}\right)\right] && (481)\\
&= -\frac{\dot{\alpha}_t}{\alpha_t}\mathbb{P}(x_1=y \mid X_t=x) && (482)\\
&= -\frac{\dot{\alpha}_t}{\alpha_t} p_{1|t}(y|x). && (483)
\end{aligned}$$

The diagonal entry is then fixed by the generator row-sum condition:

$$\begin{aligned}
R_t^F(x,x) &= -\sum_{y\neq x} R_t^F(x,y) && (484)\\
&= \frac{\dot{\alpha}_t}{\alpha_t}\sum_{y\neq x} p_{1|t}(y|x) && (485)\\
&= \frac{\dot{\alpha}_t}{\alpha_t}\left(1 - p_{1|t}(x|x)\right) && (486)\\
&= -\frac{\dot{\alpha}_t}{\alpha_t}\left(p_{1|t}(x|x) - 1\right). && (487)
\end{aligned}$$

Combining the off-diagonal and diagonal cases gives the marginal forward generator is

$$R_t^F(x,y) = -\frac{\dot{\alpha}_t}{\alpha_t}\left[p_{1|t}(y|x) - \delta_x(y)\right], \tag{488}$$

Similarly, the marginal backward generator is

$$R_t^B(x,y) = \frac{\dot{\beta}_t}{\beta_t}\left[p_{0|t}(y|x) - \delta_x(y)\right], \tag{489}$$

In practice, this case suggests a simple parameterization. We can directly parameterize $q^F_{\theta,t}(\cdot|x)$, $q^B_{\phi,t}(\cdot|x)$ for posterior endpoint distribution $p_{1|t}(\cdot|x)$, $p_{0|t}(\cdot|x)$. Then learn the generator by Cross-entropy loss:

$$\mathcal{J}_F(\theta) = \mathbb{E}_{t,x_0,x_1,s_t}\left[-\log q^F_{\theta,t}(x_1|x_t)\right], \tag{490}$$

$$\mathcal{J}_B(\phi) = \mathbb{E}_{t,x_0,x_1,s_t}\left[-\log q^B_{\phi,t}(x_0|x_t)\right]. \tag{491}$$

Note that this case coincides with the discrete flow matching formulation of Gat et al. [2024, DFM]. However, DFM mainly considers one transport direction in its experiments, and does not make explicit that the two directional formulations are time reversals of one another.

DFM also proposes a more general construction for the forward and backward processes, involving mixtures with additional components [Gat et al., 2024, Theorem 3, Eq. 21]. However, the forward and backward processes designed by this construction are NOT time-reversal of each other in general. This is not an issue for the application in generative modeling, as we only care about the endpoint marginal of the transport. However, for our application, we require the forward and backward *path* to be time-reversal in optimality to reduce the variance of the estimator. Therefore, the form of DFM is not general enough for our usage.

Our design framework is hence more general and stronger: it gives a concrete recipe for designing forward and backward discrete dynamics that are explicit time reversals of one another. In the following case 2, we consider the second case where we introduce noise, and we can see that this is distinct from the form of DFM.

Before considering the second case, we also note that the objective by Equation (490) is not the only way to learn the forward and backward rate matrices. We can also directly use generator matching and the conclusion of Proposition E.2 allows us to directly marginalize the latent forward and backward generators.

Precisely, for a sampled tuple (t, x_0, x_1, s_t), define the sample-level forward and backward rate targets

$$\widehat{R}_t^F(y|x_t,x_0,x_1,s_t) := \sum_{s'\in\mathcal{S}} \mathbf{1}\{\phi(x_0,x_1,s') = y\}A_t(s_t,s'), \tag{492}$$

$$\widehat{R}_t^B(y|x_t,x_0,x_1,s_t) := \sum_{s'\in\mathcal{S}} \mathbf{1}\{\phi(x_0,x_1,s') = y\}B_t(s_t,s'). \tag{493}$$

Then

$$R_t^F(x,y) = \mathbb{E}\left[\widehat{R}_t^F(y|x_t,x_0,x_1,s_t)\,\middle|\, X_t = x\right], \tag{494}$$

and similarly for R_t^B.

The generator-matching loss following [Holderrieth et al., 2024] is then

$$\mathcal{J}_F(\theta) = \mathbb{E}_{t,x_0,x_1,s_t}\left[\sum_{y\neq x_t}\left(R^F_{\theta,t}(x_t,y) - \widehat{R}_t^F(y|x_t,x_0,x_1,s_t)\log R^F_{\theta,t}(x_t,y)\right)\right], \tag{495}$$

and

$$\mathcal{J}_B(\phi) = \mathbb{E}_{t,x_0,x_1,s_t}\left[\sum_{y\neq x_t}\left(R^B_{\phi,t}(x_t,y) - \widehat{R}_t^B(y|x_t,x_0,x_1,s_t)\log R^B_{\phi,t}(x_t,y)\right)\right]. \tag{496}$$

Case 2: mask state with time-reversal latent dynamics. We now consider the interpolant with a mask state:

$$p_t(x_t|x_0,x_1) = \alpha_t\delta_{x_0}(x_t) + \beta_t\delta_{x_1}(x_t) + \gamma_t\delta_M(x_t). \tag{497}$$

In our experiments and also in this discussion, we choose the schedule for simplicity. *We note that our discussion is not limited to this design choice; we choose it to avoid overloading of notations.*

$$\alpha_t = (1-a_t)^2, \qquad \beta_t = a_t^2, \qquad \gamma_t = 2a_t(1-a_t), \tag{498}$$

where $a_0 = 0$, $a_1 = 1$, and $\dot{a}_t > 0$.

The latent state space is

$$\mathcal{S} = \{0, M, 1\}, \tag{499}$$

with decoding map

$$\phi(x_0, x_1, 0) = x_0, \qquad \phi(x_0, x_1, M) = M, \qquad \phi(x_0, x_1, 1) = x_1. \tag{500}$$

The latent marginal is

$$p_t(S_t = 0) = \alpha_t, \qquad p_t(S_t = M) = \gamma_t, \qquad p_t(S_t = 1) = \beta_t. \tag{501}$$

We then design a conditional latent generator realizing this latent path. The latent generator (3×3 matrix) has 9 elements, with 6 out of them to be free (diagonal elements are not free). We then need the generator to satisfy the forward Kolmogorov Equation. This Equation has 3 states, with 2 of them to be free (as they need to sum to one). Therefore, we still have 4 degrees of freedom to choose. For simplicity, we design the following forward latent generator:

$$A_t = \frac{2\dot{a}_t}{1-a_t} \begin{bmatrix} -1 & 1-a_t & a_t \\ 0 & -\frac{a_t}{2} & \frac{a_t}{2} \\ 0 & 0 & 0 \end{bmatrix}. \tag{502}$$

Equivalently, the nonzero forward transitions are

$$0 \to M: \qquad \text{with rate } 2\dot{a}_t, \tag{503}$$

$$0 \to 1: \qquad \text{with rate } \frac{2a_t\dot{a}_t}{1-a_t}, \tag{504}$$

$$M \to 1: \qquad \text{with rate } \frac{a_t\dot{a}_t}{1-a_t}. \tag{505}$$

One can verify the forward Kolmogorov Equation:

$$\partial_t \begin{bmatrix} \alpha_t \\ \gamma_t \\ \beta_t \end{bmatrix} = A_t^\top \begin{bmatrix} \alpha_t \\ \gamma_t \\ \beta_t \end{bmatrix}. \tag{506}$$

The corresponding backward latent generator is

$$B_t = \frac{2\dot{a}_t}{a_t} \begin{bmatrix} 0 & 0 & 0 \\ \frac{1-a_t}{2} & -\frac{1-a_t}{2} & 0 \\ 1-a_t & a_t & -1 \end{bmatrix}. \tag{507}$$

Equivalently, the nonzero backward transitions are

$$1 \to M: \qquad \text{with rate } 2\dot{a}_t, \tag{508}$$

$$1 \to 0: \qquad \text{with rate } \frac{2(1-a_t)\dot{a}_t}{a_t}, \tag{509}$$

$$M \to 0: \qquad \text{with rate } \frac{(1-a_t)\dot{a}_t}{a_t}. \tag{510}$$

By direct calculation, we can verify that for $s \neq s'$, we have

$$B_t(s, s') = A_t(s', s)\frac{p_t(s')}{p_t(s)}. \tag{511}$$

i.e., the latent forward and backward generators are time reversals w.r.t the latent marginal $p_t(s)$.

Unfortunately, unlike case 1, we cannot the clean form for the forward and backward generator involving only $p_{1|t}(\cdot|x)$ and $p_{0|t}(\cdot|x)$. Therefore, we directly parameterize the marginal generators in the form of nonnegative forward and backward rates

$$R^F_{\theta,t}(x,y), \qquad R^B_{\phi,t}(x,y). \tag{512}$$

More precisely, given the current state x, the network's output is one nonnegative vector over possible next states for each token position, corresponding to the off-diagonal entries of $R^F_{\theta,t}(x,\cdot)$ and $R^B_{\phi,t}(x,\cdot)$. The diagonal entries are then reconstructed by the CTMC convention $R^F_{\theta,t}(x,x) = -\sum_{y\neq x} R^F_{\theta,t}(x,y), R^B_{\phi,t}(x,x) = -\sum_{y\neq x} R^B_{\phi,t}(x,y)$.

By Proposition E.2, we want to learn the generator as marginalizations of the latent forward and backward generators. We train this marginalization by generator matching. This yields the same loss as those we discuss at the end of case 1. For easy reference, we restate them here.

For a sampled tuple (t, x_0, x_1, s_t), define the sample-level forward and backward rate targets

$$\widehat{R}^F_t(y|x_t,x_0,x_1,s_t) := \sum_{s'\in\mathcal{S}} \mathbf{1}\{\phi(x_0,x_1,s')=y\}A_t(s_t,s'), \tag{513}$$

$$\widehat{R}^B_t(y|x_t,x_0,x_1,s_t) := \sum_{s'\in\mathcal{S}} \mathbf{1}\{\phi(x_0,x_1,s')=y\}B_t(s_t,s'). \tag{514}$$

Then

$$R^F_t(x,y) = \mathbb{E}\left[\widehat{R}^F_t(y|x_t,x_0,x_1,s_t)\,\middle|\, X_t = x\right], \tag{515}$$

and similarly for R^B_t.

The generator-matching loss following [Holderrieth et al., 2024] is then

$$\mathcal{J}_F(\theta) = \mathbb{E}_{t,x_0,x_1,s_t}\left[\sum_{y\neq x_t}\left(R^F_{\theta,t}(x_t,y) - \widehat{R}^F_t(y|x_t,x_0,x_1,s_t)\log R^F_{\theta,t}(x_t,y)\right)\right], \tag{516}$$

and

$$\mathcal{J}_B(\phi) = \mathbb{E}_{t,x_0,x_1,s_t}\left[\sum_{y\neq x_t}\left(R^B_{\phi,t}(x_t,y) - \widehat{R}^B_t(y|x_t,x_0,x_1,s_t)\log R^B_{\phi,t}(x_t,y)\right)\right]. \tag{517}$$

After learning, during simulation, we discretize the CTMC with an Euler step, and given the current state x at time t, the forward one-step transition kernel is

$$K^F_{\theta,t}(x,y) = \delta_x(y) + \Delta t\, R^F_{\theta,t}(x,y), \tag{518}$$

Similarly, the backward one-step transition kernel is

$$K^B_{\phi,t}(x,y) = \delta_x(y) + \Delta t\, R^B_{\phi,t}(x,y), \tag{519}$$

In practice, Δt is chosen small enough so that the diagonal probabilities are mostly nonnegative; otherwise, the transition probabilities are clipped and renormalized for numerical stability. The forward-backward path RND can then be calculated using these categorical distribution probabilities.

Case 3: mask state with only the path $x_0 \to M \to x_1$. In the above case, we mentioned that there are many different generator designs to realize the same path (with 4 extra DOF). In this section, we discuss the simplest case, where we only allow the path $x_0 \to M \to x_1$ for the latent generator. However, as we will see, this design is problematic, and we can only rescue it with a special design of the interpolant schedule.

We now still consider the following masked interpolant:

$$p_t(x_t|x_0,x_1) = \alpha_t\delta_{x_0}(x_t) + \beta_t\delta_{x_1}(x_t) + \gamma_t\delta_M(x_t). \tag{520}$$

We choose a latent forward generator that only enables the edges

$$0 \to M, \qquad M \to 1. \tag{521}$$

This generator matrix is

$$A_t = \begin{bmatrix} -r_{0\to M}(t) & r_{0\to M}(t) & 0 \\ 0 & -r_{M\to 1}(t) & r_{M\to 1}(t) \\ 0 & 0 & 0 \end{bmatrix}. \tag{522}$$

The forward Kolmogorov equation in latent space gives

$$\dot{\alpha}_t = -r_{0\to M}(t)\alpha_t, \tag{523}$$

$$\dot{\beta}_t = r_{M\to 1}(t)\gamma_t. \tag{524}$$

Therefore,

$$r_{0\to M}(t) = -\frac{\dot{\alpha}_t}{\alpha_t}, \qquad r_{M\to 1}(t) = \frac{\dot{\beta}_t}{\gamma_t}. \tag{525}$$

Thus,

$$A_t = \begin{bmatrix} \frac{\dot{\alpha}_t}{\alpha_t} & -\frac{\dot{\alpha}_t}{\alpha_t} & 0 \\ 0 & -\frac{\dot{\beta}_t}{\gamma_t} & \frac{\dot{\beta}_t}{\gamma_t} \\ 0 & 0 & 0 \end{bmatrix}. \tag{526}$$

So far so good. However, let's take a closer look at the highlighted number in the above matrix. The denominator $\alpha_t = 0$, when $t = 1$, ensuring that the marginal density at $t = 1$ is $\pi_1 = \pi_B$. However, the numerator is typically non-zero unless $\alpha_t = 0$ when t is larger than some threshold.

What does this mean? In the latent space, the interpretation is simple: any trajectory that remains in latent state 0 is forced to jump to the mask state as $t \to 1$. This is harmless at the latent level, because the probability of the latent state 0 also vanishes.

However, after projecting to the original data space, this behavior can become problematic. The observed state $x_t = x \neq M$ does not necessarily reveal whether the latent state is 0 or 1, especially when the same token can appear in both x_0 and x_1. Therefore, near the endpoint, the observed generator may assign a very large masking rate to any sample that *looks compatible* with having come from x_0. Therefore, this marginalized generator on the observed space assigns a large transition probability towards masks and will not be stabilized. This makes the dynamics unstable.

A simple rescue is to design a cutoff, for example:

$$\alpha_t = \max\{0, 1 - 2a_t\}, \qquad \beta_t = \max\{0, 2a_t - 1\}, \qquad \gamma_t = \min\{2a_t, 2 - 2a_t\}, \tag{527}$$

where $a_0 = 0$, $a_1 = 1$, and $\dot{a}_t > 0$. This stops the transition $0 \to M$ before its rate blows up. In fact, this schedule also implies that α_t and β_t are never nonzero at the same time. Therefore, we can view this construction as concatenating two masked diffusion models [Lou et al., 2023, Shi et al., 2024] back-to-back: first $x_0 \to M$, and then $M \to x_1$.

F Composition of Local Generators

In the previous section, we mainly discussed generators for a single modality, or even for a single coordinate in the CTMC case. In practice, however, the systems of interest may contain multiple coordinates or coordinates in different modalities. Without loss of generality, suppose the full state is

$$Z = (X, Y) \in \mathcal{Z} = X \times \mathcal{Y}. \tag{528}$$

where X and Y may represent different modalities or different parts of the system. In this section, we show that it is sufficient to define a generator for each component, conditioned on the full state, and then combine these components to obtain the full generator on $\mathcal{Z}$.

Let's define the X-component generator by freezing $Y_t = y$ and updating only X_t:

$$\mathcal{L}_{X,t} f(x,y) := \lim_{\delta \downarrow 0} \frac{\mathbb{E}[f(X_{t+\delta}, y) \,|\, X_t = x, Y_t = y] - f(x,y)}{\delta}. \tag{529}$$

Similarly, the Y-component generator is defined by freezing $X_t = x$ and updating only Y_t:

$$\mathcal{L}_{Y,t} f(x,y) := \lim_{\delta \downarrow 0} \frac{\mathbb{E}[f(x, Y_{t+\delta}) \,|\, X_t = x, Y_t = y] - f(x,y)}{\delta}. \tag{530}$$

Then, we can define the full generator as the summation of these local generators.

$$\mathcal{L}_t f(x,y) = \mathcal{L}_{X,t} f(x,y) + \mathcal{L}_{Y,t} f(x,y) \tag{531}$$

This means that we can learn a generator for each component separately, while conditioning each generator on the full joint state. After learning, we simply add these component generators to obtain the full generator and evolve the whole state simultaneously. Importantly, this does not mean the two components are independent. The generators $\mathcal{L}_{X,t}$ and $\mathcal{L}_{Y,t}$ may both depend on the full state (x, y), so the dynamics can still couple the two through their rate matrices or drifts.

Additionally, we note that this summation is not the general form. In fact, the most general form is $\mathcal{L}_t = \mathcal{L}_{X,t} + \mathcal{L}_{Y,t} + \mathcal{L}_{XY,t}$, where we need an extra interaction term $\mathcal{L}_{XY,t}$. Here, we make this modeling choice to simplify our parameterization and learning.

Why can we make this design choice? This is valid because our learned transport dynamics are determined by our designed conditional dynamics. If we design the conditional generator to be "separable" across components, then its marginalization is also "separable", conditional on the joint states, i.e.,

$$\mathcal{L}_t^{x_0,x_1} = \mathcal{L}_{X,t}^{x_0,x_1} + \mathcal{L}_{Y,t}^{x_0,x_1} \quad \Longrightarrow \quad \mathcal{L}_t = \mathcal{L}_{X,t} + \mathcal{L}_{Y,t}, \tag{532}$$

$$\text{where} \quad \mathcal{L}_{X,t} f = \mathbb{E}_{x_0,x_1|z_t}\left[\mathcal{L}_{X,t}^{x_0,x_1} f\right], \mathcal{L}_{Y,t} f = \mathbb{E}_{x_0,x_1|z_t}\left[\mathcal{L}_{Y,t}^{x_0,x_1} f\right]. \tag{533}$$

We also note that this component-wise "conditional-independent" construction has also been used in graph generation [Jo et al., 2022], and is a common choice in discrete diffusion [Lou et al., 2023, Shi et al., 2024]. In discrete diffusion, one typically learns a rate matrix (typically in the form of the concrete score, or posterior distribution) for each coordinate, which corresponds to a sparse joint rate matrix where only transitions between states that differ in one token are allowed. This construction is also widely used for designing multimodal diffusion models[Holderrieth et al., 2024, Rojas et al., 2025].

During simulation, the component updates do not need to be applied sequentially. Changes in different components can occur simultaneously, known as τ-leaping [Gillespie, 2001].

G Related Works on Neural Free Energy Estimation

Probabilistic methods in machine learning have been increasingly used to accelerate free energy estimation. In this section, we review several major directions.

Targeted FEP seeks to learn an invertible map between two states so as to reduce the variance of free energy estimates. Both discrete-time normalizing flows and continuous-time flow matching methods have been used for this purpose, relying on the change-of-variables formula to account for the work associated with the transformation [Wirnsberger et al., 2020, Zhao and Wang, 2023, Ding and Zhang, 2021].

Thermodynamic integration (TI) estimates free energy differences by integrating along an interpolation between two boundary states in the quasi-static limit. NeuralTI [Máté et al., 2024a,b] parameterizes this interpolation using diffusion models or stochastic interpolants, while simultaneously learning the energy associated with each intermediate state.

The main class of methods considered in this paper is based on finite-time averaging, or Jarzynski equalities, which use continuous-time non-equilibrium processes to bridge the two boundary states. Within this class, one line of work fixes the non-equilibrium process and seeks to reduce dissipation along the process [Doucet et al., 2022, Rosa-Raíces and Limmer, 2025]. A complementary line of

work learns an additional escort to reduce dissipation not only along the prefixed non-equilibrium process, but also across the intermediate states during the transformation. Most approaches construct this transformation along an annealing path between the energies of the two boundary states [Vargas et al., 2024, Zhong et al., 2024, Albergo and Vanden-Eijnden, 2024]. FEAT constructs the transformation through a stochastic interpolant between samples from the two boundary states [He and Du et al., 2025]. Zhang et al. [2025] combine parallel tempering on the path with escorted Jarzynski for free energy estimation.

When estimating relative free energies from equilibrium samples at both endpoints, Bennett's acceptance ratio (BAR) [Bennett, 1976] provides an asymptotically efficient estimator. This principle has also been developed for nonequilibrium path ensembles [Shirts et al., 2003, Hahn and Then, 2009, Minh and Chodera, 2009, Vaikuntanathan and Jarzynski, 2011], and underlies the bidirectional estimator used in FEAT [He and Du et al., 2025] and in Equation (30). For settings involving more than two states, the multistate Bennett acceptance ratio (MBAR) [Shirts and Chodera, 2008] further generalizes BAR by combining samples from all states into a statistically efficient estimate of their relative free energies.

Recently, Guo et al. [2026] conducted complexity analysis for the standard Jarzynski equality and anneal importance sampling, providing theoretical guidance. A future direction of our work is to conduct a similar analysis for our generalized formulation.

H Simplified Forward and Backward Ratio for Mask Diffusion

Assume we are not learning between two complicated targets. Instead, we are learning from a fully masked target to an unmasked target. We do not need to learn both the forward and backward generators. This is because the masking process has a known generator, and we only need to learn its marginal score, or directly learn the backward. This covers the case of the commonly used masked diffusion [Shi et al., 2024].

Interestingly, in the setup of mask diffusion following Shi et al. [2024], we do not need to explicitly calculate the mask and unmask probabilities in the work. Instead, we only need to calculate the unmask posterior.

More precisely, consider the unnormalized importance weight (negative work) calculated by forward and backward kernel ratios:

$$W(x_{0:N}) = \tilde{P}_0(x_0)\frac{P^F(x_1|x_0)\cdots P^F(x_N|x_{N-1})}{P^B(x_0|x_1)\cdots P^B(x_N-1|x_N)}\frac{1}{P_M(x_N)} \tag{534}$$

Following Shi et al. [2024] and defining α_n in the same way as Shi et al. [2024] (e.g., $(1-t$ or $1-\cos(\frac{\pi}{2}(1-t)))$, the transition kernel can have the following types:

1. both x_n and x_{n+1} are masks:

$$P^F(x_{n+1}|x_n) = 1 \tag{535}$$

$$P^B(x_n|x_{n+1}) = 1 - \mathbb{P}\{\text{unmask}\} = \frac{1-\alpha_n}{1-\alpha_{n+1}} \tag{536}$$

2. Neither x_n or x_{n+1} is mask:

$$P^F(x_{n+1}|x_n) = 1 - \mathbb{P}\{\text{mask}\} = \frac{\alpha_{n+1}}{\alpha_n} \tag{537}$$

$$P^B(x_n|x_{n+1}) = 1 \tag{538}$$

3. x_n is not mask, while x_{n+1} is mask:

$$P^F(x_{n+1}|x_n) = \mathbb{P}\{\text{mask}\} = \frac{\alpha_n - \alpha_{n+1}}{\alpha_n} \tag{539}$$

$$P^B(x_n|x_{n+1}) = P(x_0 = x_n|x_{n+1})\mathbb{P}\{\text{unmask}\} = P(x_0 = x_n|x_{n+1})\frac{\alpha_n - \alpha_{n+1}}{1-\alpha_{n+1}} \tag{540}$$

Assume for a trajectory $x_{0:N}$, the case 3 happen at step m to $m+1$. Then the importance weight is given by

$$W(x_{0:N}) = \tilde{P}_0(x_0)\frac{\alpha_1/\alpha_0}{1}\frac{\alpha_2/\alpha_1}{1}\cdots \\ \frac{\alpha_m/\alpha_{m-1}}{1}\frac{(\alpha_m-\alpha_{m+1})/\alpha_m}{P(x_0=x_n|x_{n+1})(\alpha_m-\alpha_{m+1})/(1-\alpha_{m+1})}\frac{1}{(1-\alpha_{m+1})/(1-\alpha_{m+2})}\cdots \\ \frac{1}{(1-\alpha_{N-2})/(1-\alpha_{N-1})}\frac{1}{(1-\alpha_{N-1})/(1-\alpha_N)}\frac{1}{P_M(x_N)} \tag{541}$$

Hence

$$W(x_{0:N}) = \tilde{P}_0(x_0)\frac{\alpha_1}{\alpha_0}\frac{\alpha_2}{\alpha_1}\cdots \\ \frac{\alpha_m}{\alpha_{m-1}}\frac{(\alpha_m-\alpha_{m+1})/\alpha_m}{P(x_0=x_n|x_{n+1})(\alpha_m-\alpha_{m+1})/(1-\alpha_{m+1})}\frac{1-\alpha_{m+2}}{1-\alpha_{m+1}}\cdots \\ \frac{1-\alpha_{N-1}}{1-\alpha_{N-2}}\frac{1-\alpha_N}{1-\alpha_{N-1}}\frac{1}{P_M(x_N)} \tag{542}$$

After cancellation, we obtain:

$$W(x_{0:N}) = \tilde{P}_0(x_0)\frac{1-\alpha_N}{\alpha_0}\frac{1}{P(x_0=x_n|x_{n+1})}\frac{1}{P_M(x_N)} = \tilde{P}_0(x_0)\frac{1}{P(x_0=x_n|x_{n+1})}\frac{1}{P_M(x_N)} \tag{543}$$

Therefore, we do not need to track the probability of masking/unmasking when estimating free energy with standard mask diffusion.

In fact, one may realize that this simplification leads to a weight similar to the AR case, but along an arbitrary reveal order. Therefore, using a masked diffusion as a transport can be interpreted as a random-order autoregressive transport, also for the free energy estimation case. Conversely, the AR free energy formulation can be viewed as a special case of this construction, where the random reveal order happens to coincide with the fixed autoregressive order.

I A Note on Heavy-tailed Distributions

By Remark 3.7, if we learn a diffusion process with finite time and a bounded, Lipschitz drift function, then we may obtain an estimator with infinite variance if the distribution on one side is heavy-tailed:

Remark I.1 (Heavy-tailed endpoints can force infinite path divergence)**.** *Consider the forward process*

$$\mathrm{d}X_t = f_t(X_t)\,\mathrm{d}t + \mathrm{d}W_t, \qquad X_0 \sim p_0, \tag{544}$$

and a backward path law $\overleftarrow{Q}$ whose terminal marginal is q_1. Assume that f_t is bounded and Lipschitz on the finite time interval $[0,1]$. If p_0 has infinite second moment while q_1 is sub-Gaussian, then

$$D_{\mathrm{KL}}(\overrightarrow{P}\,\|\,\overleftarrow{Q}) = \infty. \tag{545}$$

Proof. By the chain rule for KL divergence on path space,

$$D_{\mathrm{KL}}(\overrightarrow{P}\,\|\,\overleftarrow{Q}) = D_{\mathrm{KL}}(p_1\,\|\,q_1) + \mathbb{E}_{X_1\sim p_1}\left[D_{\mathrm{KL}}\big(\overrightarrow{P}(\,\cdot\,|X_1)\,\|\,\overleftarrow{Q}(\,\cdot\,|X_1)\big)\right]. \tag{546}$$

Therefore,

$$D_{\mathrm{KL}}(\overrightarrow{P}\,\|\,\overleftarrow{Q}) \geq D_{\mathrm{KL}}(p_1\,\|\,q_1). \tag{547}$$

It remains to show that $D_{\mathrm{KL}}(p_1\,\|\,q_1) = \infty$. To simplify the proof, we first write down the corresponding ODEs of the forward processes:

$$\mathrm{d}x_t = a_t(x_t)\mathrm{d}t, \quad x_0 \sim p_0 \tag{548}$$

where $a_t = f - \frac{1}{2}\nabla \log p_t$. As the ODEs have the same marginals as the SDEs, we will focus our discussion on the ODEs.

We now consider x_t and a fixed $y_t = \mathbb{E}_{p_t}[X_t]$ for the first ODE. We denote x_0, y_0 as the results by solving the ODE backward from t to 0.

$$x_0 - y_0 = (x_t - y_t) - \int [a_s(x_s) - a_s(y_s)]\,\mathrm{d}t \tag{549}$$

From the Lipschitz assumption, we get

$$\|x_0 - y_0\| \le \|x_t - y_t\| + \| \int [a_s(x_s) - a_s(y_s)]\,\mathrm{d}t\| \tag{550}$$

$$\le \|x_t - y_t\| + L\| \int [x_s - y_s]\,\mathrm{d}t\| \tag{551}$$

$$\le \|x_t - y_t\| + L \int \|x_s - y_s\|\mathrm{d}t \tag{552}$$

From Grönwall's inequality and take square on both sides, we get

$$\|x_0 - y_0\|^2 \le \|x_t - y_t\|^2 e^{2Lt} \tag{553}$$

We now take expectation over $x_t \sim p_t$, and then we have

$$\mathbb{E}\|x_t - y_t\|^2 \ge e^{-2Lt}\mathbb{E}_{x_0 \sim p_0}[\|x_0 - y_0\|^2] = \infty \tag{554}$$

Now use the sub-Gaussian assumption on q_1. In particular, there exists $\lambda > 0$ such that

$$\mathbb{E}_{y \sim q_1}\left[\exp\left(\lambda \|y\|^2\right)\right] < \infty. \tag{555}$$

Assume $p_1 \ll q_1$, and we write

$$r(x) = p_1(x)/q_1(x) \tag{556}$$

We use the inequality

$$uv \le u \log u - u + e^v, \qquad u \ge 0,\ v \in \mathbb{R}. \tag{557}$$

Let $u = r(x)$ and $v = \lambda\|x\|^2$, we obtain

$$r(x)\lambda\|x\|^2 \le r(x)\log r(x) - r(x) + \exp\left(\lambda\|x\|^2\right). \tag{558}$$

Taking expectation under q_1 gives

$$\lambda \mathbb{E}_{p_1}\|x\|^2 \le \mathbb{E}_{q_1}[r(X)\log r(X)] - \mathbb{E}_{q_1}[r(X)] + \mathbb{E}_{q_1}\left[\exp\left(\lambda\|x\|^2\right)\right] \tag{559}$$

$$= D_{\mathrm{KL}}(p_1\|q_1) - 1 + \mathbb{E}_{q_1}\left[\exp\left(\lambda\|x\|^2\right)\right]. \tag{560}$$

LHS is unbounded, while $-1 + \mathbb{E}_{q_1}\left[\exp\left(\lambda\|x\|^2\right)\right]$ is bounded. Therefore,

$$D_{\mathrm{KL}}(p_1\|q_1) = \infty. \tag{561}$$

□

This conclusion is a generalized version of the results discussed by Ren et al. [2025], where they show that the heavy-tailed property will be maintained with an affine drift, while we show that this is also a potential issue for a Lipschitz drift function.

J Additional Experimental Details

J.1 CTMC FEAT: Lattice Ising Models

J.1.1 System

We estimate free-energy differences between Ising models on 15×15, 25×25, 32×32, and 50×50 lattices. At each endpoint, we generate 100,000 configurations in batches of 10,000 parallel Gibbs

chains, each run for 1,000,000 cyclic single-spin updates. Of these, 98,000 are used for training and 2,000 are reserved for testing. On the 50×50 lattice (2,500 spins), we evaluate only the unmasked DFM transport.

The lattice Ising model is defined as $p(x) \propto \exp(-\beta H(x))$, where β denotes the inverse temperature of the system, and $H(x) : \{-1, 1\}^{L\times L} \to \mathbb{R}$ is the Hamiltonian given by $H(x) = -\sum_{\langle i,j\rangle} x_i x_j$. Here, $\langle i, j\rangle$ runs over all neighboring spin pairs on a 2D cyclic lattice. The task of CTMC FEAT is then to estimate the free energy difference:

$$\Delta F = F_{\beta_b} - F_{\beta_a} = -\log \frac{Z_{\beta_b}}{Z_{\beta_a}}, \qquad F_\beta = -\log Z_\beta, \qquad Z_\beta = \sum_x \exp(-\beta H(x)), \tag{562}$$

Thus, F_β denotes the dimensionless free energy (with β absorbed into the reduced energy), consistent with Equation (2) and the units of the reported tables. For the 15×15–32×32 experiments, we consider $(\beta_a, \beta_b) = (0.2, 0.4)$ and $(0.2, 0.6)$. The 50×50 experiment considers only $(0.2, 0.4)$; its analytical reference value is -144.83 in the units $\Delta F/D \times 10^3$ used in Tables 2 and 5.

J.1.2 CTMC FEAT Variants

Following Section E.2.3, we propose three CTMC FEAT variants.

Flow Matching (FM). *FM* follows Case 1 in Section E.2.3, where the discrete bridge is constructed as $x_t \sim \alpha_t \delta_{x_0}(x_t) + \beta_t \delta_{x_1}(x_t)$ with $\alpha_t = 1 - t$ and $\beta_t = t$. The method directly parameterizes the endpoint posterior distributions $p_{1|t}(\cdot|x)$ and $p_{0|t}(\cdot|x)$ using $q^F_{\theta,t}(\cdot|x)$ and $q^B_{\phi,t}(\cdot|x)$, respectively, and learns them via the cross-entropy loss in Equations (490) and (491). After training, the marginal generators are obtained according to Equations (488) and (489).

Generator Matching without Masking (GM-no mask). *GM-no mask* employs the same discrete bridge of *FM*, namely $x_t \sim \alpha_t \delta_{x_0}(x_t) + \beta_t \delta_{x_1}(x_t)$ with $\alpha_t = 1 - t$ and $\beta_t = t$, which induces the conditional generators defined in Equations (471) and (473). Unlike *FM*, which parametrizes the endpoint posterior distributions, *GM-no mask* directly parameterizes the marignal generators $R^F_t(x, y)$ and $R^B_t(x, y)$ as $R^F_{\theta,t}(x_t, y)$ and $R^B_{\phi,t}(x_t, y)$, respectively. These generators are trained using the generator-matching loss in Equations (516) and (517).

Table 4: Hyperparameters of CTMC FEAT. Parentheses indicate the 50×50 FM settings at $\beta = 0.2 \leftrightarrow 0.4$.

Hyperparameter	FM	GM-no mask	GM-w. mask
System			
Lattice sizes	15×15, 25×25, 32×32 (50×50)		
Inverse temperatures	$\beta = 0.2 \leftrightarrow 0.4$, $\beta = 0.2 \leftrightarrow 0.6$		
Scheduler			
Discrete bridges	$x_t \sim \alpha_t \delta_{x_0}(x_t) + \beta_t \delta_{x_1}(x_t)$		$x_t \sim \alpha_t \delta_{x_0}(x_t)$ $+ \beta_t \delta_{x_1}(x_t) + \gamma_t \delta_M(x_t)$
Schedulers	$\alpha_t = 1 - t, \beta_t = t$		$\alpha_t = (1 - a_t)^2, \beta_t = a_t^2$ $\gamma_t = 2a_t(1 - a_t); a_t = t$
Model			
Parameterization	$q^F_{\theta,t}(\cdot|x), q^B_{\phi,t}(\cdot|x)$	$R^F_{\theta,t}(x_t, y), R^B_{\phi,t}(x_t, y)$	
Network architecture	2D CNN		
Hidden channels	256	256	384
Depth (layers)	12 (16)	12	
Kernel size	5×5		
Batch normalization	Yes		
Residual connections	Yes		
Input encoding	scalar	scalar	one-hot
Training			
Optimizer	AdamW		
Learning rate	3×10^{-4}	3×10^{-4}	2×10^{-4}
Batch size	256		
Iterations	10,000–20,000 (20,000)	10,000–20,000	20,000
Weight decay	0	0	10^{-4}
LR schedule	No (Cosine)	No	Cosine
EMA decay	0.995		
OT pair	Yes		
Simulation and Estimation			
Step size Δt	0.05		
Evaluation sample size	2,000		
Estimator	BAR		

Generator Matching with Masking (GM-w. mask). *GM-w. mask* follows Case 2 in Section E.2.3, where the discrete bridge is constructed as $x_t \sim \alpha_t \delta_{x_0}(x_t) + \beta_t \delta_{x_1}(x_t) + \gamma_t \delta_M(x_t)$ with $\alpha_t = (1 - a_t)^2, \beta_t = a_t^2, \gamma_t = 2a_t(1 - a_t)$ and $a_t = t$, which induces the conditional generators defined in Equations (502) and (507). Similar to *GM-no mask*, *GM-w. mask* parameterizes the marignal generators directly and learns them using the generator-matching loss in Equations (516) and (517).

J.1.3 Hyperparameters and Settings for CTMC FEAT

Network. For all three CTMC FEAT variants, we use a convolutional neural network (CNN) that treats each spin configuration as a 2D spatial feature map. The input spin state $x_t \in \{0, 1\}^{L\times L}$ (or $\{0, 1, M\}^{L\times L}$ for the masking variant) is reshaped into a grid and combined with a sinusoidal

time embedding via channel concatenation before being passed through the convolutional backbone. The backbone consists of 12 convolutional layers (16 for the 50×50 FM experiment), with kernel size 5×5, circular padding, batch normalization, ReLU activations, and residual connections. The network produces two output heads, one for the forward direction and one for the backward direction, via 1×1 convolutions. For FM and GM-no mask, the input is encoded as a scalar (yielding 2 input channels), while for GM-w. mask the input is encoded as a one-hot vector over the extended vocabulary $\{0, 1, M\}$ (yielding 4 input channels).

Training details. We train all models using the AdamW optimizer with a batch size of 256. FM and GM-no mask models are trained for 10,000 iterations on the 15×15 and 25×25 lattices, and for 20,000 iterations on the 32×32 lattice. The 50×50 FM model is also trained for 20,000 iterations. For FM and GM-no mask, we use a learning rate of 3×10^{-4} with no weight decay. The learning rate is constant except for the 50×50 FM experiment, where it is cosine-annealed to zero over training. For GM-w. mask, we train for 20,000 iterations using a learning rate of 2×10^{-4}, a weight decay of 10^{-4}, and a cosine annealing learning rate schedule. We maintain an exponential moving average (EMA) of the model weights with decay rate 0.995 for all variants, and use the EMA model for evaluation. To facilitate training, we sample paired data (x_0, x_1) via an optimal transport (OT) plan computed within each mini-batch, following He et al. [2025]. Gradients are clipped to unit norm.

Estimation details. We use 2,000 held-out samples from each state for evaluation. The forward and backward trajectories are discretized with step size $\Delta t = 0.05$, corresponding to 20 discretization steps. The free energy difference is estimated using the bidirectional importance-weighted estimator and refined with the Bennett Acceptance Ratio (BAR) method as discussed in Section 5. The hyperparameters are summarized in Table 4. We report the mean and sample standard deviation over three independently trained models with different random seeds.

J.2 AR FEAT: Lattice Ising Models

J.2.1 System

We consider the same lattice Ising system as CTMC FEAT for the AR setting. Unlike CTMC FEAT, which learns a transport between two target distributions, AR FEAT learns an autoregressive proposal distribution for each target:

$$q_\theta(x) = \prod_{i=1}^{L^2} q_\theta(x_i \mid x_{<i}), \tag{563}$$

where the sites are ordered in raster scan order.

We evaluate inverse temperature $\beta \in \{0.2, 0.4, 0.6\}$, on lattices of size 15×15, 25×25, and 32×32. For each lattice size and temperature, we use the same data as the CTMC case. Each dataset contains 100,000 samples, of which 98,000 are used for training and 2,000 are held out for evaluation.

J.2.2 Hyperparameters and Settings for AR FEAT

Network. We use an autoregressive convolutional neural network to model the Ising distribution in raster order. For a configuration $x \in \{0, 1\}^{L \times L}$, the model outputs one Bernoulli logit for each lattice site, where the logit at site i parameterizes $q_\theta(x_i = 1 \mid x_{<i})$. The architecture follows a PixelCNN-style [Van den Oord et al., 2016] masked convolutional design: the first layer uses a type-A mask, so the current site and all future sites are hidden, while subsequent layers use type-B masks, which allow the current hidden representation but still exclude future sites. The backbone has 12 masked convolutional layers, hidden width 384, kernel size 5×5, and GELU activations. In addition, we include explicit causal periodic-neighbor features so that periodic boundary information that is already available in raster order can be used without leaking future spins.

Training details. We train the autoregressive model by teacher forcing, minimizing the cross-entropy, All models are trained using AdamW with learning rate 2×10^{-4}, batch size 128, and no weight decay. We train for 5 epochs and clip gradients to norm 5. For each setting, 98,000 samples are used for training, and 2,000 samples are held out for evaluation.

Estimation details. After training, we estimate the absolute free energy using the trained autoregressive model as a tractable proposal. We draw 2,000 samples from q_θ by ancestral sampling, and

also use 2,000 held-out samples from the target distribution. We combine the forward and reverse estimates using Bennett's Acceptance Ratio (BAR), initialized from the average of the two one-sided estimates. We report the estimated free energy normalized by the number of spins, F/L^2, and in the main table scale it as $F/L^2 \times 10^3$. All results are run three times, and the mean and standard deviation are calculated.

J.3 Multi FEAT: Ising Fluids

J.3.1 System

We consider a system with both continuous components (coordinates) and discrete components (spin state). We take the system from [Omelyan et al., 2004]. Precisely, denote the particle positions $\mathbf{r} = (r_1, \ldots, r_N)$, and the spin $s_i \in \{-1, +1\}$, and pairwise distances $r_{ij} = \|r_i - r_j\|$. The target energy is

$$U(\mathbf{r}, \mathbf{s}) = \sum_{1 \le i < j \le N} \left[\phi_{\mathrm{sc}}(r_{ij}) - I_R(r_{ij}) - J(r_{ij}) s_i s_j\right]. \tag{564}$$

where

$$J(r) = \frac{2(z_1\sigma)^2}{z_1\sigma + 1} \frac{\epsilon_J \sigma}{r} \exp[-z_1(r - \sigma)], \tag{565}$$

$$I(r) = \frac{2(z_2\sigma)^2}{z_2\sigma + 1} \frac{\epsilon_I \sigma}{r} \exp[-z_2(r - \sigma)]. \tag{566}$$

and

$$\phi_{\mathrm{sc}}(r) = \begin{cases} 4\epsilon_{\mathrm{LJ}} \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] + \epsilon_{\mathrm{LJ}}, & r < 2^{1/6}\sigma, \\ 0, & r \ge 2^{1/6}\sigma. \end{cases} \tag{567}$$

In our experiments, we use

$$\sigma = \epsilon_J = \epsilon_{\mathrm{LJ}} = z_1 = z_2 = \lambda = 1, \qquad H = 0, \tag{568}$$

and parameterize the nonmagnetic attraction by

$$R = \frac{\epsilon_J}{\epsilon_I}, \qquad \epsilon_I = \frac{1}{R}. \tag{569}$$

Under this parameterization,

$$J(r) = \frac{\exp[-(r-1)]}{r}, \tag{570}$$

$$I(r) = \frac{1}{R} \frac{\exp[-(r-1)]}{r}. \tag{571}$$

We then vary the value R to obtain different targets.

We also found that this target might easily diffuse when R is large. Therefore, we add a harmonic potential to avoid such diffusion:

$$U_h(\mathbf{r}) = k \frac{1}{2} \|\bar{r} - r_i\|^2 \tag{572}$$

where we set $k = 0.1$ and $\bar{r}$ represents the center of the system.

We visualize some samples in Figure 10.

We estimate for the 55-particle systems. We estimated ΔF between $R = 0.5$ and $R = 1.0$, $R = 0.5$ and $R = 1.5$ as well as $R = 0.5$ and $R = 2.0$. For the reference value, we draw samples from intermediate targets $R = 0.5, 0.6, \cdots, 1.9, 2.0$, and run MBAR [Shirts and Chodera, 2008].

J.3.2 Hyperparameters and Settings

For simplicity of learning, we did not directly bridge between the two targets. Instead, we learn transport from a Gaussian and mask distribution to the target with coordinates and ± 1 spins.

Network. We use a mixed continuous-discrete EGNN to jointly model the particle positions and spin variables. The model takes as input the continuous particle positions, the discrete spin tokens, and the time variable. The discrete state uses two output tokens corresponding to the two spin values. For the 55-particle system, token embeddings have dimension 32, and the time variable is embedded using a 32-dimensional sinusoidal time embedding. The EGNN backbone uses 8 layers with hidden dimension 96 and SiLU activations. We also include explicit spin features (± 1), consisting of the current spin value and a binary indicator for whether the token is masked. Edges are fully connected, with squared pairwise distances used as edge features. The model outputs both a continuous vector for the particle positions and discrete logits for the spin variables.

Training details. When one side is Gaussian and mask, this means that the forward process has been already fixed. We learn the backward process directly by generator matching or score matching. Here, we train the model using the sum of a continuous denoising score-matching loss and a discrete cross-entropy loss. For the continuous variables, we corrupt the data as

$$x_t = (1-t)x_0 + t\epsilon, \tag{573}$$

where ϵ is centered Gaussian noise, and train the model to predict the target vector field $\epsilon - x_0$. For the discrete variables, we use a masking corruption process corresponding to Shi et al. [2024]. At time t, each token is kept with probability

$$\alpha_t = 1 - \cos\left(\frac{\pi}{2}(1-t)\right), \tag{574}$$

and otherwise replaced by the mask token. The discrete loss is a weighted cross-entropy loss on the masked tokens, with weight

$$\pi \tan\left(\frac{\pi}{2}(1-t)\right). \tag{575}$$

All models are trained with AdamW using learning rate 1×10^{-4}, batch size 64, and no weight decay. We train for 200,000 iterations with gradient clipping at norm 1. We maintain an exponential moving average of the model parameters with decay 0.995 and use the EMA model for evaluation. We also scale the target by a factor of 2 to make learning more stable.

Estimation details. For free-energy estimation, we use the BAR estimator in Equation (30) with 2,000 held-out samples to estimate the free energy difference between the centered Gaussian particle positions and fully masked spin variables, to the target distribution. We discretize the path using 300 time steps. For the stochastic continuous updates used in work estimation, we directly obtain the score from vector field, and then set the noise scale to $\sigma_t = 0.5$. For the discrete part, following the discussion in Section H, we directly use the posterior density.

The free energy difference between two states is given by the difference between their own free energy difference against the centered Gaussian with fully masked spin.

J.4 Momentum-Augmented FEAT: Gaussian Mixture Models

J.4.1 System

We next consider a continuous-state experiment on Gaussian mixture models (GMMs). The task is to estimate the free energy difference between two 40-dimensional GMM targets. The first target contains 40 Gaussian components, while the second target contains 16 Gaussian components. Both targets use location scaling 2.0 and log-variance scaling -3.0, with different random seeds for the two mixtures.

Here, we transport the joint state

$$z_t = (x_t, y_t) \in \mathbb{R}^{40} \times \mathbb{R}^{40}, \tag{576}$$

where y_t plays the role of momentum.

J.4.2 Hyperparameters and Settings

Network. We parameterize two neural networks. The first network learns a momentum-space score $s_\theta(x_t, y_t, t)$, and the second network learns a vector field $v_\phi(x_t, y_t, t)$. Both networks are MLPs with sinusoidal time embeddings, 5 hidden layers, hidden width 400, GELU activations. The networks take the concatenated state (x_t, y_t) and time t as input and output a 40-dimensional vector.

Training details. We use Adam to optimize both networks jointly. We maintain an exponential moving average of both networks with decay rate of 0.995 and use the EMA networks for evaluation. Gradients are clipped to norm 5.

Estimation Details. For free energy estimation, we simulate forward and backward trajectories in the augmented state space using an OA integrator. We use 5,000 samples for evaluation.

For the 1D visualization on Figure 3, we use the same setting, but apply it to GMM in 1D.

K Licenses for Existing Assets

Momentum FEAT was modified from FEAT's codebase: https://github.com/jiajunhe98/FEAT, released under the MIT License.

CTMC FEAT was modified from https://github.com/J-zin/DNFS, and the computation of the reference value is adapted from https://github.com/ml-jku/DiffUCO/blob/main/IsingTheoryBaselines/IsingTheory.py. Both are released under the MIT License.

L Visualizations and Additional Results

L.1 Comparison with Classical Free-Energy Estimators

We compare CTMC–FEAT against free energy perturbation (FEP), equilibrium Bennett acceptance ratio (BAR), annealed importance sampling (AIS), and AIS–BAR in Table 5. The experiments use 2,000 evaluation samples per method, with 1,000 annealing steps for AIS and 20 transport steps for the CTMC–FEAT evaluation. The 50×50 comparison is reported for the DFM variant, the only CTMC–FEAT variant evaluated at that lattice size. The comparison concerns estimation accuracy at these evaluation budgets and does not equate end-to-end training and simulation costs.

Table 5: Comparison of free-energy estimators for lattice Ising models ($\beta = 0.2 \leftrightarrow 0.4$). Values are $\Delta F/D \times 10^3$ (mean $\pm$ reported standard deviation), where D is the number of spins. Our method uses the unmasked discrete-flow-matching (DFM) variant; dashes denote unreported experiments.

Method	25×25	50×50
Reference	-144.88	-144.83
FEP, forward	-55.92 ± 0.54	–
FEP, backward	-296.96 ± 4.61	–
Equilibrium BAR	-161.47 ± 4.84	–
AIS, forward	-129.36 ± 0.25	-128.69 ± 0.86
AIS, backward	-283.93 ± 8.11	-245.94 ± 1.40
AIS–BAR	-144.75 ± 0.31	-150.62 ± 1.17
CTMC–FEAT (DFM), forward	-138.84 ± 3.47	-143.80 ± 0.39
CTMC–FEAT (DFM), backward	-152.78 ± 6.14	-146.12 ± 0.28
CTMC–FEAT (DFM), BAR	$\mathbf{-145.30 \pm 0.28}$	$\mathbf{-144.93 \pm 0.23}$

L.2 Additional Ising Estimator Ablations

We evaluate the dependence of CTMC–FEAT estimates on the choice of free-energy estimator, the number of Monte Carlo samples, and the pairing strategy used to learn the transport. All experiments in this section use the Ising temperature pair $\beta = 0.2$ and $\beta = 0.4$, and report free-energy differences normalized as $\Delta F/D \times 10^3$.

Forward, backward and BAR estimators. Although the generalized identity allows free-energy estimates from either direction, the two-sided BAR estimator reduces the finite-sample error in these experiments. Table 6 reports the three estimators for all transport variants. The BAR results are substantially closer to the reference, especially on the larger lattices.

Table 6: Estimator ablation on Ising models. Values are $\Delta F/D \times 10^3$, mean $\pm$ reported standard deviation. DFM denotes discrete flow matching; GM denotes generator matching, both instantiated as CTMC–FEAT transports.

Transport	Estimator	15×15	25×25	32×32
Reference	–	-145.41	-144.88	-144.84
DFM, no mask	Forward	-145.18 ± 2.08	-138.84 ± 3.47	-133.94 ± 4.89
	Backward	-147.26 ± 1.71	-152.78 ± 6.14	-152.55 ± 4.59
	BAR	-145.44 ± 0.44	-145.30 ± 0.28	-143.89 ± 0.60
GM, no mask	Forward	-142.73 ± 1.70	-136.10 ± 1.71	-131.03 ± 12.66
	Backward	-148.24 ± 1.53	-151.19 ± 5.02	-159.92 ± 11.40
	BAR	-145.41 ± 0.33	-144.45 ± 0.91	-143.93 ± 0.75
GM, with mask	Forward	-141.06 ± 3.80	-126.51 ± 4.04	-123.08 ± 1.97
	Backward	-148.49 ± 3.31	-164.97 ± 1.15	-166.16 ± 1.93
	BAR	-145.14 ± 0.57	-144.74 ± 1.92	-144.62 ± 1.93

Monte Carlo sample count. We fix the DFM transport and the BAR estimator and vary the number of evaluation samples. Table 7 shows estimates becoming more stable with increasing sample count.

Table 7: Evaluation-sample ablation for DFM transport with BAR, with the same units as Table 6.

Samples	15×15	25×25	32×32
Reference	-145.41	-144.88	-144.84
10	-147.91 ± 1.98	-142.97 ± 3.63	-140.99 ± 2.60
100	-143.80 ± 1.44	-145.59 ± 0.63	-144.62 ± 2.56
500	-144.49 ± 1.15	-144.47 ± 0.82	-144.87 ± 0.59
1000	-144.62 ± 1.01	-144.69 ± 0.90	-144.93 ± 0.39

Endpoint sample pairing. For the DFM variant, both independent random pairing and minibatch optimal-transport (OT) pairing yield estimates near the reference values (Table 8). Minibatch OT is used in the primary experiments.

Table 8: Random versus minibatch-OT endpoint pairing with DFM transport and BAR. Same units as above.

Coupling	15×15	25×25	32×32
Reference	-145.41	-144.88	-144.84
Random pairing	-145.39 ± 0.30	-144.93 ± 0.96	-145.26 ± 0.70
Minibatch OT	-145.44 ± 0.44	-145.30 ± 0.28	-143.89 ± 0.60

L.3 CTMC FEAT on Ising Models

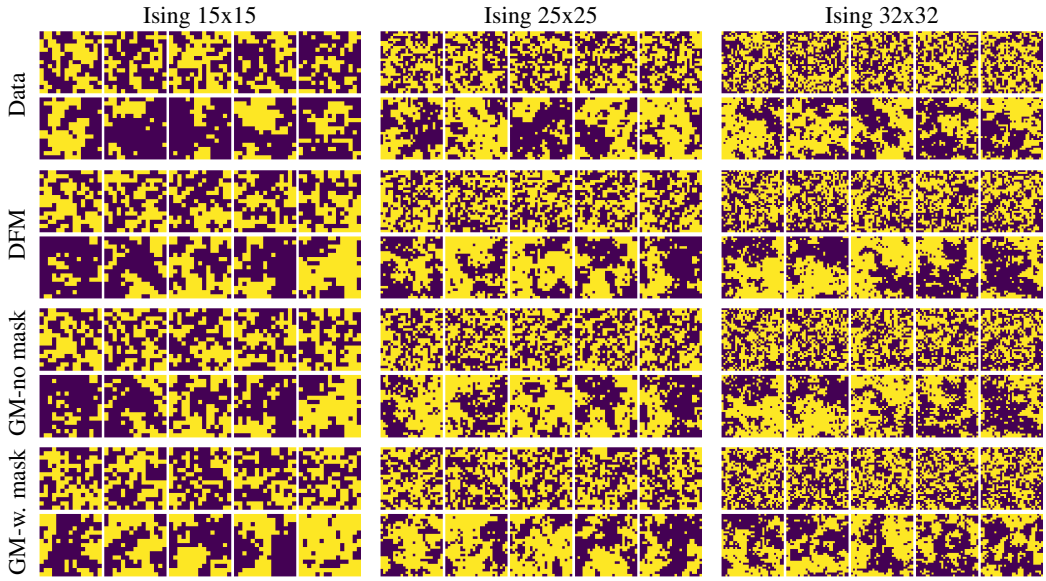


Figure 5: Generated samples from CTMC FEAT variants vs. ground truth on Ising model transport ($\beta = 0.2 \leftrightarrow 0.4$) across different lattice sizes. For each method, the two sub-rows show samples from backward and forward transport, both initialized from test data.

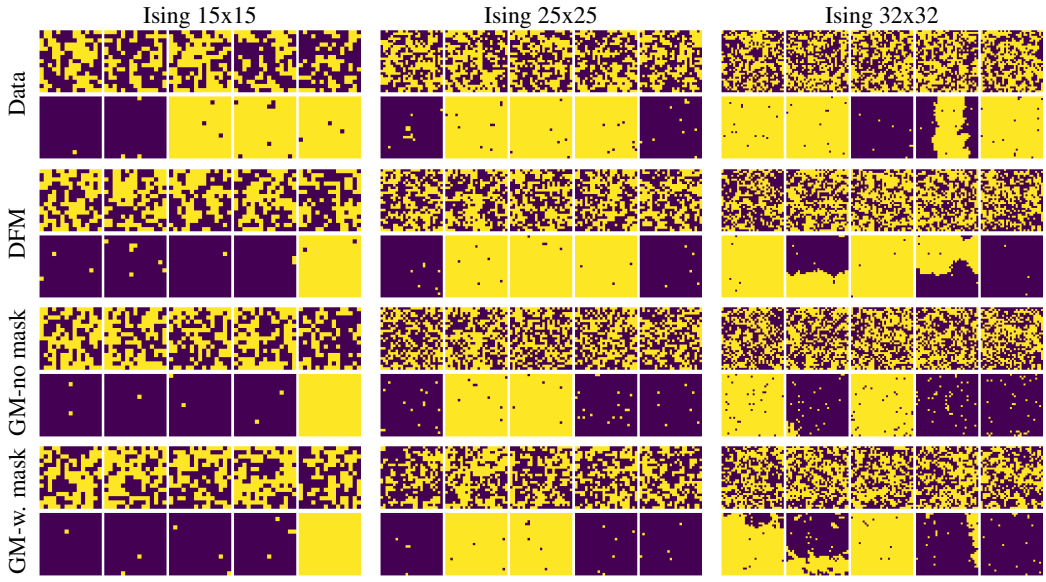


Figure 6: Generated samples from CTMC FEAT variants vs. ground truth on Ising model transport ($\beta = 0.2 \leftrightarrow 0.6$) across different lattice sizes. For each method, the two sub-rows show samples from backward and forward transport, both initialized from test data.

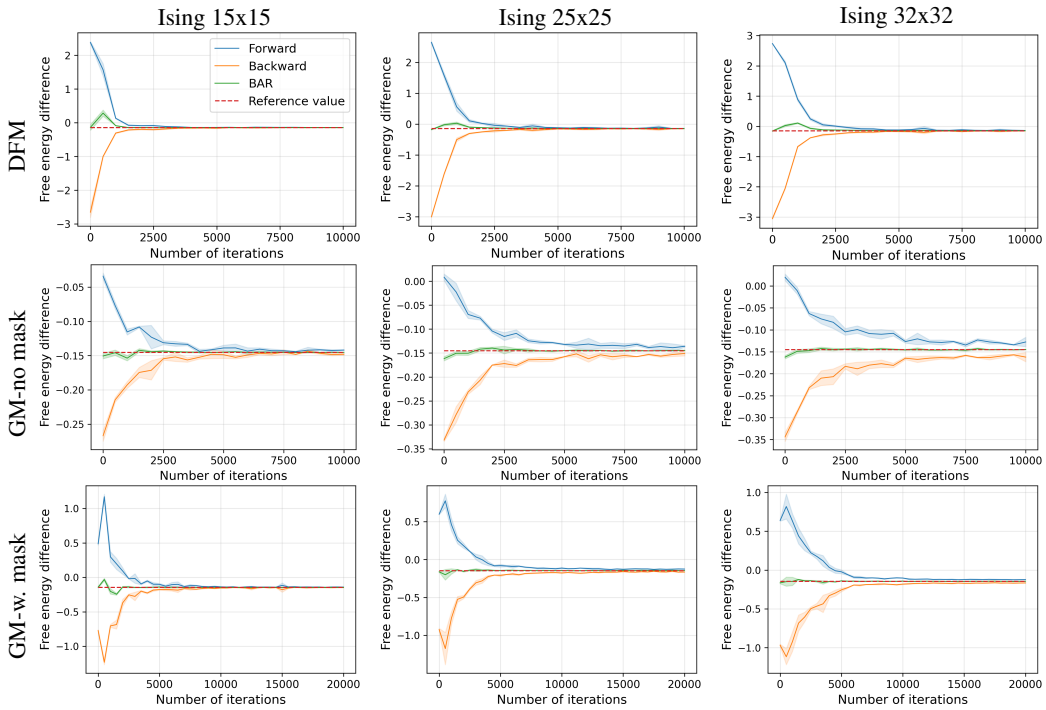


Figure 7: Free energy difference estimates ($\widehat{\Delta F}_F$, $\widehat{\Delta F}_B$, $\widehat{\Delta F}_{\mathrm{BAR}}$) over training iterations for CTMC FEAT variants on Ising model transport ($\beta = 0.2 \leftrightarrow 0.4$) across lattice sizes. The dashed red line indicates the reference value.

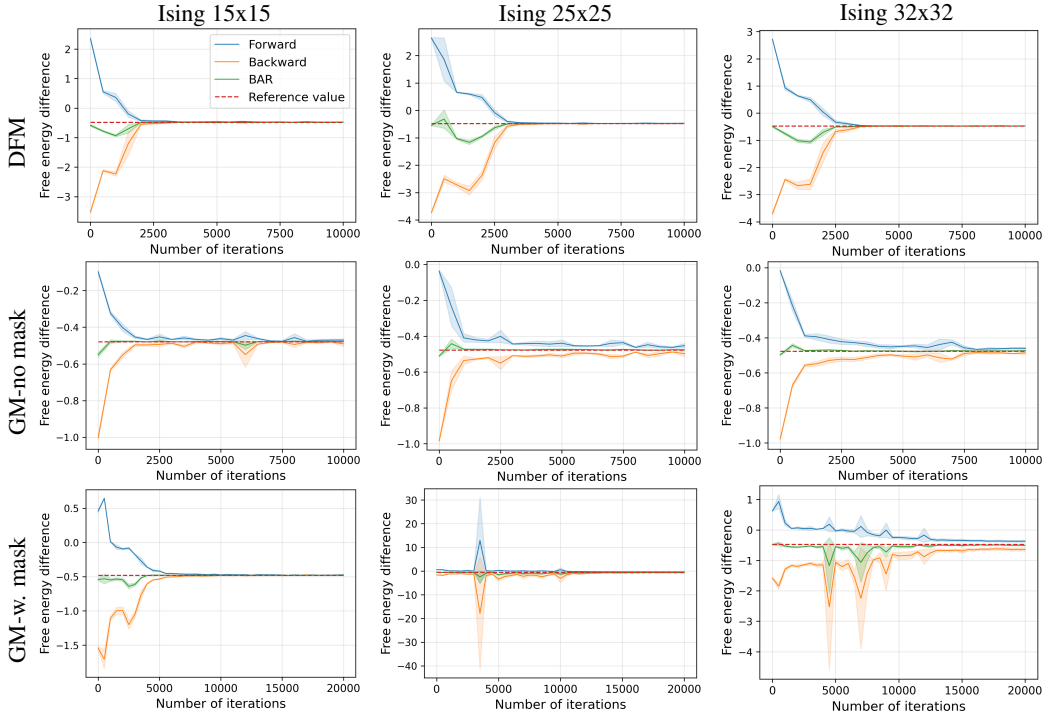


Figure 8: Free energy difference estimates ($\widehat{\Delta F}_F$, $\widehat{\Delta F}_B$, $\widehat{\Delta F}_{\mathrm{BAR}}$) over training iterations for CTMC FEAT variants on Ising model transport ($\beta = 0.2 \leftrightarrow 0.6$) across lattice sizes. The dashed red line indicates the reference value.

Table 9: Ising model using CTMC-FEAT $\beta = 0.2 \leftrightarrow 0.6$.

Dim D	**Method**	**Estimates** $(\Delta F/D \times 10^3)$	**Reference value** $(\Delta F/D \times 10^3)$
15×15 (225)	DFM	-478.68 ± 0.19	-478.68
	GM-no mask	-478.81 ± 1.03	
	GM-w. mask	-478.91 ± 0.80	
25×25 (625)	DFM	-476.21 ± 0.81	-476.71
	GM-no mask	-475.53 ± 2.69	
	GM-w. mask	-539.61 ± 35.3	
32×32 (1024)	DFM	-474.32 ± 0.60	-476.28
	GM-no mask	-473.16 ± 1.75	
	GM-w. mask	-506.83 ± 15.9	

L.4 AR FEAT on Ising Model

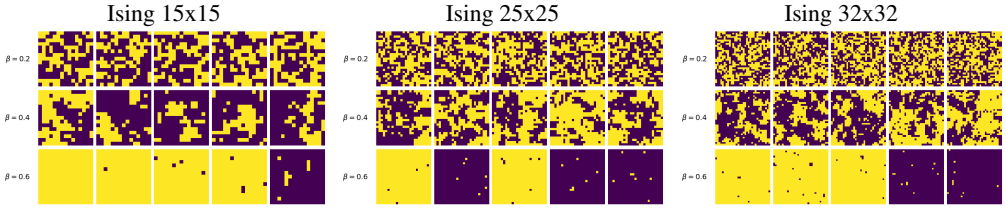


Figure 9: Samples generated by AR model at different β.

L.5 Ising Fluid

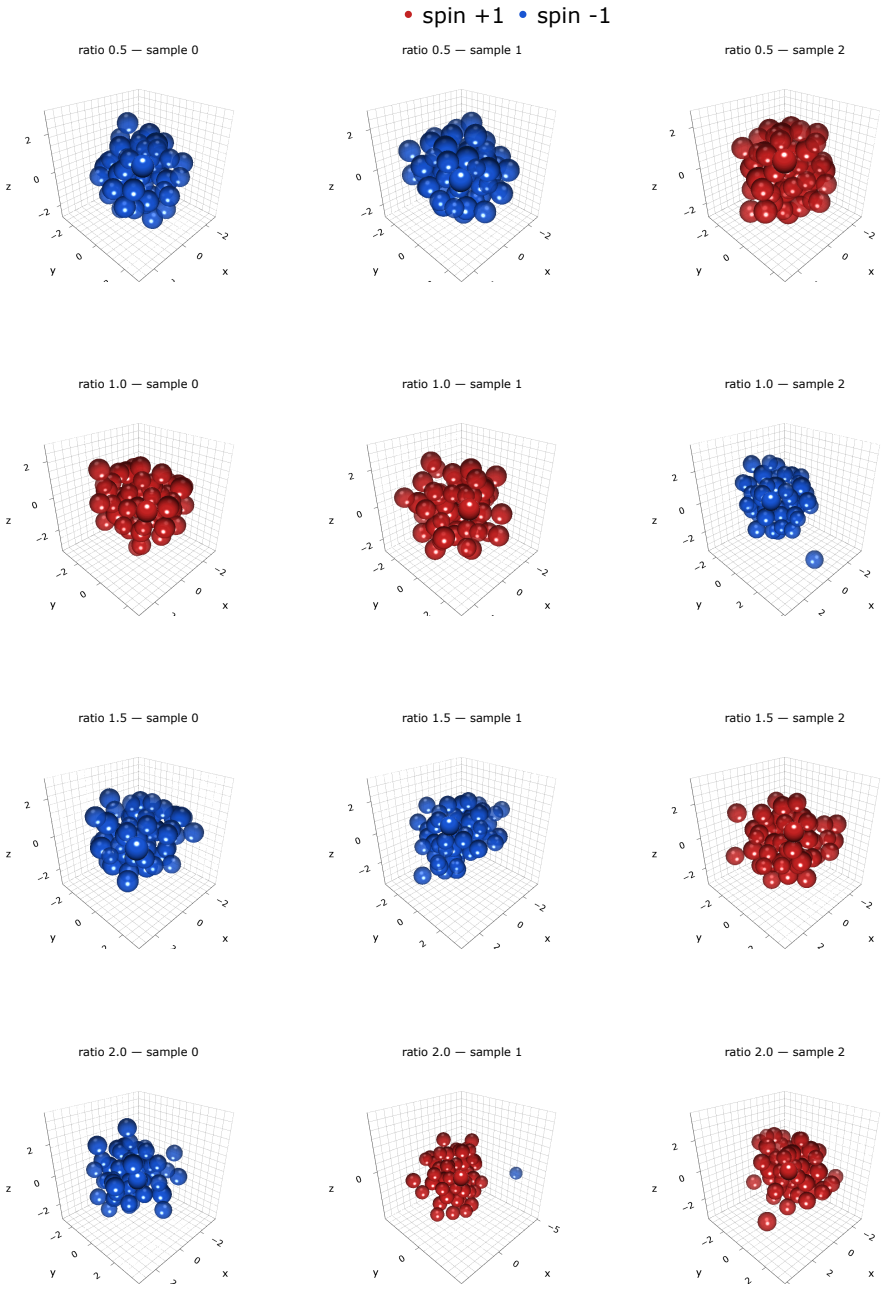
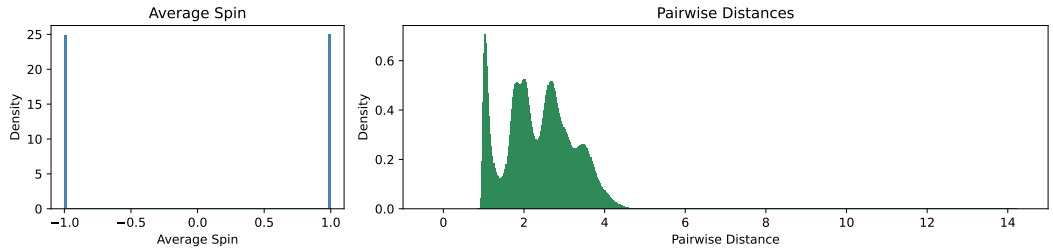
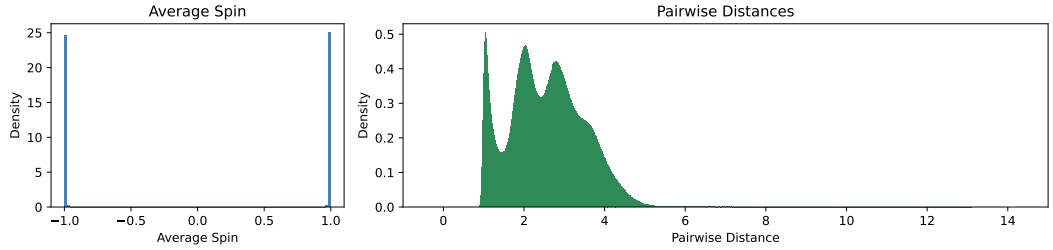


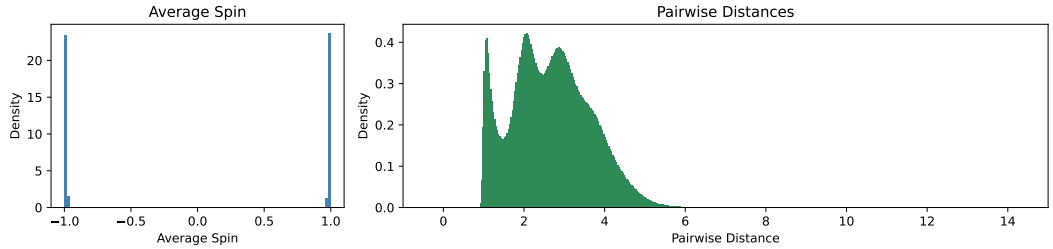
Figure 10: Visualization of 55 particles of Ising fluid (first row: $R = 0.5$, second row: $R = 1.0$, third row: $R = 1.5$, last row: $R = 2.0$). Particle size is fixed across plots.



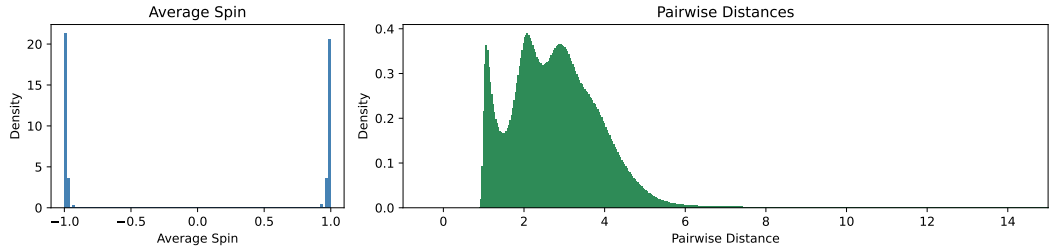
(a) $R = 0.5$



(b) $R = 1.0$



(c) $R = 1.5$



(d) $R = 2.0$

Figure 11: Average spin and pairwise distances for different values of R for 55-particle Ising fluid.

M Broader Impact

This paper presents work whose goal is to advance the field of machine learning. We expect it to have positive impacts in accelerating drug and materials discovery. However, more efficient free energy estimation also carries the risk of enabling the development of harmful chemicals or toxins. We therefore advocate for their responsible use.

N Declaration of Experiments compute resources

We run our results on a single NVIDIA A40 GPU and a GeForce RTX 4090 Graphics Card. None of our experiments require multiple GPUs, and all of the experiments can finish in several hours.

O Declaration of LLM usage

We used LLMs to assist with coding, hyperparameter search, manuscript polishing, and proofreading of the proofs.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: We proposed a generalized EJE and Crooks, and described a general framework to learn transport for estimation of the free energy.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: We discussed limitations in Conclusion section.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [Yes]

Justification: We include them in details in Section D.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: We include details of system and hyperparameter in Section J.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [Yes]

Justification: https://github.com/jiajunhe98/Free-energy-Estimation-on-Any-State-Space.

Guidelines:

- The answer [N/A] means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (https://neurips.cc/public/guides/CodeSubmissionPolicy) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (https://neurips.cc/public/guides/CodeSubmissionPolicy) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: In Section J.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: We report Error bars (std) for the estimated free energy across three runs.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)

- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: In Section N

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: The research conducted in the paper conforms, in every respect, with the NeurIPS Code of Ethics.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: In Section M.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.
    - If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.

- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: We do not have such risks.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: In Section K.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.

- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: The paper does not release new assets

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing nor research with human subjects

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

Answer: [N/A]

Justification: We used LLMs to assist with coding, hyperparameter search, manuscript polishing, and proofreading of proofs. None of them are important, original, or non-standard components.

Guidelines:

- The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.