CONTINUOUSLY TEMPERED DIFFUSION SAMPLERS

Ezra Erives, Bowen Jing, Peter Holderrieth & Tommi Jaakkola
CSAIL, Massachusetts Institute of Technology
{erives,bjing,phold}@mit.edu, tommi@csail.mit.edu

ABSTRACT

Annealing-based neural samplers seek to amortize sampling from unnormalized distributions by training neural networks to transport a family of densities interpolating from source to target. A crucial design choice in the training phase of such samplers is the *proposal distribution* by which locations are generated at which to evaluate the loss. Previous work has obtained such a proposal distribution by combining a partially learned transport with annealed Langevin dynamics. However, isolated modes and other pathological properties of the annealing path imply that such proposals achieve insufficient exploration and thereby lower performance post training. To remedy this, we propose *continuously tempered diffusion samplers*, which leverage exploration techniques developed in the context of molecular dynamics to improve proposal distributions. Specifically, a family of distributions across different temperatures is introduced to lower energy barriers at higher temperatures and drive exploration at the lower temperature of interest. We empirically validate improved sampler performance driven by extended exploration.

1 INTRODUCTION

A challenging task in Bayesian statistics and natural sciences is to sample from distributions of the form

$$\pi(x) = \frac{1}{Z}\hat{\pi}(x), \tag{1}$$

where $\hat{\pi}(x)$ denotes a known unnormalized density and $Z = \int_{\mathbb{R}^d} \hat{\pi}(x)\, \mathrm{d}x$ denotes the unknown *partition function*. Traditional MCMC-based sampling approaches, including Metropolis-Hastings and Hamiltonian Monte-Carlo (HMC), have been observed to suffer from pseudo-ergodic behavior for high-dimensional or otherwise complex choices of the density $\hat{\pi}$. Techniques such as annealed importance sampling (AIS) (Neal, 1998), and sequential Monte-Carlo-based approaches (Del Moral et al., 2006) work instead over an annealed sequence of densities $\{\hat{\pi}_t\}$ which interpolates a tractable source $\hat{\pi}_0$ and the desired target $\hat{\pi}_1 = \hat{\pi}$. Recently, *neural samplers* have been proposed which leverage a wide array of learned transport techniques, parameterized with neural networks, to optimize the annealed auxiliary distribution (Thin et al., 2021; Wu et al., 2020; Geffner & Domke, 2021; Zhang et al., 2021; Doucet et al., 2022; Geffner & Domke, 2023; Vargas & Nüsken, 2023; Richter & Berner, 2024; Albergo & Vanden-Eijnden, 2024).

In this work, we consider neural samplers which learn a *control* or *transport* to generate the evolution of a prespecified annealing path $\hat{\pi}_t$ via a so-called physics-informed neural network (PINN) loss, which penalizes the degree to which the learned vector field fails to satisfy the continuity equation (Raissi et al., 2019; Sun et al.; Máté & Fleuret, 2023; Albergo & Vanden-Eijnden, 2024). The off-policy nature of the PINN loss means that a crucial design choice in PINN-based sampler training is the *proposal distribution* $\tilde{\pi}_t$ by which samples are generated at which to evaluate the PINN loss. To this end, recent work has proposed to construct $\tilde{\pi}_t$ as the marginals of annealed Langevin dynamics steered with the partially-learned control. By augmenting the learned control in this manner, samples are driven towards high density regions of the density path, allowing for boostrapped training in which exploration gradually improves with the quality of the learned control. However, for many choices of density path, including the commonly used linear interpolation (see eq. (6)), isolated modes, high energy barriers, and other problematic features of the energy landscape obstruct this exploration process, resulting in poor sampler performance Máté & Fleuret (2023).

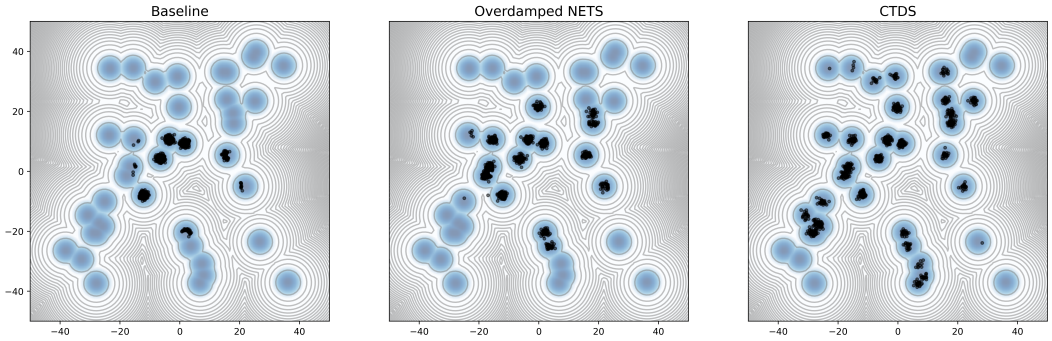

Figure 1: A comparison of samples generated by simulating the ODE given by a control $\mu_t^\theta(x_t)$ learned with various choices of proposal for the 40-mode Gaussian mixture in section 4.1. Left: Trained with a baseline proposal obtained by simulating the control with no added Langevin dynamics. Center: Trained with a Jarzynski-reweighted controlled overdamped Langevin proposal as in (Albergo & Vanden-Eijnden, 2024). Right: Trained with a Jarzynski-reweighted CTDS proposal.

To alleviate this problem, we turn toward *tempering techniques*, which are widely used in e.g., molecular dynamics to overcome high energy barriers between separate metastable states. Their success is based primarily on the insight that such barriers are lowered when the temperature is raised (Marinari & Parisi, 1992; Swendsen & Wang, 1986; Hansmann, 1997; Sugita & Okamoto, 1999; Lenner & Mathias, 2016). Specifically, we draw inspiration from recently proposed *continuous tempering* approaches, in which Langevin dynamics is run over an augmented state space given by both position and a continuous temperature variable Gobbo & Leimkuhler (2015); Graham & Storkey (2017); Luo et al. (2019). We propose to use such a continuously tempered dynamics as the basis for the proposal distribution, in doing so exploiting lower energy barriers at higher temperatures to drive exploration and thus better training. Specifically, we make the following contributions:

1. We extend the notion of a time-indexed density path to a time-and-temperature-indexed *density continuum*. We propose a multi-temperature PINN objective over such a continuum which allows for sampler training to be amortized *across temperatures*.
2. To learn dynamics across the time-temperature continuum, we propose the *continuously tempered diffusion sampler (CTDS)*, a scheme that allows for efficient exploration during sampler training by leveraging continuous tempering techniques.
3. We validate the performance of CTDS-trained samplers against those trained with existing proposals and demonstrate improved sampler performance.

2 Background

Notation We will denote normalized and unnormalized probability densities by π and $\hat{\pi}$ respectively. Further, we will call $F = -\log Z$ the *free energy* and $U = -\log \hat{\pi}$ the *energy*, so that $\pi(x) = \frac{1}{Z}\hat{\pi}(x) = e^{-U(x)+F}$. Additionally, we will denote by $\mathcal{T}$ the unit-time interval $[0, 1]$, and by $\mathcal{P}(\mathcal{X})$ the space of twice-differentiable, fully-supported, normalized densities on the space $\mathcal{X}$. When we work over phase space coordinates, we will denote the momentum as p.

2.1 Controlled Langevin Annealing

To overcome pseudo-ergodic behavior in traditional MCMC-based sampling approaches (Metropolis et al., 1953; Neal, 2012), techniques such as annealed importance sampling (AIS) Neal (1998), and later sequential Monte Carlo (SMC) based approaches Del Moral et al. (2006), work by simulating MCMC over an annealed *density path*

$$\{\hat{\pi}_t\}_{t\in\mathcal{T}}, \qquad \hat{\pi}_0 \propto \pi_{\text{source}}, \quad \text{and} \quad \hat{\pi}_1 \propto \pi_{\text{target}}, \tag{2}$$

interpolating between π_{source} and π_{target}. Recent work has since emphasized combining the sampling process with a learned transformation using e.g., normalizing flows (Wu et al., 2020; Arbel et al., 2021; Matthews et al., 2023), and, in the continuous time setting, a learned vector field (Vargas & Nüsken, 2023; Tian et al., 2024; Fan et al., 2024; Albergo & Vanden-Eijnden, 2024). Following the second approach, and given a prescribed density path as in eq. (2), we may consider a controlled (underdamped) Langevin dynamics

$$\begin{aligned} dX_t &= \mu_t^\theta(X_t) + \gamma_t M^{-1} P_t \,\mathrm{d}t \\ dP_t &= \gamma_t \left[\nabla \log \pi_t(X_t) - \varepsilon_t M^{-1} P_t\right] \mathrm{d}t + \sqrt{2\gamma_t \varepsilon_t}\,\mathrm{d}W_t \end{aligned} \tag{3}$$

as in (Zhong et al., 2024; Albergo & Vanden-Eijnden, 2024), as well as the overdamped limit given by

$$dX_t = \left[\mu_t^\theta(X_t) + \varepsilon_t \nabla \log p_t(X_t)\right] \mathrm{d}t + \sqrt{2\varepsilon_t}\,\mathrm{d}W_t, \tag{4}$$

as in Albergo & Vanden-Eijnden (2024); Vargas & Nüsken (2023), where in both cases μ_t^θ is the learned, time-dependent control.[1] Here, W_t denotes a standard Brownian motion on $\mathbb{R}^d$, ($d = d_x$ being the dimensionality of the data), and γ_t, ε_t, and M denote scaling and damping coefficients, and particle weight, respectively. The advantage of the controlled Langevin schemes in eq. (3) and eq. (4) is that the Langevin dynamics can be seen as correcting for an imperfectly learned control, and thus allowing for a bootstrapped training procedure in which the marginals of the forward process converge on the desired density path as the control improves (Albergo & Vanden-Eijnden, 2024). In particular, perfect sampling is recovered as when $\gamma_t \to \infty$ in eq. (3) and $\varepsilon_t \to \infty$ in eq. (4), or when the the control is learned perfectly so that it satisfies the continuity equation

$$\partial_t \pi_t(x_t) = -\nabla \cdot \left[\mu_t^\theta(x_t)\pi_t(x_t)\right], \qquad \forall (x_t, t) \in \mathbb{R}^d \times \mathcal{T}, \tag{5}$$

and thus facilitates an instantaneous transition along the density path for *any* choice of γ_t or ϵ_t. Several popular families of training objective have emerged to learn such a control, including minimizing a path KL divergence (Vargas & Nüsken, 2023; Zhang & Chen, 2022; Berner et al., 2022; Vargas et al., 2023), log-variance objectives (Richter et al., 2020; Nüsken & Richter, 2023; Vargas & Nüsken, 2023; Richter & Berner, 2024), and physics-informed neural network (PINN) losses (Raissi et al., 2019; Wang et al., 2023; Sun et al.). The last of these is the focus of this work and will be discussed more below.

The success of controlled annealing approaches depends heavily on the choice of density path, for which the most common choice is linear path

$$U_t^{\text{linear}}(x) \triangleq (1-t)U_0(x) + tU_1(x). \tag{6}$$

While nearly universally available, the linear path has been shown to suffer from certain regularity issues including the so-called teleportation of mass phenomenon described in (Máté & Fleuret, 2023), in certain cases causing any control satisfying eq. (5) to exhibit singularities making it hard to integrate numerically, let alone learn. One approach toward alleviating these issues is proposed in Máté & Fleuret (2023), in which an additional learned term is added to the linear path from eq. (6) to obtain

$$U_t^{\text{linear}}(x) \triangleq (1-t)U_0(x) + tU_1(x) + t(1-t)U_t^\theta(x), \tag{7}$$

which we shall refer to as the *learned path*.

2.2 PINN-Based Objectives

In the sampler setting, PINN-based objectives seek to learn the control μ_t^θ by minimizing the expected pointwise residual error of the continuity equation eq. (5). While such an objective, as originally formulated, requires access to the generally intractable free energy F_t, the free energy may be jointly learned along with the control, yielding the modified PINN objective

$$\mathcal{L}_{\text{PINN}}(\mu_t^\theta, F_t^\theta; \tilde{\pi}_t) \triangleq \int_{\mathcal{T}\times\mathbb{R}^d} |\partial_t F_t^\theta - \partial_t U_t(x) + \nabla_x \cdot \mu_t^\theta(x) - (\nabla_x U_t(x))^T \mu_t^\theta(x)|^2\, \tilde{\pi}_t(x)\,\mathrm{d}x\,\mathrm{d}t \tag{8}$$

for which the true free energy F_t can be shown to be the unique minimizer (Máté & Fleuret, 2023; Albergo & Vanden-Eijnden, 2024). More details on the PINN loss, including a derivation, can be

[1] The underdamped parameterization is slightly different than analogous formulations presented in Geffner & Domke (2023); Vargas & Nüsken (2023).

found in appendix A. Besides the choice of density path π_t, the fundamental object underpinning eq. (8) is the *proposal distribution* $\tilde{\pi}_t$, by which samples are drawn at which to evaluate the objective. A good choice of proposal distribution $\tilde{\pi}_t$ is given by the density path itself (i.e., when $\tilde{\pi}_t = \pi_t$), but such a choice is nearly always intractable. As an alternative, recent work has proposed to construct such a proposal distribution using a controlled annealed Langevin scheme (Albergo & Vanden-Eijnden, 2024).

Key Idea 2.1 (Proposal via controlled Langevin annealing)**.** Construct the PINN proposal $\tilde{\pi}$ as the marginals of a controlled Langevin annealing process, as is given in eq. (3) and eq. (4). Intuitively, as the quality of the learned control increases, so will the mode-coverage of the proposal. *Optionally*, one may additionally reweigh samples toward the ground-truth marginal π using a Jarzynski-like equality (Albergo & Vanden-Eijnden, 2024).

2.3 Tempering Techniques for Enhanced Exploration

One family of techniques to improving exploration are tempering-based methods, based on the fundamental insight that energy barriers are lowered and mixing times are improved when one *raises the temperature* of the target to obtain $\pi_t^\beta(x) \propto \pi_t(x)^\beta$, for inverse temperature $\beta \triangleq \frac{1}{T} < 1$. To exploit this insight, approaches such as parallel tempering Swendsen & Wang (1986); Hansmann (1997); Sugita & Okamoto (1999) and simulated tempering Marinari & Parisi (1992) have been proposed. These are MCMC-based sampling schemes over a temperature ladder along with mechanisms by which samples may jump between temperatures, thereby allowing faster-mixing dynamics at higher temperatures to drive convergence at a lower temperature of interest.

More recently, continuous-tempering schemes have been proposed involving a Hamiltonian dynamics over a temperature-augmented extended state thereby allowing for a smoothly-varying temperature variable Gobbo & Leimkuhler (2015); Graham & Storkey (2017); Luo et al. (2019), and demonstrating improvements over simulated and parallel tempering Graham & Storkey (2017). Such continuous tempering schemes have found utility in MD simulations Lenner & Mathias (2016), Bayesian posterior inference Luo et al. (2019), and in improving the convergence of piece-wise deterministic Markov processes Sutton et al. (2022). In this work, we extend the application of continuous tempering to the annealed neural sampler setting.

3 Continuously Tempered Diffusion Samplers

In this section, we present our main contribution: combining the controlled Langevin annealing schemes of section 2.1 with the continuous tempering approaches outlined in section 2.3, obtaining a new, tempered class of annealing schemes over a continuously-varying temperature variable, which we refer to as continuously tempered diffusion samplers (CTDS).

3.1 Tempered Density Continuums

Recall that a density path is a t-indexed annealing path $\hat{\pi}_t(x) \triangleq e^{-U_t(x)}$ which agrees with a prescribed source density $\hat{\pi}_0$ and target density $\hat{\pi}_1$. It is then natural that we extend the notion of density path to the tempered setting by affixing an additional indexing variable - the inverse-temperature variable β - to obtain the annealing continuum

$$\hat{\pi}_t^\beta(x) \triangleq e^{-U_t^\beta(x)}, \qquad U_t^\beta(x) \propto \beta U_t^1(x) \qquad \forall (t,b) \in \mathcal{T} \times \mathcal{B}, \tag{9}$$

where we have defined $\mathcal{B} \triangleq [\beta_{\text{min}}, 1]$ as the range of inverse-temperatures β of interest. We refer to the resulting continuously-indexed collection of densities $\{\hat{\pi}_t^\beta(x)\}_{t,\beta \in \mathcal{T}\times\mathcal{B}}$ as a *density continuum*. Given a target density U^{target} of interest, two such density continuums are the *linear continuum* and the *learned continuum* defined respectively for a prescribed target density $U_1^1(x) = U^{\text{target}}(x)$ by $\pi_0^\beta(x) = \mathcal{N}(0, \frac{1}{\beta} I_d)$ (thereby defining U_0^β) and

$$U_t^\beta(x) = (1-t)U_0^\beta(x) + tU_1^\beta(x) \qquad \blacktriangleright \text{ linear density continuum} \tag{10}$$

$$U_t^\beta(x) = (1-t)U_0^\beta(x) + tU_1^\beta(x) + \beta t(1-t)U_t^\theta(x) \qquad \blacktriangleright \text{ learned density continuum} \tag{11}$$

for all $(t, b) \in \mathcal{T} \times \mathcal{B}$. In the next two sections, we will construct an annealed Langevin dynamics over such a tempered density continuum. Doing so directly over $\mathcal{B}$ presents two challenges: First, it is difficult to handle the boundaries of $\mathcal{B}$ (at β_{min} and 1.0) without introducing some sort of boundary condition on the Langevin dynamics, or with the introduction of some confining potential, neither of which are particularly desirable. Second, samples from any continuous marginal density over β would almost surely not be equal to either $\beta = \beta_{\text{min}}$ (thereby discouraging mixing behavior at high temperatures) or $\beta = 1.0$ (thereby not providing training signal at the original temperature of interest). We therefore follow the lead of Gobbo & Leimkuhler (2015) and reparameterize the inverse-temperature coordinate by pulling back to the space $\Xi \triangleq \mathbb{R}$ via a carefully defined continuous mapping $\beta(\xi) : \Xi \to \mathcal{B}$ in a manner which resolves both issues; see appendix B.1 for details.

3.2 PINN Objectives Across Temperature

For fixed $\xi = \xi^\star$ (the case of $\beta(\xi^\star) = 1$ being of particular interest), we recover the standard annealed sampling problem, in which we would like to learn a control $\mu_t^\theta(x, \xi^\star)$ which satisfies the continuity equation

$$\partial_t \pi_t^{\xi^\star}(x) = \nabla_x \cdot [\mu_t^\theta(x, \xi^\star)\pi_t^{\xi^\star}(x)], \tag{12}$$

corresponding to $\{\hat{\pi}_t^{\xi^\star}(x)\}_{t\in\mathcal{T}}$ (and in turn, to the path $\{\hat{\pi}_t^{\beta(\xi^\star)}(x)\}_{t\in\mathcal{T}}$). To obtain such a control, let us consider a multi-temperature variant of the PINN loss over the space $\mathbb{R}^d \times \mathcal{T} \times \Xi$ from eq. (8), given by

$$\mathcal{L}_{\text{PINN}}^{\text{MT}}(F^\theta, \mu^\theta; \tilde{\pi}) \triangleq \int_{\mathcal{T}\times\Xi\times\mathbb{R}^d} |\partial_t F_t^\theta - \partial_t U_t^\xi + \nabla_x \cdot \mu_t^\theta - (\nabla_x U_t^\xi)^T \mu_t^\theta|^2 \, \tilde{\pi}_t^\xi(x, z) \, \mathrm{d}x \, \mathrm{d}\xi \, \mathrm{d}t, \tag{13}$$

for free energy estimate $F : \mathcal{T} \times \Xi \to \mathbb{R}$, control $\mu : \mathcal{T} \times \mathbb{R}^d \times \Xi \to \mathbb{R}^d$, and proposal distribution $\tilde{\pi}_t(x, \xi) \in \mathcal{P}(\mathbb{R}^d)$. Now, defining the free energy $F_t(\xi)$ as $F_t(\xi) \triangleq \int_{\mathbb{R}^d} e^{-U_t^\xi(x)} \, \mathrm{d}x$, we propose the following result which characterizing the minimizing F of eq. (13).

Theorem 3.1. *Let $\tilde{\pi}_t(x, \xi) \in \mathcal{P}(\mathbb{R}^d \times \Xi)$ and suppose there exists a pair $(F^\star, \mu^\star)$ which satisfies*

$$\mathcal{L}_{\text{PINN}}(F^\star, \mu^\star; \tilde{\pi}) = 0,$$

as well as the boundary condition $F_0^\star(\xi) = F_0(\xi)$ for all $\xi \in \Xi$. In this case, $F^\star$ is unique and is given by the free energy $F_t(\xi)$.

The result from theorem 3.1 extends an existing result in the single-temperature setting (Máté & Fleuret (2023), Lemma 1) to the multi-temperature setting, and a short proof can be found in appendix B.3. Furthermore, theorem 3.1 establishes that we may learn the control $\mu_t^\theta(x, \xi)$ by simultaneously learning the free energy $F_t^\theta(\xi)$. To construct such a proposal $\tilde{\pi}_t(x, \xi)$, we look to finally exploit the tempered nature of the density continuum $\{\hat{\pi}_t^\xi(x)\}$, as we discuss next.

3.3 Continuously Tempered Diffusion Samplers

In this section, we will at last formulate our main contribution. First, we will realize the density *continuum* as a density *path* by introducing a fictitious, time-dependent density over ξ. Second, we will construct the *continuously tempered diffusion sampler*, a controlled annnealed Langevin dynamics over such a density path. By introducing some time-dependent, *normalized* potential $\psi_t(\xi) : \mathcal{T} \times \Xi \to \mathbb{R}$, we may realize the continuum $\{\hat{\pi}_t^\xi(x)\}$ as a joint density

$$\hat{\pi}_t(x, \xi) \triangleq e^{-U_t^\xi(x) + \psi_t(\xi)} = e^{-\tilde{U}_t(x,\xi)}, \qquad \forall (t, \xi) \in \mathcal{T} \times \Xi \tag{14}$$

over the augmented state (x, ξ), where we have defined $\tilde{U}_t(x, \xi) \triangleq U_t(x, \xi) - \psi_t(\xi)$. Then it is easily seen that the marginal $\pi_t(\xi)$ is given by $\pi_t(\xi) \propto e^{\psi_t(\xi) - F_t(\xi)}$. As a consequence, we may rewrite $\psi_t^\theta(\xi) = \psi_t'(\xi) + F_t^\theta(\xi)$, where $F_t^\theta(\xi)$ is a learned approximation of the free energy as in eq. (8), and so that as $F_t^\theta(\xi) \to F_t(\xi)$, $\pi_t(\xi)$ becomes distributed proportionally to $e^{\psi_t'(\xi)}$. In doing so, we obtain

$$\hat{\pi}_t^\theta(x, \xi) = e^{-U_t^\xi(x) + F_t^\theta(\xi) + \psi_t'(\xi)} = e^{-\tilde{U}_t^\theta(x,\xi)}, \tag{15}$$

so as to "cancel-out" the bias of the unknown free energy $F_t(\xi)$ as the quality of our learned approximation $F_t^\theta(\xi)$ increases. In appendix B.2, we characterize more explicitly the construction of $\psi_t'(\xi)$ in practice, and in particularly the inclusion of a confining potential term. Denoting by $q_t = (x_t, \xi_t)$ and $p_t = (p_t^x, p_t^\xi)$ we may now define the *non-separable* Hamiltonian

$$\mathcal{H}_t^\theta(q_t, p_t) \triangleq \tilde{U}_t^\theta(q_t) + K(q_t, p_t), \qquad K(q_t, p_t) \triangleq \frac{\beta(\xi_t)}{2M_x}\|p_t^x\|^2 + \frac{1}{2M_\xi}\|p_t^\xi\|^2. \tag{16}$$

With the inclusion of the learned control μ_t^θ, we obtain the controlled annealed underdamped Langevin dynamics given by

$$\begin{aligned}
\mathrm{d}X_t &= \mu_t^\theta(X_t, \xi_t) + \gamma_t^x \nabla_{p^x} K(q_t, p_t)\,\mathrm{d}t \\
\mathrm{d}\xi_t &= \gamma_t^\xi \nabla_{p^\xi} K(q_t, p_t) \\
\mathrm{d}P_t^x &= \gamma_t^x \left[-\nabla_x \mathcal{H}_t^\theta(q_t, p_t) - \varepsilon_t^x \nabla_{p^x} K(q_t, p_t)\right]\mathrm{d}t + \sqrt{2\gamma_t^x \varepsilon_t^x}\,\mathrm{d}W_t^x \\
\mathrm{d}P_t^\xi &= \gamma_t^\xi \left[-\nabla_\xi \mathcal{H}_t^\theta(q_t, p_t) - \varepsilon_t^\xi \nabla_{p^\xi} K(q_t, p_t)\right]\mathrm{d}t + \sqrt{2\gamma_t^\xi \varepsilon_t^\xi}\,\mathrm{d}W_t^\xi \\
(X_0, P_0^x, \xi_0, P_0^\xi) &\sim \pi_0^\dagger(x, \xi, p^x, p^\xi).
\end{aligned} \tag{17}$$

where the extended joint density $\pi_t^\dagger$ is given by

$$\pi_t^\dagger(x_t, \xi_t, p_t^x, p_t^\xi) = \pi_t^\theta(x_t, \xi_t)\mathcal{N}(p_t^x; 0, \frac{M_x}{\beta(\xi_t)} I_d)\mathcal{N}(p_t^\xi; 0, M_\xi) \propto e^{-\mathcal{H}_t^\theta}(q_t, p_t). \tag{18}$$

In eq. (17), we have used $(\gamma_t^x, \gamma_t^\xi)$, $(\varepsilon_t^x, \varepsilon_t^\xi)$, and (M_x, M_ξ) to denote the respective scaling, damping, and weight coefficients for each of x and ξ, and have denoted by W_t^x and W_t^β standard Brownian motions on $\mathbb{R}^d$ and Ξ respectively. We refer to this scheme as a *continuously tempered diffusion sampler* (CTDS). We may then construct the proposal $\tilde{\pi}_t^\xi$ as the marginals of the forward process given by eq. (17).

3.4 A Continuously Tempered Controlled Jarzynski Result

For imperfectly learned control $\mu_t^\theta(x, \beta(\xi))$ and finite scaling coefficients γ_t^x and γ_t^ξ, the proposal distribution $\tilde{\pi}_t^\beta$ induced by the marginals of the CTDS forward process in eq. (17) will lag behind the true density path $\pi_t(x, \xi)$. To correct for this discrepancy, we derive a controlled Jarzynski equality, as in (Vargas & Nüsken, 2023; Albergo & Vanden-Eijnden, 2024), so that we may reweigh samples from eq. (17) to the correct joint density $\pi_t^\theta(x, \xi)$ described in eq. (15). Specifically, for fixed $T \in \mathcal{T}$, we may consider the work functional as in (Vaikuntanathan & Jarzynski, 2008; Zhong et al., 2024; Albergo & Vanden-Eijnden, 2024), given by

$$A_T(Q) \triangleq \int_0^T \nabla_x \cdot \mu_t^\theta(x_t, \xi_t) - \partial_t \log \tilde{U}_t^\theta(x_t, \xi_t) - \mu_t^\theta(x_t, \xi_t)^T \nabla_x \tilde{U}_t^\theta(x_t, \xi_t)\,\mathrm{d}t, \tag{19}$$

where we have denoted $q_t = (x_t, \xi_t)$ and $Q = \{q_t\}_{t\in[0,T]}$. Defining $\mathbb{P}_T$ as the path measure of the forward process eq. (17) on the interval $[0, T] \subseteq \mathcal{T}$, and letting F_t denote the free energy of the density path $\hat{\pi}_t(x, \xi)$, we have the following controlled Jarzynski identity.

Theorem 3.2 (Continuously Tempered Controlled Jarzynski Equality)**.** *For $T \in (0, 1]$, let $h : \mathbb{R}^d \times \Xi \to \mathbb{R}$ denote some observable. Then*

$$\mathbb{E}_{(x,\xi)\sim\pi_t(x,\xi))}\left[h(x, \xi)\right] = \frac{\mathbb{E}_{Q\sim\mathbb{P}_T}\left[h(Q)\exp\left(A_T(Q)\right)\right]}{\mathbb{E}_{Q\sim\mathbb{P}_T}\left[\exp\left(A_T(Q)\right)\right]}. \tag{20}$$

In particular,

$$\mathbb{E}_{(x,\xi)\sim\pi_t(x,\xi))}\left[\exp(A_T(Q))\right] = \exp(F_0 - F_T), \tag{21}$$

where we have defined $F_t \triangleq \int_{\Xi\times\mathbb{R}^d} e^{-\tilde{U}_t^\theta(x,z)}\,\mathrm{d}x\,\mathrm{d}z$.

We note that theorem 3.2 closely resembles recently established controlled Jarzynski results from (Vargas & Nüsken, 2023; Albergo & Vanden-Eijnden, 2024), that it is in fact a corollary of a more

general controlled variant of the Crooks fluctuation theorem as established in (Vaikuntanathan & Jarzynski, 2008; Vargas & Nüsken, 2023; Zhong et al., 2024), and that this resemblance is due entirely to the CTDS construction as an annealed Langevin dynamics. We present a self-contained proof of theorem 3.2 by way of an analogous Crooks identity in appendix C. We conclude by demonstrating how theorem 3.2 can be used to reweigh samples from the forward process eq. (17) to match the joint density $\pi_t^\theta(x, z)$ defined in eq. (15). This also allows us to reweight the loss and up-weigh "important" samples during training. Observe that by theorem 3.2, it follows that

$$\mathcal{L}_{\text{PINN}}^{\text{MT}}(F, \mu; \pi^\theta) = \int_{\mathcal{T}} \frac{\mathbb{E}_{Q\sim\mathbb{P}_t}\left[|\partial_t F_t - \partial_t U_t + \nabla_x \cdot \mu_t - (\nabla_x U_t)^T \mu_t|^2 \exp(A_t(Q))\right]}{\mathbb{E}_{Q\sim\mathbb{P}_t}\left[\exp(A_t(Q))\right]} \, \mathrm{d}t. \tag{22}$$

Details can be found in appendix C.4. We conclude from eq. (22) that we may reweight samples from the forward process eq. (17) so as obtain the proposal $\pi_t^\theta(x, \xi)$, which as we recall from eq. (15), converges to $\pi_t^\xi(x)\pi_t(\xi) = \pi_t^\xi(x)e^{-\psi'(\xi)}$ as $F_t^\theta(\xi) \to F_t(\xi)$. At the start of training, or when $F_t^\theta(\xi) \approx F_t(\xi)$ otherwise fails to hold, it is observed that samples tend to "pool" at either higher or lower temperatures (depending on the target). This effect diminishes as the quality of the free energy approximation increases; see fig. 2. It should also be noted that since theorem 3.2 reweights toward the joint distribution $\pi_t^\theta(x, \xi)$, a poor estimate of the free energy will bias the reweighting toward particular values of ξ.

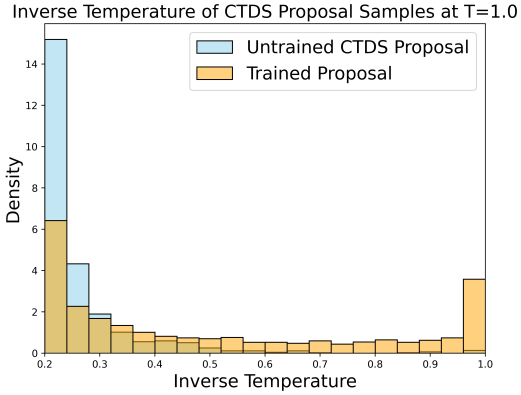

Figure 2: Distribution of sample inverse temperatures at $t = 1.0$ obtained by simulating eq. (17) before and after training for the non-Jarzynski-reweighted CTDS proposal as in section 4.1.

4 Experiments

4.1 40-Mode Gaussian Mixture

To investigate the benefit of the CTDS proposal, we compare the performance of samplers trained with various choices of proposal on a modified version challenging 40-mode Gaussian mixture from (Midgley et al., 2023).[2] We emphasize upfront, and to the strongest degree possible, that great care must be taken in comparing the results we report to existing work. Different works consider different density paths and various degrees of complexity for this problem. For example, learning a sampler for such a target *can be made artificially easier by e.g., increasing the scale of the source distribution* π_0. In this experiment, we choose to take our source distribution $\pi_0 = \mathcal{N}(0, 5.0I_2)$, thereby ensuring that modes discovered during training are done so on account of the proposal, rather than an advantageously chosen initialization. As our density path, we take the learned path given in eq. (7) (and for CTDS, the multi-temperature analog given by eq. (11)).

Training. As a baseline, we consider the proposal obtained by simulating the control by itself with no Langevin dynamics, as in (Máté & Fleuret, 2023), and which we label as the reference proposal. We additionally consider proposals obtained from the controlled overdamped (eq. (4), with $\varepsilon_t = 50.0$) and controlled underdamped/inertial (eq. (3), with $(\gamma_t, \varepsilon_t = 50.0, 2.0)$) Langevin annealing dynamics as in (Albergo & Vanden-Eijnden, 2024), which we train with and without use of the Jarzynski equality. Finally, we consider our CTDS proposal (eq. (17), with $(\gamma_t^x, \gamma_t^\xi, \varepsilon_t^x, \varepsilon_t^\xi) = (50.0, 2.0, 5.0, 2.0)$) both with and without reweighting with the Jarzynski equality (see section 3.4). We train using a replay buffer, re-sampling once per epoch for 1250 epochs for a total of 125000 training iterations at which the PINN loss is evaluated at 6250 randomly sampled elements of the buffer. The control, learnable density path, and free energy are all parameterized as

[2]We take the standard deviation of the isotropic mixture components to be $\sigma = 0.25$.

40-mode Gaussian Mixture			
Proposal	$\mathcal{W}_2 \downarrow$	**ELBO** $\uparrow$	**EUBO** $\downarrow$
Baseline	23.90 ± 0.32	-2.11 ± 0.00	23.78 ± 0.34
NETS (OD) w/o Jar.	22.61 ± 0.16	-1.65 ± 0.00	22.27 ± 0.23
NETS (OD)	20.11 ± 0.23	-1.15 ± 0.08	19.77 ± 0.22
NETS (UD) w/o Jar.	23.47 ± 0.31	-1.98 ± 0.00	22.00 ± 0.32
NETS (UD)	21.30 ± 0.23	-1.35 ± 0.02	20.95 ± 0.26
CTDS w/o Jar. (ours)	$\mathbf{12.87 \pm 0.20}$	$\mathbf{-0.24 \pm 0.01}$	19.91 ± 0.45
CTDS (ours)	14.70 ± 0.28	-0.29 ± 0.01	$\mathbf{15.35 \pm 0.34}$

Figure 3: Results for samplers trained on the 40-mode Gaussian mixture target as in section 4.1.

simple MLPs. Integration is performed with an Euler-Maruyama solver with step size $\Delta t = 0.002$. We additionally utilized Gaussian Fourier features as in Tancik et al. (2020) to encode spatial, temporal, and temperature variables, and which we found to improve training stability and sampler performance. More details can be found in appendix D.1

Evaluation. Post-training, we sample by simulating the dynamics

$$\mathrm{d}X_t = \mu_t^\theta(X_t)\,\mathrm{d}t, \qquad X_0 \sim \pi_0 \triangleq \mathcal{N}(0, 5.0 I_2). \tag{23}$$

In the CTDS case, we take $\mu_t^\theta(X_t) = \mu_t^\theta(X_t, \beta(\xi) = 1.0)$. We report mean and std. values for both 2-Wasserstein ($\mathcal{W}_2$), evidence lower bound (ELBO), and evidence upper bound (EUBO), for each using $N = 2500$ samples for each of 10 trials, as can be found in fig. 3. We observe that the CTDS outperform their baseline, overdamped and underdamped counterparts, achieving both lower $\mathcal{W}_2$ values and ELBOs, and that the Jarzynski-reweighted CTDS achieves the lowest EUBO. Curiously, while reweighted NETS based proposals seem to uniformly outperform non-reweighted counterparts, the same trend does not hold for CTDS, and in particular the non-reweighted CTDS proposal achieves both lower $\mathcal{W}_2$ and higher ELBO. We hypothesize that this is due in part to the effects discussed in section 3.4 (and visualized in fig. 2), by which an initially poor free energy approximation biases both samples and importance weights are initially biased towards high temperatures. A visualization comparing the baseline, reweighted overdamped NETS, and reweighted CTDS is provided in fig. 1.

5 Conclusion

In this work, we have proposed *continuously tempered diffusion samplers* (CTDS), a novel family of controlled Langevin annealing processes based on continuous tempering techniques from e.g., molecular dynamics. When combined with PINN-based sampler training, CTDS-based proposals outperform existing proposals based on overdamped and underdamped Langevin annealing on the challenging 40-mode Gaussian mixture from Midgley et al. (2023). We leave it to future work to extend CTDS to higher dimensional or otherwise complex targets and other scientific applications.

6 Acknowledgements

We thank Jiajun He, Yuanqi Du, and Michael Albergo for helpful discussions regarding PINN-based sampler training over the course of this project. We additionally acknowledge support from the Machine Learning for Pharmaceutical Discovery and Synthesis (MLPDS) consortium, the DTRA Discovery of Medical Countermeasures Against New and Emerging (DOMANE) threats program, and the NSF Expeditions grant (award 1918839) Understanding the World Through Code

References

Michael S. Albergo and Eric Vanden-Eijnden. Nets: A non-equilibrium transport sampler, 2024. URL `https://arxiv.org/abs/2410.02711`.

Michael Arbel, Alexander G. D. G. Matthews, and Arnaud Doucet. Annealed flow transport monte carlo, 2021. URL https://arxiv.org/abs/2102.07501.

Julius Berner, Lorenz Richter, and Karen Ullrich. An optimal control perspective on diffusion-based generative modeling. *arXiv preprint arXiv:2211.01364*, 2022.

Denis Blessing, Xiaogang Jia, Johannes Esslinger, Francisco Vargas, and Gerhard Neumann. Beyond elbos: A large-scale evaluation of variational methods for sampling, 2024. URL https://arxiv.org/abs/2406.07423.

Gavin E. Crooks. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Physical Review E*, 60(3):2721–2726, September 1999. ISSN 1095-3787. doi: 10.1103/physreve.60.2721. URL http://dx.doi.org/10.1103/PhysRevE.60.2721.

Pierre Del Moral, Arnaud Doucet, and Ajay Jasra. Sequential monte carlo samplers. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 68(3):411–436, 2006.

Arnaud Doucet, Will Grathwohl, Alexander G. D. G. Matthews, and Heiko Strathmann. Score-based diffusion meets annealed importance sampling, 2022. URL https://arxiv.org/abs/2208.07698.

Stefan Elfwing, Eiji Uchibe, and Kenji Doya. Sigmoid-weighted linear units for neural network function approximation in reinforcement learning, 2017. URL https://arxiv.org/abs/1702.03118.

Mingzhou Fan, Ruida Zhou, Chao Tian, and Xiaoning Qian. Path-guided particle-based sampling. In *Forty-first International Conference on Machine Learning*, 2024. URL https://openreview.net/forum?id=Kt4fwiuKqf.

Tomas Geffner and Justin Domke. Mcmc variational inference via uncorrected hamiltonian annealing, 2021. URL https://arxiv.org/abs/2107.04150.

Tomas Geffner and Justin Domke. Langevin diffusion variational inference, 2023. URL https://arxiv.org/abs/2208.07743.

Gianpaolo Gobbo and Benedict J Leimkuhler. Extended hamiltonian approach to continuous tempering. *Physical Review E*, 91(6):061301, 2015.

Matthew M. Graham and Amos J. Storkey. Continuously tempered hamiltonian monte carlo, 2017. URL https://arxiv.org/abs/1704.03338.

Ulrich HE Hansmann. Parallel tempering algorithm for conformational studies of biological molecules. *Chemical Physics Letters*, 281(1-3):140–150, 1997.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2017. URL https://arxiv.org/abs/1412.6980.

Hiroshi Kunita. *Stochastic flows and jump-diffusions*. Springer, 2019.

Nicolas Lenner and Gerald Mathias. Continuous tempering molecular dynamics: a deterministic approach to simulated tempering. *Journal of chemical theory and computation*, 12(2):486–498, 2016.

Robert S Liptser and Albert N Shiryaev. *Statistics of random processes: I. General theory*, volume 5. Springer Science & Business Media, 2013.

Rui Luo, Jianhong Wang, Yaodong Yang, Zhanxing Zhu, and Jun Wang. Thermostat-assisted continuously-tempered hamiltonian monte carlo for bayesian learning, 2019. URL https://arxiv.org/abs/1711.11511.

E Marinari and G Parisi. Simulated tempering: A new monte carlo scheme. *Europhysics Letters (EPL)*, 19(6):451–458, July 1992. ISSN 1286-4854. doi: 10.1209/0295-5075/19/6/002. URL http://dx.doi.org/10.1209/0295-5075/19/6/002.

Alexander G. D. G. Matthews, Michael Arbel, Danilo J. Rezende, and Arnaud Doucet. Continual repeated annealed flow transport monte carlo, 2023. URL https://arxiv.org/abs/2201.13117.

Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953.

Laurence Illing Midgley, Vincent Stimper, Gregor N. C. Simm, Bernhard Schölkopf, and José Miguel Hernández-Lobato. Flow annealed importance sampling bootstrap, 2023.

Bálint Máté and François Fleuret. Learning interpolations between boltzmann densities, 2023. URL https://arxiv.org/abs/2301.07388.

Radford M. Neal. Annealed importance sampling, 1998. URL https://arxiv.org/abs/physics/9803008.

Radford M Neal. Mcmc using hamiltonian dynamics. *arXiv preprint arXiv:1206.1901*, 2012.

Nikolas Nüsken and Lorenz Richter. Solving high-dimensional hamilton-jacobi-bellman pdes using neural networks: perspectives from the theory of controlled diffusions and measures on path space, 2023. URL https://arxiv.org/abs/2005.05409.

Maziar Raissi, Paris Perdikaris, and George E Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378:686–707, 2019.

Lorenz Richter and Julius Berner. Improved sampling via learned diffusions, 2024. URL https://arxiv.org/abs/2307.01198.

Lorenz Richter, Ayman Boustati, Nikolas Nüsken, Francisco J. R. Ruiz, and Ömer Deniz Akyildiz. Vargrad: A low-variance gradient estimator for variational inference, 2020. URL https://arxiv.org/abs/2010.10436.

Yuji Sugita and Yuko Okamoto. Replica-exchange molecular dynamics method for protein folding. *Chemical physics letters*, 314(1-2):141–151, 1999.

Jingtong Sun, Julius Berner, Kamyar Azizzadenesheli, and Anima Anandkumar. Physics-informed neural networks for sampling. In *ICLR 2024 Workshop on AI4DifferentialEquations In Science*.

Matthew Sutton, Robert Salomone, Augustin Chevallier, and Paul Fearnhead. Continuously-tempered pdmp samplers, 2022. URL https://arxiv.org/abs/2205.09559.

Robert H Swendsen and Jian-Sheng Wang. Replica monte carlo simulation of spin-glasses. *Physical review letters*, 57(21):2607, 1986.

Matthew Tancik, Pratul P. Srinivasan, Ben Mildenhall, Sara Fridovich-Keil, Nithin Raghavan, Utkarsh Singhal, Ravi Ramamoorthi, Jonathan T. Barron, and Ren Ng. Fourier features let networks learn high frequency functions in low dimensional domains, 2020. URL https://arxiv.org/abs/2006.10739.

Achille Thin, Nikita Kotelevskii, Arnaud Doucet, Alain Durmus, Eric Moulines, and Maxim Panov. Monte carlo variational auto-encoders, 2021. URL https://arxiv.org/abs/2106.15921.

Yifeng Tian, Nishant Panda, and Yen Ting Lin. Liouville flow importance sampler, 2024. URL https://arxiv.org/abs/2405.06672.

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

Francisco Vargas and Nikolas Nüsken. Transport, variational inference and diffusions: with applications to annealed flows and schr\" odinger bridges. *arXiv preprint arXiv:2307.01050*, 2023.

Francisco Vargas, Will Grathwohl, and Arnaud Doucet. Denoising diffusion samplers, 2023. URL https://arxiv.org/abs/2302.13834.

Sifan Wang, Shyam Sankaran, Hanwen Wang, and Paris Perdikaris. An expert's guide to training physics-informed neural networks, 2023. URL https://arxiv.org/abs/2308.08468.

Hao Wu, Jonas Köhler, and Frank Noé. Stochastic normalizing flows, 2020. URL https://arxiv.org/abs/2002.06707.

Guodong Zhang, Kyle Hsu, Jianing Li, Chelsea Finn, and Roger B Grosse. Differentiable annealed importance sampling and the perils of gradient noise. *Advances in Neural Information Processing Systems*, 34:19398–19410, 2021.

Qinsheng Zhang and Yongxin Chen. Path integral sampler: a stochastic control approach for sampling, 2022. URL https://arxiv.org/abs/2111.15141.

Adrianne Zhong, Ben Kuznets-Speck, and Michael R. DeWeese. Time-asymmetric fluctuation theorem and efficient free-energy estimation. *Physical Review E*, 110(3), September 2024. ISSN 2470-0053. doi: 10.1103/physreve.110.034121. URL http://dx.doi.org/10.1103/PhysRevE.110.034121.

A A Brief Overview of PINN Objectives

We briefly derive the PINN objective from eq. (8). Recall the the continuity equation

$$\partial_t \pi_t(x_t) = -\nabla \cdot \left[\mu_t^\theta(x_t)\pi_t(x_t)\right], \qquad \forall (x_t, t) \in \mathbb{R}^d \times \mathcal{T}, \tag{24}$$

from eq. (5). Dividing both sides by $\pi_t(x_t)$ yields

$$\partial_t \log \pi_t(x_t) = -\nabla \mu_t^\theta(x_t) - \mu_t^\theta(x_t)^T \nabla \log \pi_t(x_t). \tag{25}$$

Finally, plugging. in $\log \pi_t(x_t) = -U_t(x) + F_t$ and rearranging yields

$$\partial_t F_t - \partial_t U_t(x) + \nabla \mu_t^\theta(x_t) - (\nabla_x U_t(x))^T \mu_t^\theta(x) = 0 \tag{26}$$

from which we obtain the PINN objective

$$\mathcal{L}_{\text{PINN}}(\mu_t^\theta; \tilde{\pi}_t) \triangleq \int_{\mathcal{T}\times\mathbb{R}^d} |\partial_t F_t - \partial_t U_t(x) + \nabla_x \cdot \mu_t^\theta(x) - (\nabla_x U_t(x))^T \mu_t^\theta(x)|^2 \, \tilde{\pi}_t(x) \, \mathrm{d}x \, \mathrm{d}t \tag{27}$$

for fully-supported *proposal distribution* $\tilde{\pi}_t$ (Raissi et al., 2019; Sun et al.; Wang et al., 2023). The free energy F_t however is generally unknown, and thus must be estimated (Tian et al., 2024; Fan et al., 2024), or jointly-learned with the control Máté & Fleuret (2023); Albergo & Vanden-Eijnden (2024). In line with this second approach, subsequent work has established a modified PINN objective by which the free energy may be simultaneously learned along with the control, viz.,

$$\mathcal{L}_{\text{PINN}}(\mu_t^\theta, F_t^\theta; \tilde{\pi}_t) \triangleq \int_{\mathcal{T}\times\mathbb{R}^d} |\partial_t F_t^\theta - \partial_t U_t(x) + \nabla_x \cdot \mu_t^\theta(x) - (\nabla_x U_t(x))^T \mu_t^\theta(x)|^2 \, \tilde{\pi}_t(x) \, \mathrm{d}x \, \mathrm{d}t \tag{28}$$

for which the true free energy F_t can be shown to be the unique minimizer Máté & Fleuret (2023); Albergo & Vanden-Eijnden (2024). To better facilitate the learning of the free energy, past work has proposed a curriculum training procedure in which the domain of integration along the time dimension is slowly annealing from zero to one, so as to initially learn the free energy on some smaller interval before all of $\mathcal{T}$.

B Extended Details on Continuously Tempered Diffusion Samplers

In this section, we elaborate an additional design choices for continuously tempered diffusion samplers, including the choice of reparameterization, the use of confining and biasing potentials for temperature. Additionally, we provide proofs of theorem 3.1 and **??**.

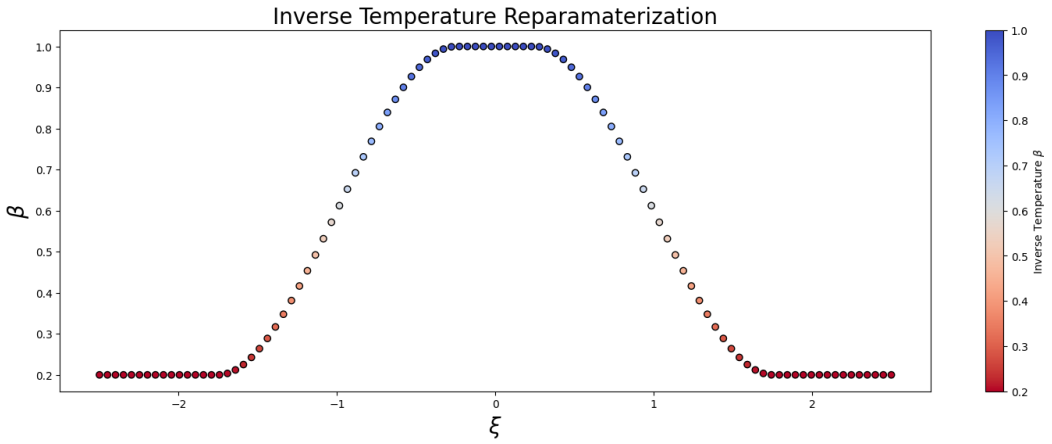

Figure 4: An illustration of the inverse temperature reparameterization given in eq. (30).

B.1 Temperature Reparameterization Details

Let us recall the definition of a density continuum over $(t, \beta) \in \mathcal{T} \times \mathcal{B}$ from eq. (29), given by

$$\hat{\pi}_t^{\beta}(x) \triangleq e^{-U_t^{\beta}(x)}, \qquad U_t^{\beta}(x) \propto \beta U_t^1(x) \qquad \forall (t, b) \in \mathcal{T} \times \mathcal{B}, \tag{29}$$

where we have defined $\mathcal{B} \triangleq [\beta_{\text{min}}, 1]$ as the range of inverse-temperatures β of interest. Our goal of constructing an annealed Langevin dynamics over such a continuum necessitates a joint density over both x and β, and in turn, an associated marginal distribution over β. Working directly over β is thus undesirable for two reasons: First, assuming that such a marginal over β is continuous, samples from this marginal would almost surely not be equal to either β_{min} or 1.0, the two temperatures of particular relevance. This is problematic because it would mean that any proposal constructed with e.g., Langevin dynamics over such a joint density would provide little training signal at the desired temperature of 1.0, nor would it spend much exploring at $t = \beta_{\text{min}}$. Secondly, we would like to constrain the temperature range to only the interval $[\beta_{\text{min}}, 1]$, and enforcing these boundary conditions directly presents complications. We therefore follow the lead of Gobbo & Leimkuhler (2015) by reparameterizing via

$$\beta(\xi) \triangleq \begin{cases} 1 & \text{if } |\xi| < \Delta \\ 1 - (1 - \beta_{\text{min}}) \left[3 \left(\frac{|\xi| - \Delta}{\Delta' - \Delta} \right)^2 - 2 \left(\frac{|\xi| - \Delta}{\Delta' - \Delta} \right)^3 \right] & \text{if } \Delta \leq |\xi| \leq \Delta' \\ \beta_{\text{min}} & \text{if } |\xi| > \Delta' \end{cases} \tag{30}$$

for all $\xi \in \Xi \triangleq \mathbb{R}$. A visualization of eq. (30) is given in fig. 4.

B.2 Biasing and Confining Potentials

Recall the joint density $\hat{\pi}_t(x, \xi)$ given in eq. (15) as

$$\hat{\pi}_t^{\theta}(x, \xi) = e^{-U_t^{\xi}(x) + F_t^{\theta}(\xi) + \psi_t'(\xi)} = e^{-\tilde{U}_t^{\theta}(x, \xi)}. \tag{31}$$

In practice, and to encourage the ξ-component to stay close to the critical interval $[-\Delta', \Delta']$ (see eq. (30)), we may introduce choose $\psi_t'(\xi)$ so as to include a *confining potential* $\psi^{\text{conf}}(\xi)$. In this paper, we take

$$\psi^{\text{conf}}(\xi) = \begin{cases} \eta(\xi + \tilde{\Delta})^2 & \text{if } \xi < -\tilde{\Delta} \\ 0 & \text{if } -\tilde{\Delta} \leq \xi \leq \tilde{\Delta} \\ \eta(\xi - \tilde{\Delta})^2 & \text{if } \xi > \tilde{\Delta} \end{cases} \tag{32}$$

where $\tilde{\Delta}$ (chosen to be slightly larger than Δ' from eq. (30)) and $\eta > 0$ (denoting the sharpness of the confining potential) are hyperparameters. Additionally, one may consider alleviating a poorly

initialized F_t^θ using an additional biasing force term ψ_t^{bias} updated continuously throughout training to balance out the measured marginal distribution over ξ, so as to obtain the general form

$$\psi_t'(\xi) = \psi^{\text{conf}}(\xi) + \psi_t^{\text{bias}}(\xi). \tag{33}$$

In practice, we find that utilizing a biasing potential $\psi_t^{\text{bias}}(\xi)$ to be cumbersome and to provide little benefit over simpler solutions such as taking e.g., a larger value of σ_ξ^2 in eq. (18).

B.3 Characterizing Minimizers of the Multi-Temperature PINN Loss

Proof of theorem 3.1. We apply the basic argument of (Máté & Fleuret, 2023) in a ξ-pointwise fashion. Let us suppose that $(\mu^\star, F^\star)$ satisfies $\mathcal{L}_{\text{PINN}}(F^\star, \mu^\star; \tilde{\pi}) = 0$. Then, for all $(x, \xi, t) \in \mathbb{R}^d \times \Xi \times \mathcal{T}$ (except perhaps on a set of measure zero), we must have

$$\partial_t F_t^\star(\xi) - \partial_t U_t^\xi(x) - \nabla_x \cdot \mu_t^\star(x, \xi) + (\nabla_x U_t^\xi)^T \mu_t^\star(x, \xi) = 0. \tag{34}$$

Now, observe that the true free energy $F_t(\xi)$ satisfies

$$\partial_t F_t(\xi) = -\partial_t \log Z_t(\xi) = \frac{-\int_{\mathbb{R}^d} \partial_t e^{-U_t^\xi(x)} \,\mathrm{d}x}{Z_t(\xi)} = \frac{\int_{\mathbb{R}^d} (\partial_t U_t^\xi(x)) e^{-U_t^\xi(x)} \,\mathrm{d}x}{Z_t(\xi)}. \tag{35}$$

However, by eq. (34), we have that

$$\begin{aligned} \partial_t U_t^\xi(x, \xi) &= \partial_t F_t^\star(\xi) - \nabla_x \cdot \mu_t^\star(x, \xi) + (\nabla_x U_t^\xi)^T \mu_t^\star(x, \xi) \\ &= \partial_t F_t^\star(\xi) - e^{U_t(x)} \nabla_x \left[\mu_t^\star(x, \xi) e^{-U_t^\xi(x)} \right]. \end{aligned} \tag{36}$$

Plugging eq. (36) into eq. (35) yields

$$\begin{aligned} \partial_t F_t(\xi) &= \frac{\int_{\mathbb{R}^d} \partial_t F_t^\star(\xi) e^{-U_t^\xi(x)} \,\mathrm{d}x}{Z_t(\xi)} - \frac{\int_{\mathbb{R}^d} \nabla_x \left[\mu_t^\star(x) e^{-U_t^\xi(x)} \right] \mathrm{d}x}{Z_t(\xi)} \\ &= \frac{\int_{\mathbb{R}^d} \partial_t F_t^\star(\xi) e^{-U_t^\xi(x)} \,\mathrm{d}x}{Z_t(\xi)} \\ &= \partial_t F_t^\star(\xi), \end{aligned}$$

where the second equality follows from the divergence theorem and the third from the definition of $Z_t(\xi)$. We conclude that $\partial_t F_t^\star(\xi) = \partial_t F_t(\xi)$ on $\mathcal{T}$, and together with the boundary condition $F_0^\star(\xi) = \partial_0 F_t(\xi)$, this is enough to prove the desired result. □

C Controlled Crooks and Jarzynski Results for Continuously Tempered Dynamics

In this section, we establish a controlled Crooks fluctuation theorem for a controlled underdamped Langevin dynamics given by a *non-separable Hamiltonian*, and as a corollary obtain an associated controlled Jarzynski equality. Our results partially generalize existing work for *separable* Hamiltonians (Vaikuntanathan & Jarzynski, 2008; Vargas & Nüsken, 2023; Zhong et al., 2024; Albergo & Vanden-Eijnden, 2024).

C.1 Background on Stochastic Integration

In the remainder of this section, we'll make heavy use of forward and backward Itô integrals, as well as the Stratonovich integral. We largely follow the notation of (Vargas & Nüsken, 2023), and refer thereto for a more elaborate and technical discussion. For now, we summarize the three important Stochastic integrals - forward and backward Itô, and Stratonovich - in fig. 5, and provide some useful identities which relate this integrals, and which we shall make heavy use of. In particular, let us consider the prototypical forward SDE

$$\mathrm{d}\overrightarrow{X}_t = \mu_t(\overrightarrow{X}_t) \,\mathrm{d}t + \sigma_t \,\mathrm{d}W_t, \tag{37}$$

Integral Name	Notation	Discretization
Forward Itô	$\int_0^T X_t \,\mathrm{d}Y_t$	$\sum_i X_{t_i}(Y_{t_{i+1}} - Y_{t_i})$
Backward Itô	$\int_0^T X_t \,\overleftarrow{\mathrm{d}}\, Y_t$	$\sum_i X_{t_{i+1}}(Y_{t_{i+1}} - Y_{t_i})$
Stratonovich	$\int_0^T X_t \circ \mathrm{d}Y_t$	$\sum_i \frac{1}{2}(X_{t_{i+1}} + X_{t_i})(Y_{t_{i+1}} - Y_{t_i})$

Figure 5: Three flavors of stochastic integration.

as shorthand for the forward Itô integral $\overrightarrow{X}_t = \overrightarrow{X}_0 + \int_0^t \mu_s(\overrightarrow{X}_s)\,\mathrm{d}s + \int_0^t \sigma_s \,\mathrm{d}W_s$, and

$$\mathrm{d}\overleftarrow{X}_t = \mu_t(\overleftarrow{X}_t)\,\mathrm{d}t + \sigma_t \,\overleftarrow{\mathrm{d}}\, W_t, \tag{38}$$

as shorthand for the backward Itô integral $\overleftarrow{X}_t = \overleftarrow{X}_T - \int_t^T \mu_s(\overleftarrow{X}_s)\,\mathrm{d}s - \int_t^T \sigma_s \,\overleftarrow{\mathrm{d}}\, W_s$. We conclude with two useful identities which we make heavy use of, and which we provide intuitive, discretized intuition for. We again direct the reader to e.g., (Vargas & Nüsken, 2023; Kunita, 2019) for a more technical treatment. First, as is directly observed from the discretizations shown in fig. 5, we may relate the forward and backward Itô integrals to a corresponding Stratonovich integral via

$$\int_0^T X_t(\mathrm{d}Y_t + \overleftarrow{\mathrm{d}}\, Y_t) = 2\int_0^T X_t \circ \mathrm{d}Y_t. \tag{39}$$

Second, and less intuitively obvious, is the fact that for $X = \overrightarrow{X}$ from eq. (37),

$$\int_0^T \mu_t(X_t)(\overleftarrow{\mathrm{d}}\, X_t - \mathrm{d}X_t) = \lim_{\Delta t \to 0} \sum_i (\mu_{t+\Delta t}(X_{t+\Delta t}) - \mu_t(X_t))(X_{t+\Delta t} - X_t). \tag{40}$$

Writing

$$\begin{aligned} \mu_{t+\Delta t}(X_{t+\Delta t}) &= \mu_t(X_t) + \partial_t \mu_t(X_t) + \nabla \mu_t(X_t)\,\mathrm{d}X_t && (41)\\ &= \mu_t(X_t) + \partial_t \mu_t(X_t) + \nabla \mu_t(X_t)\left[\mu_t(\overrightarrow{X}_t)\,\mathrm{d}t + \sigma_t\,\mathrm{d}W_t\right], && (42) \end{aligned}$$

plugging this into eq. (40), and allowing the contribution of all $o(\Delta t)$ terms to vanish, we obtain

$$\int_0^T \mu_t(X_t)(\overleftarrow{\mathrm{d}}\, X_t - \mathrm{d}X_t) = \int_0^T \sigma_t^2 \nabla \cdot \mu_t(X_t)\,\mathrm{d}t. \tag{43}$$

C.2 Radon-Nikodym Derivatives of Forward and Backward Langevin Dynamics

Let us start by considering some time-dependent *non-separable* Hamiltonian of the form

$$\mathcal{H}_t(q_t, p_t) \triangleq U_t(q_t) + K(q_t, p_t). \tag{44}$$

Then, we may define the controlled underdamped Langevin dynamics as the forward process $\{\overrightarrow{X}_t\}_{t\in[0,1]}$ given by

$$\mathrm{d}\overrightarrow{X}_t = \begin{bmatrix}\mathrm{d}q_t\\ \mathrm{d}p_t\end{bmatrix} = \underbrace{\begin{bmatrix}\mu_t(q_t)\\ 0\end{bmatrix}\mathrm{d}t}_{\text{control}} + \underbrace{\begin{bmatrix}\Gamma_t \nabla_p \mathcal{H}(q_t,p_t)\\ -\Gamma_t \nabla_q \mathcal{H}(q_t,p_t)\end{bmatrix}\mathrm{d}t}_{\text{Hamiltonian dynamics}} + \underbrace{\begin{bmatrix}0\\ -\Gamma_t E_t \partial_p K(q_t,p_t)\end{bmatrix}\mathrm{d}t + \begin{bmatrix}0\\ \sqrt{2\Gamma_t E_T}\end{bmatrix}\mathrm{d}W_t}_{\text{Langevin dynamics}}. \tag{45}$$

where $X_0 \sim \pi_0 \propto e^{-\mathcal{H}_0}$ and where $\{W_t\}$ is a standard Brownian motion on $\mathbb{R}^d$. Here, we have defined the *scaling* and *damping coefficients*

$$\Gamma_t \triangleq \begin{bmatrix}\gamma_t^1 & \dots & 0\\ \vdots & \ddots & \vdots\\ 0 & \dots & \gamma_t^d\end{bmatrix} \quad \text{and} \quad E_t \triangleq \begin{bmatrix}\varepsilon_t^1 & \dots & 0\\ \vdots & \ddots & \vdots\\ 0 & \dots & \varepsilon_t^d\end{bmatrix}. \tag{46}$$

Then, for $T \in (0,1]$ may similarly consider the backward process $\{\overleftarrow{X}_t\}_{t\in[0,1]}$.

$$\mathrm{d}\overleftarrow{X}_t = \begin{bmatrix}\mathrm{d}q_t\\ \mathrm{d}p_t\end{bmatrix} = \begin{bmatrix}\mu_t(q_t)\\ 0\end{bmatrix}\mathrm{d}t + \begin{bmatrix}\Gamma_t\nabla_p\mathcal{H}(q_t,p_t)\\ -\Gamma_t\nabla_q\mathcal{H}(q_t,p_t)\end{bmatrix}\mathrm{d}t - \begin{bmatrix}0\\ -\Gamma_t E_t\partial_p K(q_t,p_t)\end{bmatrix}\mathrm{d}t + \begin{bmatrix}0\\ \sqrt{2\Gamma_t E_T}\end{bmatrix}\overleftarrow{\mathrm{d}}W_t, \tag{47}$$

where $\overleftarrow{X}_T \sim \pi_T \propto e^{-\mathcal{H}_T}$ and where $\overleftarrow{\mathrm{d}}$ denotes the backward Itô differential (see appendix C.1). Let $\mathbb{P}_T$ and $\mathbb{Q}_T$ denote the path measures of $\{\overrightarrow{X}_t\}_{t\in[0,T]}$ and $\{\overleftarrow{X}_t\}_{t\in[0,T]}$. In what follows, we will denote by Q_t and P_t the q and p components respectively of the process X_t, and when necessary, by $\overleftarrow{Q}_t$ and $\overleftarrow{P}_t$ the respective q and p components of the backward process $\overleftarrow{X}_t$. In this section we will establish the following result characterizing the Radon-Nikodym derivative (RND) of the forward and backward processes in eq. (45) and eq. (47).

Theorem C.1 (RND of Forward and Backward Processes)**.** *The path measures $\mathbb{P}_T$ and $\mathbb{Q}_T$ are absolutely continuous with respect to one another, and for a sufficiently well-behaved proces $\{X_t\} = \{(Q_t, P_t\}$, the Radon-Nikodym derivative $\frac{\mathrm{d}\mathbb{Q}_T}{\mathrm{d}\mathbb{P}_T}(X)$ is given by*

$$\log\frac{\mathrm{d}\mathbb{Q}_T}{\mathrm{d}\mathbb{P}_T}(X) = \log\pi_T(X_T) - \log\pi_0(X_0) \tag{48}$$

$$+\int_0^T \nabla_p K(q_t,p_t)\circ\mathrm{d}P_t + \int_0^T \nabla_q\cdot\mu_t(q_t) + (\Gamma_t\nabla_q\mathcal{H}_t(q_t,p_t))^T\nabla_p K(q_t,p_t)\,\mathrm{d}t. \tag{49}$$

One challenging aspect of eq. (45) (and, in turn, of eq. (47)) is the fact that we are only injecting noise into the momentum component, so that the driving Brownian motion W_t has dimension only half that of the phase space on which these two processes are defined. Accordingly, we recall the following Girsanov result which accommodates such diffusion processes involving a lower dimensional driving Brownian motion.

Theorem C.2 (Multivariate Extension of Theorem 7.19, Liptser & Shiryaev (2013))**.** *Let $T \in [0,1]$ and consider the two diffusion processes $\{Y\}_{t\in[0,T]}, \{Z\}_{t\in[0,T]}$, on $\mathbb{R}^n$ given by*

$$\mathrm{d}Y_t = \phi_t(Y_t)\,\mathrm{d}t + \Sigma_t\,\mathrm{d}W_t$$
$$\mathrm{d}Z_t = \psi_t(Z_t)\,\mathrm{d}t + \Sigma_t\,\mathrm{d}W_t$$

with $Y_0 = Z_0$ fixed, so that μ, ϕ, Σ satisfy mild regularity conditions, where $\{W_t\}$ is a standard Brownian motion on $\mathbb{R}^k$, and where $\Sigma_t \in \mathbb{R}^{n\times k}$. Additionally, let $\mathbb{Q}_T^Y$ and $\mathbb{Q}_T^Z$ denote the respective path measures of Y and Z. Then, if

$$\int_0^T \left[\phi_t(X_t)^T(\Sigma_t\Sigma_t^T)^+\phi_t(X_t) + \psi_t(X_t)^T(\Sigma_t\Sigma_t^T)^+\psi_t(X_t)\right]\mathrm{d}t < \infty \tag{50}$$

both $\mathbb{Q}_T^Y$ and $\mathbb{Q}_T^Z$ almost surely, then $\mathbb{Q}_T^Y \sim \mathbb{Q}_T^Z$ are absolutely continuous with respect to one another and their Radon-Nikodym derivative is given by

$$\frac{\mathrm{d}\mathbb{Q}_T^Y}{\mathrm{d}\mathbb{Q}_T^Z}(Z) = \exp\left(\int_0^T (\phi_t(z)-\psi_t(z))^T(\Sigma_t\Sigma_t^T)^+(\mathrm{d}Z_t - \frac{1}{2}(\phi_t(z)+\psi_t(z))\,\mathrm{d}t)\right), \tag{51}$$

where A^+ denotes the Moore-Penrose pseudoinvere of a matrix A.

We may now proceed to the proof of Theorem C.1.

Proof of theorem C.1. First, let us recall the forward process from eq. (45), given as

$$\mathrm{d}\overrightarrow{X}_t = \begin{bmatrix}\mathrm{d}q_t\\ \mathrm{d}p_t\end{bmatrix} = \begin{bmatrix}\mu_t(q_t)\\ 0\end{bmatrix}\mathrm{d}t + \begin{bmatrix}\Gamma_t\nabla_p\mathcal{H}(q_t,p_t)\\ -\Gamma_t\nabla_q\mathcal{H}(q_t,p_t)\end{bmatrix}\mathrm{d}t + \begin{bmatrix}0\\ -\Gamma_t E_t\partial_p K(q_t,p_t)\end{bmatrix}\mathrm{d}t + \begin{bmatrix}0\\ \sqrt{2\Gamma_t E_t}\end{bmatrix}\mathrm{d}W_t, \tag{52}$$

and a corresponding backward process given by

$$\mathrm{d}\overleftarrow{X}_t = \begin{bmatrix}\mu_t(q_t)\\ 0\end{bmatrix}\mathrm{d}t + \begin{bmatrix}\Gamma_t\nabla_p\mathcal{H}(q_t,p_t)\\ -\Gamma_t\nabla_q\mathcal{H}(q_t,p_t)\end{bmatrix}\mathrm{d}t - \begin{bmatrix}0\\ -\Gamma_t E_t\partial_p K(q_t,p_t)\end{bmatrix}\mathrm{d}t + \begin{bmatrix}0\\ \sqrt{2\Gamma_t E_T}\end{bmatrix}\overleftarrow{\mathrm{d}}W_t. \tag{53}$$

whose path measures on $[0, T]$ we have denoted by $\mathbb{P}_T$ and $\mathbb{Q}_T$ respectively. At a high level, we will employ the same trick as was utilized concurrently in both Vargas & Nüsken (2023) and Richter & Berner (2024) of introducing forward and backward references processes with identical path measures, so as to obtain an expression, via cancellation, of the desired RND between the original forward and backward processes. However, the approach must be adapted to handle the singular diffusion coefficient Σ_t. Explicitly, for some choice of drift $\nu' : \mathcal{T} \times \mathbb{R}^{2d} \to \mathbb{R}^d$, let us first consider a forward reference process $\{\overrightarrow{X}_t^{\nu',\mathrm{ref}}\}$ of the form

$$\mathrm{d}\overrightarrow{X}_t^{\nu',\mathrm{ref}} = \begin{bmatrix}\mathrm{d}q_t \\ \mathrm{d}p_t\end{bmatrix} = \begin{bmatrix}\mu_t(q_t) + \Gamma_t \nabla_p \mathcal{H}(q_t, p_t) \\ \nu'_t(q_t, p_t)\end{bmatrix} \mathrm{d}t + \begin{bmatrix} 0 \\ \sqrt{2\Gamma_t E_t}\end{bmatrix} \mathrm{d}W_t, \qquad X_0^{\phi,\mathrm{ref}} \sim \Lambda \tag{54}$$

where Λ denotes the Lebesgue measure on phase space $\mathbb{R}^{2d}$. Let us now choose ν so that $X_t^{\nu',\mathrm{ref}} \sim \Lambda$ is distributed like the Lebesgue measure for all $t \in [0, 1]$. To do so, it is sufficient to take ν' so that the drift in eq. (54) is divergence free, or equivalently, that

$$\nabla_p \cdot \nu'_t(q_t, p_t) = -\nabla_q \cdot (\mu_t(q_t) + \Gamma_t \nabla_p \mathcal{H}_t(q_t, p_t)) \tag{55}$$

This is easily achieved by setting

$$\nu'_t(q_t, p_t) = \nu_t(q_t, p_t) = -\operatorname{diag}(p_t) \operatorname{diag}(\partial_q \mu_t(q_t)) - \Gamma_t \nabla_q \mathcal{H}_t(q_t, p_t) \tag{56}$$

where $\partial_q \mu_t(q_t)$ denotes the Jacobian of μ_t, and where the diagonal operator diag acts on vectors and (square) matrices as

$$\operatorname{diag}(p_t) \triangleq \begin{bmatrix} p_1 & \dots & 0 \\ \vdots & \vdots & 0 \\ 0 & \dots & p_d \end{bmatrix} \in \mathbb{R}^{d\times d}, \quad \text{and} \quad \operatorname{diag}(\partial_q \mu_t(q_t)) \triangleq \begin{bmatrix} (\partial_q \mu_t(q_t))_{11} \\ \vdots \\ (\partial_q \mu_t(q_t))_{dd} \end{bmatrix} \in \mathbb{R}^d, \tag{57}$$

respectively embedding a vector as a diagonal, or extracting from a matrix its diagonal. Plugging $\nu' = \nu$ from eq. (56) into eq. (54), we obtain

$$\mathrm{d}\overrightarrow{X}_t^{\mathrm{ref}} \triangleq \mathrm{d}\overrightarrow{X}_t^{\nu,\mathrm{ref}} = \begin{bmatrix}\mu_t(q_t) + \Gamma_t \nabla_p \mathcal{H}_t(q_t, p_t) \\ \nu(q_t, p_t)\end{bmatrix} \mathrm{d}t + \begin{bmatrix} 0 \\ \sqrt{2\Gamma_t E_t}\end{bmatrix} \mathrm{d}W_t, \qquad X_0^{\mathrm{ref}} \sim \Lambda \tag{58}$$

Denoting by $\mathbb{P}^{\mathrm{ref}}$ the path measure of the process $\overrightarrow{X}^{\mathrm{ref}}$, we may verify that the regularity conditions of theorem C.2 hold, so that we may then apply the result to $Y = \overrightarrow{X}$ and $Z = \overrightarrow{X}^{\mathrm{ref}}$ with

$$\phi_t \triangleq \begin{bmatrix}\mu_t(q_t) + \Gamma_t \nabla_p \mathcal{H}_t(q_t, p_t) \\ -\Gamma_t \nabla_q \mathcal{H}_t(q_t, p_t) - \Gamma_t E_t \nabla_p K(q_t, p_t)\end{bmatrix}, \quad \psi_t \triangleq \begin{bmatrix}\mu_t(q_t) + \Gamma_t \nabla_p \mathcal{H}_t(q_t, p_t) \\ \nu_t(q_t, p_t)\end{bmatrix}, \quad \Sigma_t \triangleq \begin{bmatrix} 0 \\ \sqrt{2\Gamma_t E_t}\end{bmatrix}.$$

It follows that $\mathbb{P}_T \sim \mathbb{P}_T^{\mathrm{ref}}$ are absolutely continuous with respect to one another. Observe now that $\Sigma_t \Sigma_t^T = \begin{bmatrix} 0 & 0 \\ 0 & 2\Gamma_t E_t \end{bmatrix}$, so that its pseudo-inverse is given by

$$(\Sigma_t \Sigma_t^T)^+ = \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{2}(\Gamma_T E_t)^{-1} \end{bmatrix}.$$

Writing $A_t = A_t(q_t, p_t) \triangleq \Gamma_t \nabla_q \mathcal{H}_t(q_t, p_t) + \Gamma_t E_t \nabla_p \mathcal{H}_t(q_t, p_t)$, eq. (51) then immediately yields

$$\log \frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{P}^{\mathrm{ref}}}(X) = \log\left(\frac{\mathrm{d}\pi_0}{\mathrm{d}E_0}\right)(X_0) + \int_0^T \frac{1}{2}(-A_t - \nu_t)^T (\Gamma_t E_t)^{-1} (\mathrm{d}P_t - \frac{1}{2}(-A_t + \nu_t)\,\mathrm{d}t) \tag{59}$$

Let us now turn our attention towards the backward process from eq. (47), which we recall below as

$$\mathrm{d}\overleftarrow{X}_t = \begin{bmatrix}\mu_t(q_t) \\ 0\end{bmatrix} \mathrm{d}t + \begin{bmatrix}\Gamma_t \nabla_p \mathcal{H}_t(q_t, p_t) \\ -\Gamma_t \nabla_q \mathcal{H}_t(q_t, p_t)\end{bmatrix} \mathrm{d}t - \begin{bmatrix} 0 \\ -\Gamma_t E_t \partial_p K(q_t, p_t)\end{bmatrix} \mathrm{d}t + \begin{bmatrix} 0 \\ \sqrt{2\Gamma_t E_t}\end{bmatrix} \overleftarrow{\mathrm{d}} W_t. \tag{60}$$

from which we may construct the reference process

$$\mathrm{d}\overleftarrow{X}_t^{\mathrm{ref}} = \begin{bmatrix}\mu_t(q_t) + \Gamma_t \nabla_p K(p_t) \\ \nu_t(q_t, p_t)\end{bmatrix} \mathrm{d}t + \begin{bmatrix} 0 \\ \sqrt{2\Gamma_t E_t}\end{bmatrix} \overleftarrow{\mathrm{d}} W_t, \qquad X_0^{\mathrm{ref}} \sim \Lambda \tag{61}$$

whose path measure we shall denote by $\mathbb{Q}_T^{\text{ref}}$ and so that $\overleftarrow{X}_t^{\text{ref}} \sim \Lambda$ is distributed like the Lebesgue measure for all $t \in [0, T]$. By introducing $s \triangleq T - t$, we may instead consider the forward (in s) process $\{Y_s\} \triangleq \{\overleftarrow{X}_{1-s}\}$ and $\{Y_s^{\text{ref}}\} \triangleq \{\overleftarrow{X}_{1-s}^{\text{ref}}\}$ given by

$$\mathrm{d}Y_s = -\begin{bmatrix}\tilde{\mu}_s(q_s)\\ 0\end{bmatrix}\mathrm{d}s - \begin{bmatrix}\tilde{\Gamma}_s\nabla_p\mathcal{H}_s(q_s,p_s)\\ -\tilde{\Gamma}_s\nabla_q\mathcal{H}_s(q_s,p_s)\end{bmatrix}\mathrm{d}s + \begin{bmatrix}0\\ -\tilde{\Gamma}_s\tilde{E}_s\nabla_p K(q_s,p_s)\end{bmatrix}\mathrm{d}s + \begin{bmatrix}0\\ \sqrt{2\tilde{\Gamma}_s\tilde{E}_s}\end{bmatrix}\mathrm{d}W_s, \qquad Y_0 \sim \pi_T \propto e^{-\mathcal{H}_T}, \tag{62}$$

and the associated reverse-time reference process

$$\mathrm{d}Y_s^{\text{ref}} = \begin{bmatrix}-\tilde{\mu}_s(q_s) - \tilde{\Gamma}_s\nabla_p\mathcal{H}_s(q_s,p_s)\\ -\tilde{\nu}(q_s,p_s)\end{bmatrix}\mathrm{d}s + \begin{bmatrix}0\\ \sqrt{2\tilde{\Gamma}_s\tilde{E}_s}\end{bmatrix}\mathrm{d}W_s, \qquad Y_0^{\text{ref}} \sim \Lambda \tag{63}$$

where we have defined e.g., $\tilde{\mu}_s = \mu_{T-s}$, and where $Y_s^{\text{ref}} \sim \Lambda$ is inherited from eq. (61). Note that both equations now involve the forward Itô differential d rather than the backward differential $\overleftarrow{\mathrm{d}}$. Denoting by $\tilde{\mathbb{Q}}_T$ and $\tilde{\mathbb{Q}}_T^{\text{ref}}$ the path measures of the time-reversed processes $\{Y_s\}$ and $\{Y_s^{\text{ref}}\}$, we may again verify the necessary regularity conditions and apply theorem C.2 with $Y = \{Y_s\}$ and $Z = \{Y_s^{\text{ref}}\}$ so that we have

$$\phi_s \triangleq \begin{bmatrix}-\tilde{\mu}_s(q_s) - \tilde{\Gamma}_s\nabla_p\tilde{\mathcal{H}}_s(q_s,p_s)\\ \tilde{\Gamma}_s\nabla_q\tilde{\mathcal{H}}_s(q_s,p_s) - \tilde{\Gamma}_s\tilde{E}_s\nabla_p K(q_s,p_s)\end{bmatrix}, \quad \psi_s \triangleq \begin{bmatrix}-\tilde{\mu}_s(q_s) - \tilde{\Gamma}_s\nabla_p\tilde{\mathcal{H}}_s(q_s,p_s)\\ -\tilde{\nu}_s(q_s,p_s)\end{bmatrix}, \quad \Sigma_s \triangleq \begin{bmatrix}0\\ \sqrt{2\tilde{\Gamma}_s\tilde{E}_s}\end{bmatrix}. \tag{64}$$

yielding

$$\log\frac{\mathrm{d}\tilde{\mathbb{Q}}_T}{\mathrm{d}\tilde{\mathbb{Q}}_T^{\text{ref}}}(Y) = \log\left(\frac{\mathrm{d}\pi_T}{\mathrm{d}\Lambda}\right)(X_0) + \int_0^T \frac{1}{2}(-\tilde{C}_s + \tilde{\nu}_s)^T(\tilde{\Gamma}_s\tilde{E}_s)^{-1}(\mathrm{d}P_s - \frac{1}{2}(-\tilde{C}_s - \tilde{\nu}_s)\,\mathrm{d}s). \tag{65}$$

where we have introduced $\tilde{C}_s(q_s,p_s) = C_{T-s}(q_s,p_s)$ with $C_t(q_t,p_t) \triangleq A_t(q_t,p_t) - 2\Gamma_t\nabla_q\mathcal{H}_t(q_t,p_t)$. Denoting by $\mathbb{Q}^{\text{ref}}$ the path measure of $\{\overleftarrow{X}_t^{\text{ref}}\}$, it is easily shown that

$$\log\frac{\mathrm{d}\tilde{\mathbb{Q}}_T}{\mathrm{d}\tilde{\mathbb{Q}}_T^{\text{ref}}}(Y) = \log\frac{\mathrm{d}\mathbb{Q}_T}{\mathrm{d}\mathbb{Q}_T^{\text{ref}}}(\overleftarrow{X}), \tag{66}$$

where $Y_s = \overleftarrow{X}_{T-s}$. It therefore follows from eq. (66), the relation $\tilde{f}_s\,\mathrm{d}\tilde{P}_s = -f_t\,\overleftarrow{\mathrm{d}}\,P_t$, and eq. (65) that

$$\log\frac{\mathrm{d}\mathbb{Q}_T}{\mathrm{d}\mathbb{Q}_T^{\text{ref}}}(X) = \log\left(\frac{\mathrm{d}\pi_T}{\mathrm{d}\Lambda}\right)(X_T) + \int_0^T \frac{1}{2}(-C_t + \nu_t)^T(\Gamma_t E_t)^{-1}(-\overleftarrow{\mathrm{d}}\,P_t - \frac{1}{2}(-C_t - \nu_t)\,\mathrm{d}t). \tag{67}$$

To finish, we may exploit the fact that since $\mathbb{Q}_T^{\text{ref}} = \mathbb{P}_T^{\text{ref}}$ are the same to find then

$$\log\frac{\mathrm{d}\mathbb{Q}_T}{\mathrm{d}\mathbb{P}_T}(X) = \log\frac{\mathrm{d}\mathbb{Q}_T}{\mathrm{d}\mathbb{Q}_T^{\text{ref}}}(X) - \log\frac{\mathrm{d}\mathbb{P}_T}{\mathrm{d}\mathbb{P}_T^{\text{ref}}}(X) - \log\frac{\mathrm{d}\mathbb{P}_T^{\text{ref}}}{\mathrm{d}\mathbb{Q}_T^{\text{ref}}}(X). \tag{68}$$

To see that the last term is zero, it suffices to note that $\overleftarrow{X}_t^{\text{ref}} \sim \Lambda$ is given by the Lebesgue measure for all $t \in [0, T]$, and since the Lebesgue measure has vanishing score, the forward reference process eq. (58) is exactly the time-reversal of the backward reference process eq. (61), and thus $\log\frac{\mathrm{d}\mathbb{P}_T^{\text{ref}}}{\mathrm{d}\mathbb{Q}_T^{\text{ref}}}(X) = 0$. We may therefore plug in eq. (59) and eq. (67) into eq. (68) to obtain

$$\log\frac{\mathrm{d}\mathbb{Q}_T}{\mathrm{d}\mathbb{P}_T}(X) = \log\left(\frac{\mathrm{d}\pi_T}{\mathrm{d}\Lambda}\right)(X_T) - \log\left(\frac{\mathrm{d}\pi_0}{\mathrm{d}\Lambda}\right)(X_0) \tag{69}$$

$$+ \int_0^T \frac{1}{2}(-C_t + \nu_t)^T(\Gamma_t E_t)^{-1}(-\overleftarrow{\mathrm{d}}\,P_t - \frac{1}{2}(-C_t - \nu_t)\,\mathrm{d}t) \tag{70}$$

$$- \int_0^T \frac{1}{2}(-A_t - \nu_t)^T(\Gamma_t E_t)^{-1}(\mathrm{d}P_t - \frac{1}{2}(-A_t + \nu_t)\,\mathrm{d}t). \tag{71}$$

We may now simplify the portions of the integrand correspond to each of the red and blue terms.

Red Terms: The red terms combine to

$$\int_0^T \frac{1}{2}\left[(C_t^T(\Gamma_t E_t)^{-1}\overleftarrow{\mathrm{d}}P_t + A_t^T(\Gamma_t E_t)^{-1}\,\mathrm{d}P_t\right] + \int_0^T \frac{1}{2}\nu_t^T(\Gamma_t E_t)^{-1}(\mathrm{d}P_t - \overleftarrow{\mathrm{d}}P_t). \tag{72}$$

Writing $D_t \triangleq \Gamma_t \nabla_q \mathcal{H}_t(q_t, p_t)$, so that $A_t = C_t + 2D_t$, the first term simplifies as

$$\begin{aligned}
&\int_0^T \frac{1}{2}\left[(C_t^T(\Gamma_t E_t)^{-1}\overleftarrow{\mathrm{d}}P_t + A_t^T(\Gamma_t E_t)^{-1}\,\mathrm{d}P_t\right] \\
&= \frac{1}{2}\int_0^T (A_t - D_t)^T(\Gamma_t E_t)^{-1}(\mathrm{d}P_t + \overleftarrow{\mathrm{d}}P_t) + \frac{1}{2}\int_0^T D_t^T(\Gamma_t E_t)^{-1}(dP_t - \overleftarrow{\mathrm{d}}P_t) \\
&= \frac{1}{2}\int_0^T (\nabla_p K(q_t,p_t))^T(\mathrm{d}P_t + \overleftarrow{\mathrm{d}}P_t) - \int_0^T \nabla_p \cdot D_t\,\mathrm{d}t \\
&= \int_0^T \nabla_p K(q_t,p_t)^T \circ \mathrm{d}P_t - \int_0^T \nabla_p \cdot D_t\,\mathrm{d}t,
\end{aligned} \tag{73}$$

where in the first and second equalities we have utilized the definitions of A_t, C_t, and D_t, in the second we have used eq. (43), and in the third eq. (39). The second term simplifies as

$$\begin{aligned}
\int_0^T \frac{1}{2}\nu_t^T(\Gamma_t E_t)^{-1}(\mathrm{d}P_t - \overleftarrow{\mathrm{d}}P_t) &= -\int_0^T \nabla_p \cdot \nu_t\,\mathrm{d}t \\
&= \int_0^T \nabla_q \cdot \mu_t(q_t) + \nabla_q \cdot (\Gamma_t \nabla_p \mathcal{H}_t(q_t,p_t))\,\mathrm{d}t \\
&= \int_0^T \nabla_q \cdot \mu_t(q_t) + \nabla_p \cdot D_t\,\mathrm{d}t
\end{aligned} \tag{74}$$

where we have utilized eq. (43) and eq. (55). Plugging eq. (73) and eq. (74) into eq. (72) yields

$$\int_0^T \nabla_p K(q_t,p_t)^T \circ \mathrm{d}P_t + \int_0^T \nabla_q \cdot \mu_t(q_t)\,\mathrm{d}t. \tag{★}$$

Blue Terms: The blue terms combine to

$$-\frac{1}{4}\int_0^1 \left[(-C_t+\nu_t)^T(\Gamma_t E_t)^{-1}(-C_t-\nu_t) - (-A_t-\nu_t)^T(\Gamma_t E_t)^{-1}(-A_t+\nu_t)\right]\mathrm{d}t \tag{75}$$

which readily simplifies to

$$\begin{aligned}
&-\frac{1}{4}\int_0^T C_t^T(\Gamma_t E_t)^{-1}C_t - A_t(\Gamma_t E_t)^{-1}A_t\,\mathrm{d}t \\
&= -\frac{1}{4}\int_0^T (A_t - 2D_t)^T(\Gamma_t E_t)^{-1}(A_t - 2D_t) - A_t(\Gamma_t E_t)^{-1}A_t\,\mathrm{d}t \\
&= \int_0^T D_t^T(\Gamma_t E_t)^{-1}(A_t - D_t)\,\mathrm{d}t \qquad (★) \\
&= \int_0^T (\Gamma_t \nabla_q \mathcal{H}_t(q_t,p_t))^T(\Gamma_t E_t)^{-1}(E_t\Gamma_t\nabla_p\mathcal{H}(q_t,p_t))\,\mathrm{d}t \\
&= \int_0^T (\Gamma_t \nabla_q \mathcal{H}_t(q_t,p_t))^T \nabla_p K(q_t,p_t)\,\mathrm{d}t
\end{aligned}$$

Combining eq. (★) with eq. (★), we obtain

$$\log\frac{\mathrm{d}\mathbb{Q}_T}{\mathrm{d}\mathbb{P}_T}(X) = \log\left(\frac{\mathrm{d}\pi_T}{\mathrm{d}\Lambda}\right)(X_T) - \log\left(\frac{\mathrm{d}\pi_0}{\mathrm{d}\Lambda}\right)(X_0) \tag{76}$$

$$+\int_0^T \nabla_p K(p_t)\circ\mathrm{d}P_t + \int_0^T \nabla_q\cdot\mu_t(q_t) + (\Gamma_t\nabla_q\mathcal{H}_t(q_t,p_t))^T\nabla_p K(q_t,p_t)\,\mathrm{d}t. \tag{77}$$

The desired result then follows from the observation that $\log\left(\frac{\mathrm{d}\pi_T}{\mathrm{d}\Lambda}\right)(X_T) - \log\left(\frac{\mathrm{d}\pi_0}{\mathrm{d}\Lambda}\right)(X_0) = \log\pi_T(X_T) - \log\pi_0(X_0)$. □

C.3 Controlled Crooks and Jarzynski Equalities for Non-Separable Langevin Dynamics

We now state and prove a controlled generalization of the Crooks fluctuation theorem (see (Crooks, 1999)) relating the dynamics from eq. (45) and eq. (47) for the non-separable Hamiltonian from eq. (44).

Theorem C.3 (Controlled Crooks Fluctuation Theorem for Non-Separable Langevin Dynamics)**.** *For any $T \in (0, 1]$, and with $\mathbb{P}_T$ and $\mathbb{Q}_T$ defined as as the path measures of eq.* (45) *and eq.* (47) *on the interval $[0, T]$, then for any $X \sim \mathbb{P}_T$, we have*

$$\frac{\mathrm{d}\mathbb{Q}_T}{\mathrm{d}\mathbb{P}_T}(\overrightarrow{X}) = \exp\left(F_T - F_0 + \int_0^T \nabla_q \cdot \mu_t(q_t) - \partial_t \log U_t(q_t) - \mu_t(q_t)^T \nabla_q \mathcal{H}_t(q_t, p_t)\,\mathrm{d}t\right). \tag{78}$$

Proof. By Theorem C.1,

$$\log \frac{\mathrm{d}\mathbb{Q}_T}{\mathrm{d}\mathbb{P}_T}(X) = \log \pi_T(X_T) - \log \pi_0(X_0) \tag{79}$$

$$+ \int_0^T \nabla_p K(q_t, p_t) \circ \mathrm{d}P_t + \int_0^T \nabla_q \cdot \mu_t(q_t) + (\Gamma_t \nabla_q \mathcal{H}_t(q_t, p_t))^T \nabla_p K(q_t, p_t)\,\mathrm{d}t. \tag{80}$$

Observe now that via the Stratonovich formulation of Itô's lemma applied to $\log \hat{\pi}_t = \log \pi_t + F_t$,

$$\log \hat{\pi}_1(X_1) - \log \hat{\pi}_0(X_0) = \int_0^T \partial_t \log \hat{\pi}_t(X_t)\,\mathrm{d}t + \int_0^T \partial_x \log \hat{\pi}_t(X_t) \circ \mathrm{d}X_t \tag{81}$$

$$= -\int_0^T \partial_t \mathcal{H}_t(q_t, p_t)\,\mathrm{d}t - \int_0^T (\nabla_q \mathcal{H}_t(q_t, p_t))^T \circ \mathrm{d}Q_t - \int_0^T (\nabla_p \mathcal{H}_t(q_t, p_t))^T \circ \mathrm{d}P_t \tag{82}$$

$$= -\int_0^T \partial_t U_t(q_t)\,\mathrm{d}t - \int_0^T (\nabla_q \mathcal{H}_t(q_t, p_t))^T \circ \mathrm{d}Q_t - \int_0^T (\nabla_p K(q_t, p_t))^T \circ \mathrm{d}P_t. \tag{83}$$

Since $X \sim \mathbb{P}_T$ is given by eq. (45), we may plug in

$$\mathrm{d}Q_t = [\mu_t(q_t) + \Gamma_t \nabla_p \mathcal{H}_t(q_t, p_t)]\,\mathrm{d}t = [\mu_t(q_t) + \Gamma_t \nabla_p K(q_t, p_t)]\,\mathrm{d}t \tag{84}$$

to obtain

$$\int_0^T (\Gamma_t \nabla_q \mathcal{H}(q_t, p_t))^T \nabla_p K(q_t, p_t)\,\mathrm{d}t = \int_0^T \nabla_q \mathcal{H}_t(q_t, p_T)^T \circ \mathrm{d}Q_t - \int_0^T \mu_t(q_t)^T \nabla_q \mathcal{H}_t(q_t, p_t)\,\mathrm{d}t. \tag{85}$$

Equation 83 thus implies that

$$\int_0^T (\Gamma_t \nabla_q \mathcal{H}(q_t, p_t))^T \nabla_p K(q_t, p_t)\,\mathrm{d}t = \log \hat{\pi}_0(X_0) - \log \hat{\pi}_1(X_1) - \int_0^T \partial_t \log U_t(q_t)\,\mathrm{d}t \tag{86}$$

$$- \int_0^T (\nabla_p \mathcal{H}_t(q_t, p_t))^T \circ \mathrm{d}P_t - \int_0^T \mu_t(q_t)^T \nabla_q \mathcal{H}_t(q_t, p_t)\,\mathrm{d}t. \tag{87}$$

Plugging Equation 87 back into Equation 80 we have

$$\log \frac{\mathrm{d}\mathbb{Q}_T}{\mathrm{d}\mathbb{P}_T}(X) = \log \pi_T(X_T) - \log \pi_0(X_0) + \log \hat{\pi}_0(X_0) - \log \hat{\pi}_T(X_T)$$

$$+ \int_0^T \nabla_p K(q_t, p_t) \circ \mathrm{d}P_t + \int_0^T \nabla_q \cdot \mu_t(q_t)\,\mathrm{d}t$$

$$- \int_0^T \partial_t \log U_t(q_t)\,\mathrm{d}t - \int_0^T \nabla_p K(q_t, p_t) \circ \mathrm{d}P_t - \int_0^T \mu_t(q_t)^T \nabla_q \mathcal{H}_t(q_t, p_t)\,\mathrm{d}t.$$

Simplifying then yields

$$\log \frac{\mathrm{d}\mathbb{Q}_T}{\mathrm{d}\mathbb{P}_T}(X) = F_T - F_0 + \int_0^T \nabla_q \cdot \mu_t(q_t) - \partial_t \log U_t(q_t) - \mu_t(q_t)^T \nabla_q \mathcal{H}_t(q_t, p_t)\,\mathrm{d}t \tag{88}$$

from which the desired result immediately follows. □

We finish by noting that the Radon-Nikodym derivative of Theorem C.3 allows us to importance sample from $\mathbb{Q}_T$ using $\mathbb{P}_T$, as is formalized in the following controlled Jarzynski equality.

Corollary C.4 (Controlled Jarzynski Equality)**.** *For $T \in (0, 1]$, let $h \in \mathcal{C}^1([0, T], \mathbb{R})$ denote some observable. Then*

$$\mathbb{E}_{X\sim\mathbb{Q}_T}\left[h(X)\right] = \frac{\mathbb{E}_{X\sim\mathbb{P}_T}\left[h(X)\exp\left(A_t(X)\right)\right]}{\mathbb{E}_{X\sim\mathbb{P}_T}\left[\exp\left(A_t(X)\right)\right]}. \tag{89}$$

where the work functional $A_T(X)$ is given by

$$A_T(X) \triangleq \int_0^T \nabla_q \cdot \mu_t(q_t) - \partial_t \log U_t(q_t) - \mu_t(q_t)^T \nabla_q \mathcal{H}_t(q_t, p_t)\,\mathrm{d}t. \tag{90}$$

In particular,

$$\mathbb{E}_{X\sim\mathbb{P}_T}\left[\exp(A_T(X))\right] = \exp(F_0 - F_T) = \frac{Z_T}{Z_0}. \tag{91}$$

A similar result is established in (Albergo & Vanden-Eijnden, 2024), for an overdamped Langevin dynamics, and in the special case that $h(X) = h(X_T)$ is a function of the terminal point.

Proof. Observe that by theorem C.3,

$$\mathbb{E}_{X\sim\mathbb{P}_T}\left[\exp(A_T(X))\right] = \exp(F_0 - F_T)\mathbb{E}_{X\sim\mathbb{P}_T}\left[\frac{\mathrm{d}\mathbb{Q}_T}{\mathrm{d}\mathbb{P}_T}(X)\right] = \exp(F_0 - F_T).$$

Thus,

$$\begin{aligned}\frac{\mathbb{E}_{X\sim\mathbb{P}_T}\left[h(X)\exp\left(A_T(X)\right)\right]}{\mathbb{E}_{X\sim\mathbb{P}_T}\left[\exp\left(A_T(X)\right)\right]} &= \exp(F_0 - F_T)\mathbb{E}_{X\sim\mathbb{P}_T}\left[h(X)\exp\left(A_t(X)\right)\right] \\ &= \mathbb{E}_{X\sim\mathbb{P}_T}\left[h(X)\frac{\mathrm{d}\mathbb{Q}_T}{\mathrm{d}\mathbb{P}_T}(X)\right] \\ &= \mathbb{E}_{X\sim\mathbb{Q}_T}\left[h(X)\right],\end{aligned}$$

as desired. □

C.4 Specialization to the Continuously Tempered Setting

Proof of theorem 3.2. Recall the CTDS dynamics from eq. (17), which may be rewritten in the form of eq. (45) as

$$\mathrm{d}X_t^{\text{CTDS}} = \begin{bmatrix}\mathrm{d}q_t \\ \mathrm{d}p_t\end{bmatrix} = \underbrace{\begin{bmatrix}\mu_t^\theta(q_t) \\ 0\end{bmatrix}\mathrm{d}t}_{\text{control}} + \underbrace{\begin{bmatrix}\Gamma_t \nabla_p \mathcal{H}_t(q_t, p_t) \\ -\Gamma_t \nabla_q \mathcal{H}_t(q_t, p_t)\end{bmatrix}\mathrm{d}t}_{\text{Hamiltonian dynamics}} + \underbrace{\begin{bmatrix}0 \\ -\Gamma_t E_t \partial_p K(q_t, p_t)\end{bmatrix}\mathrm{d}t + \begin{bmatrix}0 \\ \sqrt{2\Gamma_t E_T}\end{bmatrix}\mathrm{d}W_t}_{\text{Langevin dynamics}}. \tag{92}$$

where $q_t = (X_t, \xi_t)$, $p_t = (P_t^x, P_t^\xi)$, $U_t(q_t) = \tilde{U}_t^\theta(X_t, \xi_t)$, and $K(q_t, p_t) = \frac{\beta(\xi_t)}{2M_x}\|P_t^x\|^2 + \frac{1}{2M_\xi}\|P_t^\xi\|^2$, and

$$\Gamma_t \triangleq \begin{bmatrix}\gamma_t^x I_d & 0 \\ 0 & \gamma_t^\xi\end{bmatrix} \qquad \text{and} \qquad E_t \triangleq \begin{bmatrix}\varepsilon_t^x I_d & 0 \\ 0 & \varepsilon_t^\xi\end{bmatrix}. \tag{93}$$

Thus, theorem 3.2 follows immediately from theorem C.4 in the special case that $h(X) = h(X_T)$ is a function of the terminal point. □

Now recall from eq. (15) that

$$\hat{\pi}_t^\theta(x,\xi) = e^{-U_t^\xi(x)+F_t^\theta(\xi)+\psi_t'(\xi)} = e^{-\tilde{U}_t^\theta(x,\xi)}, \tag{94}$$

and define by $\pi_t^\theta(x,\xi)$ the normalized version of $\hat{\pi}_t^\theta(x,\xi)$. Additionally recall that $\mathbb{P}_t$ and $\mathbb{Q}_t$ denote the forward and backward path measures on the interval $[0,t]$, for $t \in [0,1]$. In particular, if $X \sim \mathbb{Q}_t$, then $X_t \sim \pi_t^\theta$. Then, theorem 3.2 may be applied to reweight the multi-temperature PINN objective from eq. (13) via

$$\begin{aligned}
\mathcal{L}_{\text{PINN}}^{\text{MT}}(F,\mu;\pi^\theta) &\triangleq \int_{\mathcal{T}} \mathbb{E}_{(x,z)\sim\pi^\theta}\left[|\partial_t F_t - \partial_t \tilde{U}_t^\theta - \nabla_x \cdot \mu_t + (\nabla_x \tilde{U}_t^\theta)^T \mu_t|^2\right] \mathrm{d}t \\
&= \int_{\mathcal{T}} \mathbb{E}_{Q\sim\mathbb{Q}_t}\left[|\partial_t F_t - \partial_t \tilde{U}_t^\theta - \nabla_x \cdot \mu_t + (\nabla_x \tilde{U}_t^\theta)^T \mu_t|^2\right] \mathrm{d}t \\
&= \int_{\mathcal{T}} \frac{\mathbb{E}_{Q\sim\mathbb{P}_t}\left[|\partial_t F_t - \partial_t \tilde{U}_t^\theta - \nabla_x \cdot \mu_t + (\nabla_x \tilde{U}_t^\theta)^T \mu_t|^2 \exp(A_t(Q))\right]}{\mathbb{E}_{Q\sim\mathbb{P}_t}\left[\exp(A_t(Q))\right]} \mathrm{d}t.
\end{aligned} \tag{95}$$

yielding eq. (22). To arrive at the expression for $A_t(Q)$ (from eq. (19)) from the general work functional in eq. (90), we have used the fact the ξ-component of the control $\mu_t^\theta(q_t)$ is zero to obtain $\nabla_q \cdot \mu_t^\theta = \nabla_x \cdot \mu_t^\theta$, and this together with the fact that the kinetic energy depends only on the temperature and momenta to obtain that $(\mu_t^\theta)^T \nabla_q \mathcal{H}_t^\theta = (\mu_t^\theta)^T \nabla_x \tilde{U}_t^\theta$.

D Experimental Details

D.1 40-Mode Gaussian Mixture

Additional Training Details. At train time, we consider the four proposal types as outlined in fig. 6. For each, we parameterize the learned control $\mu_t^\theta(x)$ ($\mu_t^\theta(x,\beta(\xi))$ for CTDS), the free energy F_t^θ ($F_t^\theta(\beta(\xi))$ for CTDS), and the learned potential U_t^θ as feed-forward neural networks with width 256 and depth three, and using the SiLU non-linearity (Elfwing et al., 2017). We train using a replay buffer, re-sampling once per epoch for 1250 epochs for a total of 125000 training iterations at which the PINN loss is evaluated at 6250 randomly sampled elements of the buffer. We utilize the Adam optimizer (Kingma & Ba, 2017) with learning rate 1×10^{-3}, reducing by a factor of $\gamma = 0.97$ every 1000 iterations, and after an initial burn-in period of 15000 training iterations. Additionally, we follow the lead of (Wang et al., 2023; Albergo & Vanden-Eijnden, 2024) in utilizing curriculum-based training whereby at $T \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$ (in that order) we spend $\{1000, 1000, 1000, 1000, 2000, 2000, 2000, 3000, 3000, 3000\}$ iterations respectively, and the remaining iterations at $T = 1.0$. For CTDS, we reparameterize using eq. (30) with $\beta_{\min} = 0.2$, $\Delta = 0.25$, and $\Delta' = 1.9$, and additionally set $\psi_t'(\xi) = \psi^{\text{conf}}(\xi)$ as in eq. (32) with $\eta = 10.0$ and $\tilde{\Delta} = 2.0$. Finally, we utilize Gaussian Fourier features as in (Tancik et al., 2020) to encode position, time, and temperature with 100, 20, and 20 features respectively, drawn from isotropic Gaussians with standard deviations 0.1, 5, and 1, respectively.

Sampling Details. In practice, we discretize using an Euler solver with with $\Delta t = 0.004$ (250 timesteps). Letting $\phi^\theta : \mathbb{R}^d \to \mathbb{R}^d$, $\phi^\theta : X_0 \mapsto X_1$ denote the corresponding unit-time flow of eq. (23), we sample from the pushforward $\tilde{\pi}^\theta \triangleq [\phi^\theta]_\sharp \mathcal{N}(0, 5.0 I_2)$. The density $\tilde{\pi}^\theta$ may then be computed via the continuous change of variables formula as

$$\log \tilde{\pi}^\theta(X_1) = \log \pi_0(X_0) - \int_0^1 \nabla \cdot \mu_t^\theta(X_t)\, \mathrm{d}t. \tag{96}$$

Evidence Lower Bound. Letting $\tilde{\pi}^\theta$ denote our sample density as given in eq. (96), and $\hat{\pi}_1$ the unnormalized target, we define the evidence lower bound (ELBO) as

$$\mathbb{E}_{x\sim\tilde{\pi}^\theta}\left[\log\left(\frac{\hat{\pi}_1(x)}{\tilde{\pi}^\theta(x)}\right)\right] = -D_{\text{KL}}(\tilde{\pi}^\theta \parallel \pi_1) + \log Z \leq \log Z, \tag{97}$$

where $Z = \int \hat{\pi}_1(x)\, \mathrm{d}x$ denotes the partition function of $\hat{\pi}$.

Name	Dynamics	Hyperparameters	Density Path
Reference	$\mathrm{d}X_t = \mu_t^\theta(X_t)\,\mathrm{d}t$	N/A	$U_t = (1-t)U_0 + tU_1 + t(1-t)U_t^\theta$
NETS (OD)	eq. (4)	$\varepsilon_t = 50.0$	$U_t = (1-t)U_0 + tU_1 + t(1-t)U_t^\theta$
NETS (UD)	eq. (3)	$\gamma_t = 50.0$, $\varepsilon_t = 2.0$, $M = 1.0$	$U_t = (1-t)U_0 + tU_1 + t(1-t)U_t^\theta$
CTDS	eq. (17)	$\gamma_t^x = 50.0$, $\varepsilon_t^x = 2.0$, $\gamma_t^\xi = 5.0$, $\varepsilon_t^\xi = 2.0$, $M_x = M_\xi = 1.0$	$U_t^\xi = (1-t)U_0^\xi + tU_1^\xi + \beta(\xi)t(1-t)U_t^\theta$

Figure 6: The proposals and their accompanying density paths for the 40-mode Gaussian mixture experiment described in appendix D.1.

Evidence Upper Bound. Following Blessing et al. (2024), we define the evidence upper bound (EUBO) as

$$\mathbb{E}_{x\sim\pi_1}\left[\log\left(\frac{\hat{\pi}_1(x)}{\tilde{\pi}^\theta(x)}\right)\right] = D_{\mathrm{KL}}(\pi_1 \parallel \tilde{\pi}^\theta) + \log Z \geq \log Z, \tag{98}$$

where Z is defined as before.