INTER-TRAJECTORY IMPORTANCE SAMPLING: A PRINCIPLED MITIGATION OF MODE COLLAPSE

Anonymous authors
Paper under double-blind review

ABSTRACT

Sampling from unnormalized distributions is a fundamental challenge in computational science. Diffusion samplers typically approach it by minimizing the reverse Kullback–Leibler divergence between path measures but remain susceptible to mode collapse. Conventional single-trajectory estimators evaluate trajectories independently, without directly assessing sample diversity. We introduce Inter-Trajectory Importance Sampling (ITIS), a multi-sample estimator that couples trajectory losses through path-measure importance weighting. ITIS remains unbiased, preserves the expected gradient and critical points, requires no additional forward passes or heuristic inter-particle forces, and yields i.i.d. samples at inference. At the finite-batch level, it assigns lower realized losses to diverse trajectory configurations than to collapsed ones. We characterize theoretically how this method mitigates mode collapse and we demonstrate its efficacy experimentally. As a drop-in loss replacement, ITIS improves established diffusion samplers on standard benchmarks and interacting-particle systems.

1 INTRODUCTION

Efficiently sampling from unnormalized densities remains a pivotal bottleneck across diverse scientific disciplines, ranging from computational chemistry (Faulkner & Livingstone, 2024) and Bayesian inference (Blei et al., 2017) to reinforcement learning (Levine, 2018) and the alignment of generative models via reward-based fine-tuning (Uehara et al., 2024; Domingo-Enrich et al., 2025; 2026). The shared problem in these domains is to generate samples from a distribution $\pi_T \propto \tilde{\pi}_T$ whose unnormalized density $\tilde{\pi}_T$ can be evaluated pointwise but whose normalization constant is intractable. Motivated by these challenges, the application of machine learning to the sampling problem has become an active research area (Noé et al., 2019; Midgley et al., 2023; Blessing et al., 2024; Havens et al., 2025). The overarching goal of these *neural samplers* is to *amortize* the cost of sampling. By investing computational effort once during training, the learned sampler can subsequently generate samples from π_T at a fraction of the per-sample cost incurred by traditional non-amortizing methods such as Markov Chain Monte Carlo (MCMC) (Hastings, 1970; Faulkner & Livingstone, 2024), Annealed Importance Sampling (Neal, 2001), or more recently Sequential Monte Carlo (SMC) (Del Moral et al., 2006).

Diffusion samplers and mode collapse. Generative diffusion models have demonstrated state-of-the-art performance on learning complex high-dimensional distributions from data samples, producing photorealistic images and coherent text at scales that establish them as among the most powerful generative models known (Sohl-Dickstein et al., 2015; Ho et al., 2020; Song et al., 2021b). *Diffusion samplers* adapt their generative process to the unnormalized sampling problem, where the absence of data samples from the target distribution necessitates a conceptually distinct training paradigm. Rather than fitting a drift to data via data-based score matching techniques, the drift is optimized with respect to a stochastic optimal control problem (Tzen & Raginsky, 2019; Zhang & Chen, 2022; Vargas et al., 2023; Berner et al., 2024). Solving this problem is equivalent to minimizing the reverse Kullback-Leibler (rKL) divergence between the path measure associated to the drift and a reference path measure whose marginal distribution is π_T. In this learning problem, point-wise evaluations of $\tilde{\pi}_T$ provide the only supervision. The rKL is known to be mode-seeking, i.e. it only penalizes the learned process for placing mass where the target π_T has none, but ignores regions of non-vanishing π_T that the learned process fails to cover. As a result, training tends to converge to a solution that captures only a subset of the modes of π_T while leaving others entirely unvisited, a failure known as

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

mode collapse (Minka, 2005; Bishop, 2006). While diffusion samplers represent the most popular family of neural samplers, their successful application in challenging sampling problems is hampered by mode collapse (Blessing et al., 2024; Richter & Berner, 2024; He et al., 2025b; Havens et al., 2025). Existing remedies predominantly address mode collapse either algorithmically, by hybridizing neural samplers with traditional Monte Carlo methods, or through architectural and procedural modifications. However, they leave the training signal itself untouched, and with it the fact that the standard estimator of the rKL cannot tell mode-collapsed from mode-covering solutions.

Inter-Trajectory Importance Sampling. The central observation guiding this work is that mode collapse is intrinsically a *multi-sample* phenomenon. It describes the relationship between the modes visited by a *collection* of sampling trajectories. The rKL landscape admits mode-collapsed local minima alongside mode-covering ones, yet, to the best of our knowledge, every existing diffusion sampler estimates the rKL by aggregating *single-trajectory* terms, each of which is blind to where the other trajectories went. We introduce *Inter-Trajectory Importance Sampling* (ITIS), a multi-sample estimator of the same objective that couples the loss contributions of a batch of simultaneously generated trajectories by reweighting each contribution with the transition kernels of the concurrently generated trajectories. ITIS is unbiased, so the expected objective and its critical points are unchanged and no heuristic interactions are introduced. The coupling instead acts on realized batches, where the loss is smaller for diverse than for collapsed trajectory configurations.

We show that ITIS implements a principled mechanism that favors mode-covering solutions in gradient-based training and helps retain them once reached. Importantly, we demonstrate the efficacy of this mechanism on dedicated problems that are prone to mode collapse. Finally, we demonstrate that ITIS serves as a drop-in replacement for the single-trajectory rKL losses of established diffusion samplers, requiring no changes to architecture or parametrization and improving performance on standard benchmarks as well as on interacting-particle systems.

2 Related Work

Neural Samplers. A wide family of neural samplers has been proposed, including flow-based methods (Noé et al., 2019; Midgley et al., 2023; Albergo & Vanden-Eijnden, 2025; Tian et al., 2024), discrete neural samplers (Sanokowski et al., 2025a; Holderrieth et al., 2025), and score-based methods that rely on off-line learning from buffers to evade scalability issues related to learning directly from simulations (Akhound-Sadegh et al., 2024). This work focuses on diffusion samplers, which represent a particularly active area of research. **Diffusion Samplers.** Grounded in stochastic optimal control and the Schrödinger bridge problem (Pavon, 1989; Tzen & Raginsky, 2019; Chen et al., 2016; Berner et al., 2024), foundational diffusion samplers include path integral samplers (PIS) (Zhang & Chen, 2022), time-reversed diffusion sampler (DIS) (Berner et al., 2024), and denoising diffusion samplers (DDS) (Vargas et al., 2023). These methods learn the parameters of a controlled (stochastic differential equation) SDE such that the terminal marginal of the induced path measure approximates the unnormalized target distribution π_T, typically by minimizing the rKL divergence between the controlled path measure and a reference path measure Blessing et al. (2024). Subsequent work expanded these methods along the following axes. **Combination with Monte Carlo methods.** Several works combine diffusion samplers with classical MC techniques, Doucet et al. (2022) embed a diffusion model inside an annealed importance sampling chain, Phillips et al. (2024b) insert SMC resampling steps between diffusion transitions, Chen et al. (2025) interleave controlled diffusion with Langevin corrections in an SMC framework.

Modified training objectives. Vargas et al. (2024) introduce an approach that learns shared parameters of the forward and reverse process simultaneously. He et al. (2025a) replace the per-step KL with a diffusive rKL, and Sendera et al. (2024) decouple the behaviour and target policies for improved off-policy training. Guo et al. (2026) introduce proximal regularization near the target density, while Blessing et al. (2025a) impose trust-region constraints on measure transport in path space. **Modified gradient estimators.** Richter & Berner (2024) introduce a framework that connects DIS, DDS, and PIS naturally to Schrödinger Bridges and they introduce a low-variance loss for diffusion samplers that enables optimizing arbitrary divergences. Sanokowski et al. (2025b) point out theoretical problems and practical challenges in the application of this loss to Schrödinger Bridges, and they propose a REINFORCE-based alternative that is theoretically grounded and which

yields improved experimental results. **Modified parametrization and Schrödinger bridge formulations.** Blessing et al. (2025b) introduce underdamped Langevin dynamics to diffusion samplers for improved mixing, while Havens et al. (2025) scale to molecular systems via adjoint matching (Domingo-Enrich et al., 2025), enabling a highly scalable simulation-free training objective. Liu et al. (2025b) use stochastic optimal control to derive a method that enables the application of adjoint matching to Schrödinger bridge samplers. Blessing et al. (2026) recast bridge training as a scalable fixed-point problem using Nelson's relation.

Mode collapse is identified as a central obstacle to deploying diffusion samplers on challenging targets in works that systematically analyze their limitations (Blessing et al., 2024; He et al., 2025b). To the best of our knowledge, our work is the first to address this problem in diffusion samplers by departing from the paradigm of aggregating independent single-trajectory loss estimates, instead developing an objective that is intrinsically sensitive to diversity across simultaneously generated trajectories. **Interacting Particle Methods.** A natural strategy against mode collapse is to simulate multiple samples simultaneously and enforce pairwise repulsion. Stein Variational Gradient Descent (SVGD) (Liu & Wang, 2016) is the canonical such method, combining attraction toward the target log-density with a kernel-gradient repulsion term that spreads particles across modes. The Ensemble Kalman Sampler (EKS) (Garbuno-Inigo et al., 2020) pursues the same goal via a gradient-flow formulation in which particles interact through their empirical covariance rather than a kernel. Both SVGD and EKS neither amortize the sampling cost nor yield independent samples.

A related electrostatic principle was made explicit by Unterthiner et al. (2018) via the Coulomb GAN model. Its generator outputs are treated as repelling charges in the attractive Coulomb field of the real data. The requirement of real data restricts this method to the problems where samples from the target distribution are available. Corso et al. (2024) transfer the same principle to diffusion models with *Particle Guidance*: the reverse SDE is augmented by a time-evolving repulsive joint potential. Zilberstein et al. (2025) apply a related RBF repulsion to diversify latent diffusion samplers for inverse problems. Both methods couple particles *at inference time*, producing correlated outputs. ITIS is, to our knowledge, the first multi-sample training objective that yields i.i.d. samples at inference.

3 Problem Description

Formally, sampling from an unnormalized target distribution as described in Sec. 1 can be denoted

$$\pi_T(X) = \frac{\tilde{\pi}_T(X)}{\mathcal{Z}}, \quad \mathcal{Z} = \int_\Omega \tilde{\pi}_T(X) \mathrm{d}X \tag{1}$$

where $\mathcal{Z}$ is the unknown and computationally intractable normalization constant. In this setting, it is assumed that we do not have access to any samples from π_T but can evaluate $\tilde{\pi}_T(X)$ for any sample candidate X. The task of sampling from distributions as described in Eq. (1) can be approached as a Variational Inference problem by using a diffusion model / sampler in the following way. Diffusion models are generative models that learn a transport map from samples $X_0 \in \mathbb{R}^d$ drawn from a simple prior distribution π_0 to samples X_T distributed according to a target distribution π_T. The transport map is defined by the reverse diffusion process, whose parameters are learned to invert a forward diffusion process. The parameters of the corresponding forward SDE may be learned or predefined. Examples of the latter include the variance-exploding and variance-preserving SDEs (Song et al., 2021a). In diffusion samplers, the forward and reverse diffusion processes are defined by the following SDEs

$$\mathrm{d}X_t = f(X_t, t)\,\mathrm{d}t + \sigma_t\,\mathrm{d}W_t, \qquad \mathrm{d}X_t = r_\theta(X_t, t)\,\mathrm{d}t + \sigma_t\,\mathrm{d}W_t$$

with diffusion coefficient σ_t and Brownian motion W_t, where θ denotes the parameters of the function, which are typically learned. Using an Euler–Maruyama scheme, their discretized counterparts can be characterized by Gaussian transition kernels

$$q(X_{t-1}|X_t) \sim \mathcal{N}(X_{t-1}; X_t + f(X_t, t)\,\Delta t, \sigma_t^2 \Delta t)$$
$$p_\theta(X_t|X_{t-1}) \sim \mathcal{N}(X_t; X_{t-1} + r_\theta(X_{t-1}, t-1)\,\Delta t, \sigma_{t-1}^2 \Delta t)$$

with discretization step size Δt.

Diffusion samplers are typically trained by minimizing the rKL between the finite-dimensional distribution of forward and reverse SDE path

$$q(X_{0:T}) = \pi_T(X_T) \prod_{t=1}^{T} q(X_{t-1}|X_t)), \quad p_\theta(X_{0:T}) = \pi_0(X_0) \prod_{t=1}^{T} p_\theta(X_t|X_{t-1})$$

such that

$$D_{\mathrm{KL}}(p_\theta(X_{0:T}) \,||\, q(X_{0:T})) = \mathbb{E}_{X_{0:T}\sim p_\theta}\left[\log \frac{p_\theta(X_{0:T})}{q(X_{0:T})}\right] = \mathbb{E}_{X_{0:T}\sim p_\theta}\left[\log \frac{\pi_0(\cdot)}{q(\cdot \mid X_1)}\right]$$
$$+ \sum_{t=1}^{T-1} \mathbb{E}_{X_{0:T}\sim p_\theta}\left[\log \frac{p_\theta(\cdot \mid X_{t-1})}{q(\cdot \mid X_{t+1})}\right] + \mathbb{E}_{X_{0:T}\sim p_\theta}\left[\log \frac{p_\theta(\cdot \mid X_{T-1})}{\pi_T(\cdot)}\right]. \tag{2}$$

For notational simplicity, the symbol p_θ is used to denote both the joint distribution $p_\theta(X_{0:T})$ and the conditional distribution $p_\theta(\cdot \mid X_{t-1})$. Marginal distributions are denoted by p_{θ,X_t}. The same conventions apply analogously to q. In practice, the outer expectation over paths in Eq. (2) is approximated with a single sample per trajectory. The rKL between the Gaussian transition kernels is evaluated in closed form, whereas the terminal term, which involves π_T, is evaluated at the sampled endpoint X_T^k. The resulting loss of trajectory k is a sum of per-step costs that depend only on $X_{0:T}^k$,

$$\mathcal{L}_{\mathrm{Single}}(X_{0:T}^k) = \sum_{t=1}^{T} C(X_t^k),$$

with $C(\cdot)$ being a generic cost function and the cost for DDS shown in Sec. B. The loss of each trajectory is therefore unaffected by where the other trajectories of the batch are, in particular by whether they occupy the same mode or different ones. Although the rKL is minimized by the target itself, its mode-seeking nature gives rise to mode-collapsed local minima, and this estimator provides no signal that distinguishes them from mode-covering ones.

4 METHOD

4.1 INTER-TRAJECTORY IMPORTANCE SAMPLING

In the following we denote the collection of random variables $\{X_{0:T}^k\}_{k=0}^{K-1}$ with uppercase letters whereas respective realized sample-paths thereof $\{x_{0:T}^k\}_{k=0}^{K-1} \overset{\text{iid}}{\sim} p_\theta$ are denoted in lowercase. The single-sample loss can then be estimated as

$$\mathcal{L}_{\mathrm{Single}}(\{X_{0:T}^k\}_{k=0}^{K-1}) \approx \sum_{t=1}^{T} \frac{1}{K} \sum_{k=0}^{K-1} \mathbb{E}_{X_t^k \sim p_\theta(X_t^k|x_{t-1}^k)}[C(X_t^k)] \approx \sum_{t=1}^{T} \frac{1}{K} \sum_{k=0}^{K-1} C(x_t^k). \tag{3}$$

To streamline notation, we define

$$p_k(\,\cdot\,) \;:=\; p_\theta\big(\,\cdot \,\big|\, X_{t-1}^k\big)\,, \quad w_{mk}(\,\cdot\,) \;:=\; p_k(\,\cdot\,)/p_m(\,\cdot\,) \tag{4}$$

representing the transition kernels from trajectory k and corresponding importance weights. The associated time indices are implicit and can be inferred from the random variable X_t. The importance weights become $w_{mk} := \frac{p_\theta(x_t^m|x_{t-1}^k)}{p_\theta(x_t^m|x_{t-1}^m)}$ when using realizations. We recast (3) in terms of an expectation with respect to the sampling distribution of a reference trajectory. We approximate this expectation using a single sample $m \in \{0, \ldots, K-1\}$ in each trajectory step t

$$\mathcal{L}(\{X_t^k\}_{k=0}^{K-1})_m \approx \frac{1}{K} \sum_{k=0}^{K-1} \mathbb{E}_{X_t^m \sim p_\theta(X_t^m|x_{t-1}^m)}[w_{mk}(X_t^m)\, C(X_t^m)] \approx \frac{1}{K} \sum_{k=0}^{K-1} w_{mk}(x_t^m)\, C(x_t^m). \tag{5}$$

Transition kernel $p_m(\cdot)$ with respective importance weight $w_{mk}(\cdot)$ is used in Eq. (5) compared to transition kernel $p_k(\cdot)$ in Eq. (3). Taking the average over all reference trajectories, i.e. for all

$m \in \{0, \dots, K-1\}$ and summing over all time steps t yields

$$\mathcal{L}_{\text{Multi}}(\{X^k_{0:T}\}_{k=0}^{K-1}) \approx \sum_{t=1}^{T} \frac{1}{K^2} \sum_{m,k=0}^{K-1} w_{mk}(x_t^m)\, C(x_t^m). \tag{6}$$

A natural choice for the set of trajectories $\{x^k_{0:T}\}_{k=0}^{K-1}$ is the set of trajectories that are used anyway in the training of diffusion samplers i.e., in each training step the respective batch of trajectories. Therefore, ITIS does not introduce additional forward passes or sampling. We introduce a *chunk size* hyperparameter that controls the partitioning of each batch into disjoint subsets, or chunks. Trajectories are mixed only within each chunk. Varying the chunk size allows interpolation between the Single and Multi loss formulations to suit the requirements of the underlying problem.

4.2 ITIS Mechanisms of Mode Coverage

The ITIS multi-sample objective (6) applies importance weighting at every diffusion step t. To analyse how the associated coupling among trajectories mitigates mode collapse, it suffices to study a single fixed step t since the total loss is a sum over time steps. We therefore fix an arbitrary step index t, write $t-1$ and t for the preceding and current time point at that step, and suppress the subscripts associated to t whenever this does not compromise clarity. The per-step multi-sample estimator (6) then reads

$$\mathcal{L}_{\text{Multi}}\big(\{X_t^k\}_{k=0}^{K-1}\big) \;=\; \frac{1}{K^2} \sum_{m,k=0}^{K-1} w_{mk}(X_t^m)\, C(X_t^m). \tag{7}$$

Proposition 1 (Mode-coverage properties of $\mathcal{L}_{\text{Multi}}$)**.** *We fix $K \geq 2$. The analysis is carried out for a single fixed diffusion step as described above. Define the per-step single-sample estimator as*

$$\mathcal{L}_{\text{Single}} \;:=\; \frac{1}{K} \sum_{k=0}^{K-1} C(X_t^k). \tag{8}$$

1. ***Unbiased gradient.*** *The two estimators agree in expectation under the joint K-trajectory path measure $p_\theta^{\otimes K}$ and*

$$\mathbb{E}_{p_\theta^{\otimes K}}\big[\nabla_\theta \mathcal{L}_{\text{Multi}}\big] \;=\; \mathbb{E}_{p_\theta^{\otimes K}}\big[\nabla_\theta \mathcal{L}_{\text{Single}}\big]. \tag{9}$$

   In particular, both share the same expected parameter update, and hence the same set of critical points.

2. ***Loss ratio on realized batches.*** *Condition on the preceding states $\{x^k_{t-1}\}_{k=0}^{K-1}$, which determine the kernels p_k, and let the step-t states $X_t^m \sim p_m$. With $\mathcal{L}_{\text{Single}} := K^{-1}\sum_{m=0}^{K-1} C(X_t^m) > 0$ and $r := \mathcal{L}_{\text{Multi}}/\mathcal{L}_{\text{Single}}$, and since $w_{mm} = 1$,*

$$\begin{aligned} &r = 1 && \text{if } p_k = p_m \;\forall\, k, m, \\ &r \to K^{-1} && \text{in probability as the overlap of } p_k \text{ and } p_m \text{ vanishes } \forall\, k \neq m. \end{aligned} \tag{10}$$

Derivations are provided in Sec. A. Given the preceding states, p_m and p_k are Gaussians of equal width, so $w_{mk} < 1$ exactly when X_t^m lands closer to the mean of p_m than to that of p_k. This is the typical outcome, since X_t^m is drawn from p_m, and it becomes increasingly likely as the kernels separate. The weights still average to one, $\mathbb{E}_{p_m}[w_{mk}] = \int p_k = 1$, which is the same change-of-measure identity that yields the unbiasedness in part (1), because the rare samples that land closer to the mean of p_k receive correspondingly large weights. Consequently typical batches have $r < 1$, approaching K^{-1} as the trajectories separate. Between a low diversity batch ($r = 1$) and a diverse batch ($r < 1$) of equal $\mathcal{L}_{\text{Single}}$ this yields the per-step loss gap

$$\Delta\mathcal{L}_{\text{Multi}} := \mathcal{L}^{\text{low}}_{\text{Multi}} - \mathcal{L}^{\text{div}}_{\text{Multi}} = (1-r)\,\mathcal{L}_{\text{Single}} \;\geq\; 0. \tag{11}$$

The sign condition $\mathcal{L}_{\text{Single}} \geq 0$ is natural, since the objective is the path rKL and hence nonnegative in expectation, and the running costs are analytic conditional rKL terms, e.g. the norm term in (20), and hence nonnegative for every batch. The terminal cost may take either sign, but the diffusion

coefficient schedules of the samplers considered make the terminal kernels narrow compared to the spread of the trajectories, so that $r \approx K^{-1}$ at the terminal step for every configuration and hence the terminal term does not discriminate between configurations.
In the following we provide intuition for how the loss gap (11) steers the parameters towards mode covering regions of parameter space and stabilizes them there.
Differentiating $\mathcal{L}_{\text{Multi}} = r\,\mathcal{L}_{\text{Single}}$ w.r.t. the model parameters yields

$$\nabla_\theta \mathcal{L}_{\text{Multi}} = r\,\nabla_\theta \mathcal{L}_{\text{Single}} + \mathcal{L}_{\text{Single}}\,\nabla_\theta r, \tag{12}$$

whose two terms play complementary roles.

Reaching mode coverage. The second term has no counterpart in the single sample gradient. Since r decreases with the separation of the trajectories and $\mathcal{L}_{\text{Single}} \geq 0$, its descent direction pushes the transition kernels of the trajectories apart. Since the rKL keeps the trajectories in regions of high target density, this favors configurations in which they occupy distinct modes. The term acts at partial overlap, which prevents the parameters from entering mode-collapsed regions from there, and vanishes in both edge cases of (10). By Proposition 1 (1), this push is balanced in expectation by rare batches with large weights, so its effect on training relies on typical batches dominating the updates, as they do under bounded-step optimizers such as Adam (Kingma & Ba, 2014) or under weight clipping.

Retaining mode coverage. The first term is the conventional gradient scaled by r. Where the trajectories of a batch are separated, a gradient step is therefore reduced in magnitude by up to a factor of K, including along the mode seeking directions of the rKL. Collapsed batches are updated at the unchanged rate. Mode-covering configurations are thus left more slowly than collapsed ones, an effect that no global learning rate schedule reproduces, since such a schedule cannot depend on the configuration of a batch. This holds for plain gradient descent. Adaptive optimizers such as Adam (Kingma & Ba, 2014) normalize the overall gradient scale, but there the effect persists as a relative downweighting of separated batches within an update.

5 Experimental Setup

We evaluate our ITIS multi-sample training objective on diffusion samplers and benchmarks that are well established in the field of diffusion samplers (Blessing et al., 2024; Chen et al., 2025; Guo et al., 2026) as well as on dedicated benchmarks designed to investigate the impact of ITIS on mode coverage. This enables a comprehensive comparison between our multi-sample objective and the conventional single-sample loss.
Mode discovery and coverage. A prior matched to the extent of the target gives a sampler a head start that can mask differences in its ability to discover modes, which is why Midgley et al. (2023) test mode discovery on the 2D GMM40 with a deliberately narrow initialization. To isolate the mechanisms of Sec. 4.2, we therefore use GMM40, whose 40 unit variance components have centres drawn uniformly from $[-40, 40]^d$, with a reduced prior scale, lowering the scale of the DDS prior from 40 in the standard benchmark to 25 for $d = 50$ and to 10 for $d = 2$. We vary the chunk size K at fixed batch size, so that $K = 1$ recovers the single sample loss exactly, tune learning rate and noise scale separately for each K with equal budget, and report the Sinkhorn distance and the entropic mode coverage (EMC) of Blessing et al. (2024), which equals 1 for a uniform allocation of samples across modes, over 5 seeds. Although lower dimensional, the 2D variant is the harder discovery problem. In 50D the centres concentrate near a sphere around the origin, so the radial path from the prior towards one mode rarely passes near another mode. In 2D, modes frequently lie between the prior and more distant modes and capture trajectories on their way, hindering the discovery of the distant ones and triggering mode collapse.
Standard benchmarks. The standard targets MW54, Robot1, Robot4, Funnel, GMM40 and MoS follow Blessing et al. (2024) and Chen et al. (2025) and are described in Sec. C.6.1.
Particle systems. We consider classical continuous particle potentials where the energy $E(x)$ depends on pairwise distances in an n-particle system. For this we follow (Akhound-Sadegh et al., 2024; Guo et al., 2026) and use well established targets DW-4 (2D double-well potential, 4 particles) and Lennard Jones potentials LJ-13 (3D, 13 particles) and LJ-55 (3D, 55 particles). We use the equivariant 2-Wasserstein distance between the ground-truth and the generated samples, with the ground-truth samples, generated via MCMC samples following Klein et al. (2023). The mathematical description of these particle potentials is detailed in Sec. C.6.2.

Methods. Our empirical framework centers on three purely neural diffusion samplers: DDS (Vargas et al., 2023), DIS (Berner et al., 2024), and PIS (Zhang & Chen, 2022), detailed in Sec. 2. All three learn a controlled SDE whose terminal marginal approximates the target, and all three are trained by minimizing the rKL between path measures (Zhang & Chen, 2022; Vargas et al., 2023; Berner et al., 2024). The per-step costs are the rKL between the Gaussian transition kernels of the respective discretization and are available in closed form. The samplers differ in the reference process and its parametrization. PIS controls a Brownian motion started at a point mass and learns the drift directly, whereas DDS and DIS start from a Gaussian prior and approximate the time reversal of a variance-preserving noising process with a score network, differing in their parametrization and discretization. We augment each base sampler with our proposed ITIS-based multi-sample objective on the aforementioned standard benchmarks. We report numbers from Chen et al. (2025) for two classical baselines: standard sequential Monte Carlo (SMC) (Del Moral et al., 2006), which transports particles through annealed intermediate distributions via importance resampling and MCMC rejuvenation; and SMC-ESS (Chen et al., 2025; Del Moral et al., 2012), an adaptive variant that selects the annealing schedule via the constant-ESS criterion. We further consider two hybrid methods that retain the SMC skeleton while substituting learned proposals: CRAFT (Matthews et al., 2022), which trains a normalizing flow between annealing temperatures within the SMC framework, and SMC-FC (Buchholz et al., 2021). Finally, we include Schrödinger bridge methods grounded in entropy-regularized optimal transport (OT), specifically Controlled Monte Carlo Diffusions (CMCD) (Vargas et al., 2024) and its variants CMCD-KL and CMCD-LV, which minimize a KL divergence and log-variance bound on the path measure respectively. Our evaluation follows the established protocol of Blessing et al. (2024).

On the particle-based benchmarks we compare DDS with ITIS to the following baselines. These include Proximal Diffusion Neural Samplers (PDNS) (Guo et al., 2026), which frame stochastic optimal control (SOC) via proximal point methods on path measure spaces, and Adjoint Sampling (AS) (Havens et al., 2025), another SOC approach. We also compare against its generalization, Adjoint Schrödinger Bridge Samplers (ASBS) (Liu et al., 2025b), which relaxes the restrictive memoryless condition of the design space. We also compare against the sequential controlled Langevin diffusions (SCLD) (Chen et al., 2025) which is a methodology that combines Monte Carlo (SMC) with diffusion samplers, and analyzes them through continuous-time path-space measures. Finally, we compare to iDEM (Akhound-Sadegh et al., 2024), the first learned sampler to successfully scale to the LJ-55 regime, and Particle Denoising Diffusion Samplers (PDDS) (Phillips et al., 2024a).

We employ standard metrics as defined in App. C.5 and App. C.5.1. In our evaluation of standard benchmarks, we prioritize Sinkhorn distance and Maximum Mean Discrepancy (MMD) as the central mode-coverage sensitive probability metrics over the Evidence Lower Bound (ELBO). While the ELBO is a standard training objective, it is well-documented to be insensitive to mode collapse and can remain deceptively high even when the model fails to capture the full support of the target distribution. In contrast, Sinkhorn and MMD provide more stringent measures of distributional alignment (Blessing et al., 2024).

6 Main Results

Mode discovery and coverage. Figure 1 gives the most direct evidence for ITIS. Because chunk size $K = 1$ is the single sample loss at identical batch size and tuning budget, the variation with K isolates the effect of the inter-trajectory coupling. In 50D every $K \geq 5$ improves on the single sample loss, reducing the Sinkhorn distance from 4417.8 at $K = 1$ to 2930.6 at $K = 100$, while EMC is close to one throughout, consistent with the easier discovery in high dimension, and still increases with K. In the harder 2D problem the effect is considerably larger. Every $K \geq 2$ improves on $K = 1$ in both metrics, with the Sinkhorn distance decreasing from 249.0 to 114.6 and EMC rising from 0.838 to 0.931 at $K = 400$. A study on the runtime and memory costs of ITIS is presented Sec. C.1 and experimental details and the hyperparameter optimization protocol is provided in Sec. C.3.

Particle systems. Tab. 1 compares DDS (Multi) with the baseline methods using particle-based Wasserstein-2 distances in sample space and energy space. DDS (Multi) achieves the best overall sample-space distance on the DW-4 and LJ-13 potentials. On LJ-55 system, DDS (Multi) obtains a competitive result of 4.31, demonstrating the ITIS-related improvement over DDS (Single) at 4.60. Notably, these results are achieved by significantly fewer diffusion steps ($T = 80, 70, 80$) for (DW-4, LJ-13, LJ-55) compared to 1000 for e.g. PDNS, resulting in only a fraction of the Number of Func-

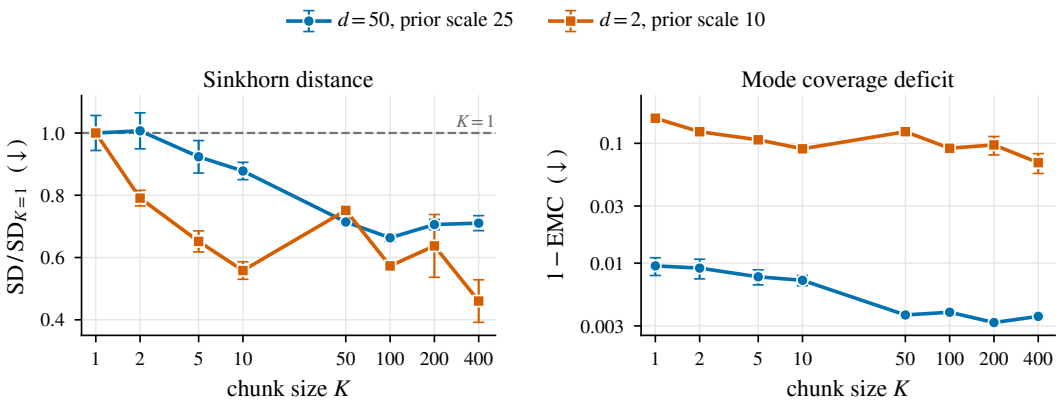


Figure 1: Mode discovery on GMM40 with reduced prior scale (DDS) as a function of the chunk size K, where $K = 1$ is the single sample loss. The left panel shows the Sinkhorn distance (SD) relative to its value at $K = 1$ (dashed line), since its absolute scale depends on the dimension. The right panel shows the coverage deficit $1 - \mathrm{EMC}$ of the entropic mode coverage on a logarithmic axis. Mean $\pm$ standard error over 5 seeds.

tion Evaluations (NFEs). Compared directly with DDS (Single), DDS (Multi) reduces all distances across all three systems. Thus, ITIS yields consistent improvements, while achieving competitive results overall.

Table 1: Evaluation of particle-based distances. Following Liu et al. (2025b); Guo et al. (2026) we report particle-based Wasserstein-2 distances w.r.t. samples $\mathcal{W}_2$, and energy values $E(\cdot)\,\mathcal{W}_2$, best results highlighted in **bold**. Baseline results taken from Guo et al. (2026), except partially for AS and ASBS from the corrected version Liu et al. (2025a) of Liu et al. (2025b).

	DW-4 ($d = 8$)		LJ-13 ($d = 39$)		LJ-55 ($d = 165$)	
Method	$\mathcal{W}_2 \downarrow$	$E(\cdot)\,\mathcal{W}_2 \downarrow$	$\mathcal{W}_2 \downarrow$	$E(\cdot)\,\mathcal{W}_2 \downarrow$	$\mathcal{W}_2 \downarrow$	$E(\cdot)\,\mathcal{W}_2 \downarrow$
PDDS	0.92 ± 0.08	0.58 ± 0.25	4.66 ± 0.87	56.01 ± 10.80	—	—
SCLD	1.30 ± 0.64	0.40 ± 0.19	2.93 ± 0.19	27.98 ± 1.26	—	—
PIS	0.68 ± 0.28	0.65 ± 0.25	1.93 ± 0.07	18.02 ± 1.12	4.79 ± 0.45	228.70 ± 131.27
DDS (Single)	0.92 ± 0.11	0.90 ± 0.37	1.99 ± 0.13	24.61 ± 8.99	4.60 ± 0.09	173.09 ± 18.01
LV-PIS	1.04 ± 0.29	1.89 ± 0.89	—	—	—	—
iDEM	0.70 ± 0.06	0.55 ± 0.14	1.61 ± 0.01	30.78 ± 24.46	4.69 ± 1.52	93.53 ± 16.31
AS	0.62 ± 0.06	0.55 ± 0.12	1.67 ± 0.01	2.40 ± 1.25	4.04 ± 0.05	30.83 ± 8.19
ASBS	0.43 ± 0.05	$\mathbf{0.20} \pm 0.11$	1.59 ± 0.00	1.99 ± 1.01	4.00 ± 0.03	28.10 ± 8.15
PDNS	0.51 ± 0.04	0.21 ± 0.03	1.57 ± 0.01	$\mathbf{1.01} \pm 0.18$	$\mathbf{3.95} \pm 0.01$	$\mathbf{21.97} \pm 3.14$
DDS (Multi)	$\mathbf{0.39} \pm 0.06$	0.25 ± 0.10	$\mathbf{1.55} \pm 0.02$	15.12 ± 3.19	4.31 ± 0.06	35.61 ± 12.45

Standard benchmarks. We benchmark our method against the established results reported in Blessing et al. (2024) for the mixture model and funnel tasks. For the high-dimensional many-well (MW54) and robotic tasks (Robot 1 and Robot 4), we adopt the performance baselines established by Chen et al. (2025). Following the protocol in Blessing et al. (2024), standard single-sample objectives are typically trained and evaluated using $T = 128$ diffusion steps to achieve competitive Probability Metric (PM) scores that quantify mode coverage. In Tab. 2 and 8, we report our findings, highlighting that our objective not only matches or exceeds these baseline PM values but does so with a substantial reduction in both the number of diffusion steps and total NFEs.
5d Many-Well benchmark. For the MW54 experiments, DDS with the multi-sample objective ($T = 35$) yields better Sinkhorn distances compared to the single-sample objectives and comparable to the strong baselines such as PIS (single) and the CMCD-LV, as detailed in Tab. 2. **10d Robot1 benchmark.** For the robot task with 1 goal, DDS with multi-sample objective ($T = 25$) clearly outperforms its single-sample counterpart in terms of the Sinkhorn values – see Tab. 2, signifying much better mode coverage. **10d Robot4 benchmark.** For the robot task with 4 goals, DDS with multi-sample objective ($T = 10$) clearly yields the best Sinkhorn values (see Tab. 2), outperforming

Table 2: **Sinkhorn Distance (↓) Comparison across Varied Target Densities.** Performance of our ITIS-based multi-sample objective (*Multi*) relative to standard single-sample baselines (*Single*) and established reference methods. Best results per target are highlighted in **bold**. Dashed lines in DIS indicate no baselines were presented in Chen et al. (2025) for direct comparison against our method. We refer the reader to Appendix D regarding the value marked with asterisk.

Method	MW54 (5d)	Robot1 (10d)	Robot4 (10d)	Funnel (10d)	GMM40 (50d)	MoS (50d)
DDS (Single)	$0.63_{\pm 0.24}$	$11.44_{\pm 1.25}$	$5.38_{\pm 2.44}$	$142.89_{\pm 9.55}$	$5551.1_{\pm 116.4}$	$2154.88_{\pm 3.86}$
DDS (Multi)	$0.53_{\pm 0.15}$	$4.32_{\pm 0.75}$	$\mathbf{1.05_{\pm 0.57}}$	$130.01_{\pm 1.30}$	$\mathbf{2441.5_{\pm 12.51}}$	$\mathbf{960.23_{\pm 47.75}}$
DIS (Single)	—	—	—	$118.95_{\pm 12.8}$	$3044.7_{\pm 464.7}$	$2200.59_{\pm 18.73}$
DIS (Multi)	—	—	—	$\mathbf{113.52_{\pm 9.71}}$	$2618.31_{\pm 98.76}$	$1828.19_{\pm 23.34}$
PIS (Single)	$0.52^{*}_{\pm 0.008}$	$1.54_{\pm 0.72}$	$2.02_{\pm 0.36}$	$171.33_{\pm 123}$	$10405.75_{\pm 69.41}$	$2113.17_{\pm 31.17}$
PIS (Multi)	$0.58_{\pm 0.015}$	$1.76_{\pm 0.05}$	$1.70_{\pm 0.03}$	$143.97_{\pm 0.25}$	$6951.19_{\pm 114.10}$	$1562.02_{\pm 9.27}$
SMC	$20.71_{\pm 5.33}$	$24.02_{\pm 1.06}$	$24.08_{\pm 0.26}$	$149.35_{\pm 4.73}$	$46370.34_{\pm 137.79}$	$3297.28_{\pm 2184.54}$
SMC-ESS	$1.11_{\pm 0.15}$	$1.82_{\pm 0.50}$	$2.11_{\pm 0.31}$	$117.48_{\pm 9.70}$	$24240.68_{\pm 50.52}$	$1477.04_{\pm 133.80}$
SMC-FC	$2.03_{\pm 0.17}$	$\mathbf{0.37_{\pm 0.08}}$	$1.23_{\pm 0.02}$	$211.43_{\pm 30.08}$	$39018.27_{\pm 159.32}$	$3200.10_{\pm 95.35}$
CRAFT	$11.47_{\pm 0.90}$	$2.92_{\pm 0.01}$	$4.14_{\pm 0.50}$	$133.42_{\pm 1.04}$	$28960.70_{\pm 354.89}$	$1918.14_{\pm 108.22}$
CMCD-KL	$0.57_{\pm 0.05}$	$3.71_{\pm 1.00}$	$2.62_{\pm 0.41}$	$124.89_{\pm 8.95}$	$22132.28_{\pm 595.18}$	$1848.89_{\pm 532.56}$
CMCD-LV	$\mathbf{0.51_{\pm 0.08}}$	$28.49_{\pm 0.07}$	$27.00_{\pm 0.07}$	$139.07_{\pm 9.35}$	$4258.57_{\pm 737.15}$	$1945.71_{\pm 48.79}$

every other baseline (especially the neural samplers) for this target and signifying better mode coverage, comparable with hybrid approaches such as SMC-ESS. **10d Funnel benchmark.** The results in Tab. 2 and 8 demonstrate that the multi-sample objectives consistently improves distributional alignment over single-sample baselines. In this setting, multi-sample objective achieves the strongest performance when integrated with the DIS sampler for both the PMs (Tab. 2: column 4, Tab. 8, column 1) outperforming every other method. Also DDS and PIS exhibit significant gains over their respective single-sample counterparts. Notably, these performance improvements are achieved using only $T = 25$ diffusion steps. **Mixture model benchmarks ($d = 50$).** The multi-sample objective demonstrates significant performance gains, as detailed in Tab. 2 and 8. Notably, the ITIS-augmented DDS sampler outperforms all baseline methods, achieving a state-of-the-art Sinkhorn distance of ~ 2442. The PIS and DIS both exhibit significant improvements over their respective single-sample counterparts. In the 10-mode Mixture of Students (MoS) task, DDS equipped with the multi-sample objective achieves a Sinkhorn distance of ~ 960, outperforming controlled versions such as CMCD-LV (~ 1945).
Crucially, these results are achieved using only $T = 15$ diffusion steps, resulting in a $\sim 10\times$ reduction in discretization to reach convergence.

7 Limitations, Conclusion & Future Work

Limitations and future work. Computing all pairwise importance weights costs $\mathcal{O}(K^2)$ per chunk. The best chunk size is problem dependent, ranging from $K = 5$ to $K = 125$ in our experiments, so this cost is not restrictive here, but it would dominate for substantially larger K. The weights decay exponentially with the distance between transition means, so even when neighboring trajectories overlap, many pairs in a chunk carry a small weight and, with it, little gradient contribution. Restricting the computation to nearby pairs, for example, via cell lists (Frenkel & Smit, 2002), would therefore leave the loss and its gradient essentially unchanged while recovering near-$\mathcal{O}(K)$ cost. We leave a systematic study of such approximations to future work.
Conclusion. We introduced Inter-Trajectory Importance Sampling (ITIS), a multi-sample training objective for diffusion samplers that couples each trajectory's loss with other trajectories in its batch. The estimator remains unbiased, preserving the expected objective, its gradients and critical points, while assigning a lower realized loss to more diverse batch configurations. This provides a principled way to mitigate mode collapse without adding heuristic repulsive forces or coupling samples at inference time. Across mode-discovery tests, standard sampling benchmarks, and interacting-particle systems, ITIS improves distributional fidelity relative to methods using the corresponding single-sample objective, even with fewer NFEs.

AI USE STATEMENT

In this work, we used generative AI tools solely to polish the writing, improving grammar, clarity, and readability. We did not use generative AI tools for any other purpose. We have reviewed all AI-assisted edits to ensure that they preserve the intended meaning and accurately convey our work. We take responsibility for the final content of this work, including all text, claims, and artifacts produced with the aid of generative AI.

ETHICS STATEMENT

This work introduces Inter-Trajectory Importance Sampling (ITIS), a training objective for diffusion samplers that mitigates mode collapse while improving sampling efficiency. The method may benefit applications in probabilistic inference, scientific computing, and generative modeling by improving sample diversity and reducing the number of diffusion steps and energy evaluations required. As a general improvement to generative sampling methods, ITIS could also indirectly strengthen systems capable of generating realistic synthetic content. Such systems may be misused for deception, impersonation, misinformation, or harassment. In addition, the method may inherit or amplify biases present in the underlying datasets or energy functions used in downstream applications. Care should therefore be taken when deploying models trained with ITIS in socially sensitive domains such as healthcare, finance, or education. The proposed method does not directly target surveillance, cybersecurity, or weapon-related applications. Moreover, the reduced computational requirements observed in our experiments may help lower the energy cost of diffusion-based sampling, although large-scale generative modeling still carries environmental costs. To mitigate potential harms, we encourage responsible release practices, including transparent documentation of datasets, model assumptions, limitations, and intended use cases, as well as the release of reproducible code and evaluation protocols. Practitioners should also ensure compliance with relevant legal, privacy, and fairness standards when applying these methods in real-world settings.

REPRODUCIBILITY STATEMENT

The method and its application to Denoising Diffusion Samplers are described in Sections 4.1 and Appendix B, while Appendix A gives the proof of Proposition 1 and states the conditions used in the argument. Experimental settings and reproduction guidance are collected in Appendix C: Sections C.3–C.4 describe the tuning and evaluation protocols, Sections C.6–C.7 specify benchmark targets and method configurations, and Section C.1 reports runtime and memory measurements. Together, these sections document the algorithm, theoretical claims, benchmark setup, and evaluation procedures used to obtain the results. The paper provides an anonymized implementation at `https://anonymous.4open.science/r/ITIS-04D4/README.md`.

REFERENCES

Tara Akhound-Sadegh, Jarrid Rector-Brooks, Avishek Joey Bose, Sarthak Mittal, Pablo Lemos, Cheng-Hao Liu, Marcin Sendera, Siamak Ravanbakhsh, Gauthier Gidel, Yoshua Bengio, Nikolay Malkin, and Alexander Tong. Iterated denoising energy matching for sampling from Boltzmann densities. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, 2024. URL `https://proceedings.mlr.press/v235/akhound-sadegh24a.html`.

Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pp. 2623–2631, 2019.

Michael S. Albergo and Eric Vanden-Eijnden. NETS: A non-equilibrium transport sampler. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 1026–1055, 2025. URL `https://proceedings.mlr.press/v267/albergo25a.html`.

Oleg Arenz, Gerhard Neumann, and Mingjun Zhong. Efficient gradient-free variational inference using policy search. In Jennifer Dy and Andreas Krause (eds.), *Proceedings of the 35th Inter-*

national Conference on Machine Learning, volume 80 of *Proceedings of Machine Learning Research*, pp. 234–243. PMLR, 10–15 Jul 2018. URL https://proceedings.mlr.press/v80/arenz18a.html.

Oleg Arenz, Mingjun Zhong, and Gerhard Neumann. Trust-region variational inference with Gaussian mixture models. *Journal of Machine Learning Research*, 21(163):1–60, 2020. URL https://jmlr.org/papers/v21/19-524.html.

Nachman Aronszajn. Theory of reproducing kernels. *Transactions of the American Mathematical Society*, 68(3):337–404, 1950. doi: 10.1090/S0002-9947-1950-0051437-7. URL https://www.ams.org/journals/tran/1950-068-03/S0002-9947-1950-0051437-7/.

Julius Berner, Lorenz Richter, and Karen Ullrich. An optimal control perspective on diffusion-based generative modeling. *Transactions on Machine Learning Research*, 2024. URL https://openreview.net/forum?id=oYIjw37pTP.

Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer, New York, 2006.

David M. Blei, Alp Kucukelbir, and Jon D. McAuliffe. Variational inference: A review for statisticians. *Journal of the American Statistical Association*, 112(518):859–877, 2017. doi: 10.1080/01621459.2017.1285773.

Denis Blessing, Xiaogang Jia, Johannes Esslinger, Francisco Vargas, and Gerhard Neumann. Beyond ELBOs: A large-scale evaluation of variational methods for sampling. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 4205–4229, 2024. URL https://proceedings.mlr.press/v235/blessing24a.html.

Denis Blessing, Julius Berner, Lorenz Richter, Carles Domingo-Enrich, Yuanqi Du, Arash Vahdat, and Gerhard Neumann. Trust region constrained measure transport in path space for stochastic optimal control and inference, 2025a. URL https://arxiv.org/abs/2508.12511. arXiv preprint arXiv:2508.12511.

Denis Blessing, Julius Berner, Lorenz Richter, and Gerhard Neumann. Underdamped diffusion bridges with applications to sampling. In *International Conference on Learning Representations*, 2025b. URL https://publikationen.bibliothek.kit.edu/1000179040.

Denis Blessing, Lorenz Richter, Julius Berner, Egor Malitskiy, and Gerhard Neumann. Bridge matching sampler: Scalable sampling via generalized fixed-point diffusion matching, 2026. URL https://arxiv.org/abs/2603.00530. arXiv preprint arXiv:2603.00530.

Alexander Buchholz, Nicolas Chopin, and Pierre E. Jacob. Adaptive tuning of Hamiltonian Monte Carlo within sequential Monte Carlo. *Bayesian Analysis*, 16(3):745–771, 2021. doi: 10.1214/20-BA1222. URL https://doi.org/10.1214/20-BA1222.

Junhua Chen, Lorenz Richter, Julius Berner, Denis Blessing, Gerhard Neumann, and Anima Anandkumar. Sequential controlled Langevin diffusions. In *International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=dImD2sgy86.

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

Gabriele Corso, Yilun Xu, Valentin de Bortoli, Regina Barzilay, and Tommi Jaakkola. Particle guidance: non-I.I.D. diverse sampling with diffusion models. In *International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=KqbCvIFBY7.

Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems*, volume 26, 2013. URL https://papers.nips.cc/paper/2013/hash/af21d0c97db2e27e13572cbf59eb343d-Abstract.html.

Marco Cuturi, Laetitia Meng-Papaxanthos, Yingtao Tian, Charlotte Bunne, Geoff Davis, and Olivier Teboul. Optimal transport tools (OTT): A JAX toolbox for all things Wasserstein, 2022. URL https://arxiv.org/abs/2201.12324. arXiv preprint arXiv:2201.12324.

Pierre Del Moral, Arnaud Doucet, and Ajay Jasra. Sequential Monte Carlo samplers. *Journal of the Royal Statistical Society: Series B*, 68(3):411–440, 2006.

Pierre Del Moral, Arnaud Doucet, and Ajay Jasra. An adaptive sequential Monte Carlo method for approximate Bayesian computation. *Statistics and Computing*, 22(5):1009–1020, September 2012. ISSN 1573-1375. doi: 10.1007/s11222-011-9271-y. URL https://doi.org/10.1007/s11222-011-9271-y.

Carles Domingo-Enrich, Michal Drozdzal, Brian Karrer, and Ricky T. Q. Chen. Adjoint matching: Fine-tuning flow and diffusion generative models with memoryless stochastic optimal control. In *International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=xQBRrtQM8u.

Carles Domingo-Enrich, Yuanqi Du, and Michael S. Albergo. A unified perspective on fine-tuning and sampling with diffusion and flow models, 2026. URL https://arxiv.org/abs/2605.00229. arXiv preprint arXiv:2605.00229.

Arnaud Doucet, Will Grathwohl, Alexander G. D. G. Matthews, and Heiko Strathmann. Score-based diffusion meets annealed importance sampling. In *Advances in Neural Information Processing Systems*, volume 35, 2022. URL https://proceedings.neurips.cc/paper_files/paper/2022/hash/86b7128efa3950df7c0f6c0342e6dcc1-Abstract.html.

Michael F. Faulkner and Samuel Livingstone. Sampling algorithms in statistical physics: a guide for statistics and machine learning. *Statistical Science*, 39(1), February 2024. ISSN 0883-4237. doi: 10.1214/23-STS893. URL https://arxiv.org/abs/2208.04751.

Daan Frenkel and Berend Smit. *Understanding Molecular Simulation: From Algorithms to Applications*. Academic Press, San Diego, 2nd edition, 2002. ISBN 978-0-12-267351-1.

Alfredo Garbuno-Inigo, Franca Hoffmann, Wuchen Li, and Andrew M. Stuart. Interacting Langevin diffusions: Gradient structure and ensemble Kalman sampler. *SIAM Journal on Applied Dynamical Systems*, 19(1):412–441, 2020. doi: 10.1137/19M1251655. URL https://arxiv.org/abs/1903.08866.

Arthur Gretton, Karsten M. Borgwardt, Malte J. Rasch, Bernhard Schölkopf, and Alexander Smola. A kernel two-sample test. *Journal of Machine Learning Research*, 13(25):723–773, 2012. URL https://jmlr.org/papers/v13/gretton12a.html.

Wei Guo, Jaemoo Choi, Yuchen Zhu, Molei Tao, and Yongxin Chen. Proximal diffusion neural sampler. In *International Conference on Learning Representations*, 2026. URL https://openreview.net/pdf?id=XTHQqS7ObC.

W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 04 1970. ISSN 0006-3444. doi: 10.1093/biomet/57.1.97. URL https://doi.org/10.1093/biomet/57.1.97.

Aaron J. Havens, Benjamin Kurt Miller, Bing Yan, Carles Domingo-Enrich, Anuroop Sriram, Daniel S. Levine, Brandon M. Wood, Bin Hu, Brandon Amos, Brian Karrer, Xiang Fu, Guan-Horng Liu, and Ricky T. Q. Chen. Adjoint sampling: Highly scalable diffusion samplers via adjoint matching. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 22204–22237. PMLR, 2025. URL https://proceedings.mlr.press/v267/havens25a.html.

Jiajun He, Wenlin Chen, Mingtian Zhang, David Barber, and José Miguel Hernández-Lobato. Training neural samplers with reverse diffusive KL divergence. In *Proceedings of The 28th International Conference on Artificial Intelligence and Statistics*, volume 258 of *Proceedings of Machine Learning Research*, pp. 5167–5175, 2025a. URL https://proceedings.mlr.press/v258/he25a.html.

Jiajun He, Yuanqi Du, Francisco Vargas, Dinghuai Zhang, Shreyas Padhy, RuiKang OuYang, Carla Gomes, and José Miguel Hernández-Lobato. No trick, no treat: Pursuits and challenges towards simulation-free training of neural samplers, 2025b. URL https://arxiv.org/abs/2502.06685. arXiv preprint arXiv:2502.06685.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin (eds.), *Advances in Neural Information Processing Systems*, volume 33, pp. 6840–6851. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper/2020/file/4c5bcfec8584af0d967f1ab10179ca4b-Paper.pdf.

Peter Holderrieth, Michael S. Albergo, and Tommi Jaakkola. LEAPS: A discrete neural sampler via locally equivariant networks. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 23397–23416. PMLR, 2025. URL https://proceedings.mlr.press/v267/holderrieth25a.html.

Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Leon Klein, Andrew Y. K. Foong, Tor Erlend Fjelde, Bruno Mlodozeniec, Marc Brockschmidt, Sebastian Nowozin, Frank Noé, and Ryota Tomioka. Timewarp: Transferable acceleration of molecular dynamics by learning time-coarsened dynamics, 2023. URL https://arxiv.org/abs/2302.01170.

Jonas Köhler, Leon Klein, and Frank Noe. Equivariant flows: Exact likelihood generative learning for symmetric densities. In Hal Daumé III and Aarti Singh (eds.), *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 5361–5370. PMLR, 13–18 Jul 2020. URL https://proceedings.mlr.press/v119/kohler20a.html.

Jonas Köhler, Leon Klein, and Frank Noé. Equivariant flows: Exact likelihood generative learning for symmetric densities, 2020. URL https://arxiv.org/abs/2006.02425.

Sergey Levine. Reinforcement learning and control as probabilistic inference: Tutorial and review, 2018. URL https://arxiv.org/abs/1805.00909. arXiv preprint arXiv:1805.00909.

Guan-Horng Liu, Jaemoo Choi, Yongxin Chen, Benjamin Kurt Miller, and Ricky T. Q. Chen. Adjoint schrödinger bridge sampler, 2025a. URL https://arxiv.org/abs/2506.22565.

Guan-Horng Liu, Jaemoo Choi, Yongxin Chen, Benjamin Kurt Miller, and Ricky T. Q. Chen. Adjoint Schrödinger bridge sampler. In *Advances in Neural Information Processing Systems*, volume 38, 2025b. URL https://proceedings.neurips.cc/paper_files/paper/2025/hash/174692c52dc84fad2b2e99dd8637ce6a-Abstract-Conference.html.

Qiang Liu and Dilin Wang. Stein variational gradient descent: A general purpose Bayesian inference algorithm. In *Advances in Neural Information Processing Systems*, volume 29, 2016. URL https://proceedings.neurips.cc/paper/2016/hash/b3ba8f1bee1238a2f37603d90b58898d-Abstract.html.

Ilya Loshchilov and Frank Hutter. SGDR: Stochastic gradient descent with warm restarts. In *International Conference on Learning Representations*, 2017. URL https://openreview.net/forum?id=Skq89Scxx.

Alexander G. D. G. Matthews, Michael Arbel, Danilo J. Rezende, and Arnaud Doucet. Continual repeated annealed flow transport Monte Carlo. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 15196–15219. PMLR, 2022. URL https://proceedings.mlr.press/v162/matthews22a.html.

Laurence Illing Midgley, Vincent Stimper, Gregor N. C. Simm, Bernhard Schölkopf, and José Miguel Hernández-Lobato. Flow annealed importance sampling bootstrap. In *International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=XCTVFJwS9LJ.

Tom Minka. Divergence measures and message passing. Technical Report MSR-TR-2005-173, Microsoft Research, 2005. URL https://www.microsoft.com/en-us/research/publication/divergence-measures-and-message-passing/.

Radford M. Neal. Annealed importance sampling. *Statistics and Computing*, 11:125–139, 2001. URL https://arxiv.org/abs/physics/9803008.

Radford M. Neal. Slice sampling. *The Annals of Statistics*, 31(3):705–767, 2003. URL https://arxiv.org/abs/physics/0009028.

Frank Noé, Simon Olsson, Jonas Köhler, and Hao Wu. Boltzmann generators: Sampling equilibrium states of many-body systems with deep learning. *Science*, 365(6457):eaaw1147, 2019. doi: 10.1126/science.aaw1147.

Michele Pavon. Stochastic control and nonequilibrium thermodynamical systems. *Applied Mathematics and Optimization*, 19(1):187–202, January 1989. ISSN 1432-0606. doi: 10.1007/BF01448198. URL https://doi.org/10.1007/BF01448198.

Gabriel Peyré and Marco Cuturi. Computational optimal transport. *Foundations and Trends in Machine Learning*, 11(5–6):355–607, 2019. URL https://arxiv.org/abs/1803.00567.

Angus Phillips, Hai-Dang Dau, Michael John Hutchinson, Valentin De Bortoli, George Deligiannidis, and Arnaud Doucet. Particle denoising diffusion sampler, 2024a. URL https://arxiv.org/abs/2402.06320.

Angus Phillips, Hai-Dang Dau, Michael John Hutchinson, Valentin de Bortoli, George Deligiannidis, and Arnaud Doucet. Particle denoising diffusion sampler. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 40688–40724, 2024b. URL https://proceedings.mlr.press/v235/phillips24a.html.

Lorenz Richter and Julius Berner. Improved sampling via learned diffusions. In *International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=h4pNROsO06.

Sebastian Sanokowski, Wilhelm Berghammer, Haoyu Wang, Martin Ennemoser, Sepp Hochreiter, and Sebastian Lehner. Scalable discrete diffusion samplers: Combinatorial optimization and statistical physics. In *International Conference on Learning Representations*, 2025a. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/d87eadbe6114b768e76c5d5e8eb39388-Abstract-Conference.html.

Sebastian Sanokowski, Lukas Gruber, Christoph Bartmann, Sepp Hochreiter, and Sebastian Lehner. Rethinking losses for diffusion bridge samplers. In *Advances in Neural Information Processing Systems*, volume 38, 2025b. URL https://proceedings.neurips.cc/paper_files/paper/2025/hash/9c2bd9467f14c5377ce551df39838a97-Abstract-Conference.html.

Marcin Sendera, Minsu Kim, Sarthak Mittal, Pablo Lemos, Luca Scimeca, Jarrid Rector-Brooks, Alexandre Adam, Yoshua Bengio, and Nikolay Malkin. Improved off-policy training of diffusion samplers. In *Advances in Neural Information Processing Systems*, volume 37, 2024. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/93c936b9e492def9c00782cab79dbc6d-Abstract-Conference.html.

Jascha Sohl-Dickstein, Eric A. Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pp. 2256–2265. PMLR, 2015. URL https://proceedings.mlr.press/v37/sohl-dickstein15.html.

Yang Song, Conor Durkan, Iain Murray, and Stefano Ermon. Maximum likelihood training of score-based diffusion models. In *Advances in Neural Information Processing Systems*, volume 34, pp. 1415–1428, 2021a.

Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2021b. URL https://openreview.net/forum?id=PxTIG12RRHS.

Yifeng Tian, Nishant Panda, and Yen Ting Lin. Liouville flow importance sampler. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 48186–48210, 2024. URL https://proceedings.mlr.press/v235/tian24c.html.

Belinda Tzen and Maxim Raginsky. Theoretical guarantees for sampling and inference in generative models with latent diffusions, 2019. URL https://arxiv.org/abs/1903.01608. arXiv preprint arXiv:1903.01608.

Masatoshi Uehara, Yulai Zhao, Ehsan Hajiramezanali, Gabriele Scalia, Gökcen Eraslan, Avantika Lal, Sergey Levine, and Tommaso Biancalani. Bridging model-based optimization and generative modeling via conservative fine-tuning of diffusion models. In *Advances in Neural Information Processing Systems*, volume 37, 2024. URL https://papers.nips.cc/paper_files/paper/2024/hash/e68274fc4f158dbcbd4dddc672f7ee9c-Abstract-Conference.html.

Thomas Unterthiner, Bernhard Nessler, Calvin Seward, Günter Klambauer, Martin Heusel, Hubert Ramsauer, and Sepp Hochreiter. Coulomb GANs: Provably optimal Nash equilibria via potential fields. In *International Conference on Learning Representations*, 2018. URL https://openreview.net/pdf?id=SkVqXOxCb.

Francisco Vargas, Will Grathwohl, and Arnaud Doucet. Denoising diffusion samplers. In *International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=8pvnfTAbu1f.

Francisco Vargas, Shreyas Padhy, Denis Blessing, and Nikolas Nüsken. Transport meets variational inference: Controlled Monte Carlo diffusions. In *International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=PP1rudnxiW.

Qinsheng Zhang and Yongxin Chen. Path integral sampler: A stochastic control approach for sampling. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=_uCb2ynRu7Y.

Nicolas Zilberstein, Morteza Mardani, and Santiago Segarra. Repulsive latent score distillation for solving inverse problems. In *International Conference on Learning Representations*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/96d328a1f6d8396d8c8a62f2beee252a-Abstract-Conference.html.

A Proof of Proposition 1

Throughout we condition on the preceding states, so that the kernels p_k are fixed while the step-t states $X_t^m \sim p_m$, and hence r, remain random. The first identity holds for every realization and the second is a statement about the distribution of r, taken along a family of kernel configurations whose pairwise separations $\delta_{mk} := \|\mu_m - \mu_k\|/s$ all grow.

We use the shorthand

$$p_k(\,\cdot\,) := p_\theta(\,\cdot\,|\,x_{t-1}^k), \quad w_{mk}(\,\cdot\,) := p_k(\,\cdot\,)/p_m(\,\cdot\,) \tag{13}$$

so that

$$\mathcal{L}_{\text{Multi}} = \frac{1}{K^2} \sum_{m,k=0}^{K-1} w_{mk}(X_t^m)\, C(X_t^m). \tag{14}$$

Since $w_{mm} = 1$ and $\mathcal{L}_{\text{Single}}(X_t^k) := C(X_t^k)$, the diagonal $m = k$ summands of (14) reproduce $K^{-1}\mathcal{L}_{\text{Single}}$, while the off-diagonal $m \neq k$ summands carry the inter-trajectory coupling.

A.1 Part (1): Unbiased gradient

Fix θ. We first establish the equality of expected values

$$\mathbb{E}_{p_\theta^{\otimes K}}\big[\mathcal{L}_{\text{Multi}}\big] = \mathbb{E}_{p_\theta^{\otimes K}}\big[\mathcal{L}_{\text{Single}}\big]; \tag{15}$$

the corresponding gradient identity (9) follows as shown below.

Importance-weight cancellation. For $k \neq m$, condition on (x_{t-1}^k, x_{t-1}^m). Since w_{mk} is the Radon–Nikodym derivative of p_k w.r.t. p_m and $X_t \sim p_m$, the change-of-measure identity

$$\mathbb{E}_{X_t \sim\, p_m}\big[w_{mk}(X_t)\, \varphi(X_t)\big] = \mathbb{E}_{X_t \sim\, p_k}\big[\varphi(X_t)\big] \tag{16}$$

holds for every p_k-integrable φ. Setting $\varphi = C(\,\cdot\,)$ and using $\mathcal{L}_{\text{Single}}(X_t^k) := C(X_t^k)$,

$$\mathbb{E}_{p_\theta^{\otimes K}}\big[w_{mk}(X_t)\, C(X_t^m)\big] = \mathbb{E}_{p_\theta}\big[C(X_t^k)\big] \tag{17}$$

$$= \mathbb{E}_{p_\theta}\big[\mathcal{L}_{\text{Single}}(X_t^k)\big] =: \mu(\theta), \tag{18}$$

which holds for all index pairs (k, m). Substituting into (14),

$$\mathbb{E}_{p_\theta^{\otimes K}}\big[\mathcal{L}_{\text{Multi}}\big] = \frac{1}{K^2} \sum_{m,k=0}^{K-1} \mu(\theta) = \mu(\theta) = \mathbb{E}_{p_\theta^{\otimes K}}\big[\mathcal{L}_{\text{Single}}\big], \tag{19}$$

which proves (15).

Gradients. Equation (15) holds for every θ, so both expectations coincide as functions of θ, and so do their gradients. The trajectories are generated by reparameterization, $X_{0:T}^k = g_\theta(\varepsilon^k)$ with θ-independent noise ε^k, and gradients flow through the samples and the importance weights. Under standard integrability conditions, differentiation and expectation therefore commute,

$$\mathbb{E}\big[\nabla_\theta \mathcal{L}_{\text{Multi}}\big] = \nabla_\theta \mathbb{E}\big[\mathcal{L}_{\text{Multi}}\big] = \nabla_\theta \mathbb{E}\big[\mathcal{L}_{\text{Single}}\big] = \mathbb{E}\big[\nabla_\theta \mathcal{L}_{\text{Single}}\big].$$

A.2 Part (2): Loss ratio on realized batches

Since $w_{mm} = 1$, Eq. (14) gives

$$r = K^{-1} + \frac{1}{K} \frac{\sum_m \sum_{k \neq m} w_{mk}\, C(X_t^m)}{\sum_m C(X_t^m)}.$$

If $p_k = p_m$ for all k, m, all weights equal one and $r = 1$.

For the Gaussian kernels $p_m = \mathcal{N}(\mu_m, s^2 I)$ and $p_k = \mathcal{N}(\mu_k, s^2 I)$ the normalizing constants cancel, so with $d := \mu_m - \mu_k$ and the reparameterization $X_t^m = \mu_m + s\varepsilon$, $\varepsilon \sim \mathcal{N}(0, I)$,

$$\log w_{mk} = \frac{\|X_t^m - \mu_m\|^2 - \|X_t^m - \mu_k\|^2}{2s^2} = \frac{s^2\|\varepsilon\|^2 - \big(\|d\|^2 + 2s\, d^\top \varepsilon + s^2\|\varepsilon\|^2\big)}{2s^2} = -\frac{\delta^2}{2} - \delta Z,$$

where $\delta := \|d\|/s$ and $Z := d^\top \varepsilon/\|d\| \sim \mathcal{N}(0,1)$. The quadratic terms in ε cancel because both kernels share the same covariance, leaving an expression that is affine in ε. Using the Gaussian moment generating function $\mathbb{E}[e^{tZ}] = e^{t^2/2}$ at $t = -\delta$,

$$\mathbb{E}[w_{mk}] = \mathbb{E}\big[e^{-\delta^2/2-\delta Z}\big] = e^{-\delta^2/2}\,\mathbb{E}\big[e^{-\delta Z}\big] = e^{-\delta^2/2}e^{\delta^2/2} = 1$$

at every separation, in accordance with (16) for $\varphi \equiv 1$. Since log is increasing and $-\delta < 0$,

$$\mathbb{P}(w_{mk} > a) = \mathbb{P}\big(-\tfrac{\delta^2}{2} - \delta Z > \log a\big) = \mathbb{P}\Big(Z < -\frac{\log a + \delta^2/2}{\delta}\Big) = \Phi\Big(-\frac{\delta}{2} - \frac{\log a}{\delta}\Big) \xrightarrow[\delta\to\infty]{} 0$$

for every $a > 0$ and with Φ denoting the standard normal cumulative distribution function. Hence $w_{mk} \to 0$ in probability while its mean stays one.

B Application to DDS loss

In this section we exemplify the application of ITIS to a concrete type of diffusion sampler, namely the highly influential Denoising Diffusion Samplers (DDS) (Vargas et al., 2023) . This method parameterizes the sampling process as the reversal of an Ornstein-Uhlenbeck process. The training objective is derived from the corresponding reverse Kullback–Leibler divergence (rKL). The resulting objective is given by

$$\mathcal{L}_{\text{DDS}} = D_{\text{KL}}(p_\theta, q) = \mathbb{E}_{X_{0:T}\sim p_\theta}\left[2\sigma^2 \sum_{t=1}^{T-1} \frac{\lambda_t^2}{\alpha_t}\|s_\theta(t, X_t)\|^2 + \log\frac{\mathcal{N}(X_T; 0, \sigma^2 I)}{\pi_T(X_T)}\right] \tag{20}$$

with an approximation s_θ of the SDE scores, SDE parameters $\sigma, \lambda_t, \alpha_t$, with σ being the diffusion coefficient and $\lambda_t := 1 - \sqrt{1-\alpha_t}$. Here, the generic cost function $C(\cdot)$ from Eq. (3) takes the form $C(X_t) := \frac{2\sigma^2\lambda_t^2}{\alpha_t}\|s_\theta(t, X_t)\|^2$ for the running cost and $C(X_T) := \log\frac{\mathcal{N}(X_T;0,\sigma^2 I)}{\pi_T(X_T)}$ for the terminal cost. The derivation of the ITIS version of the loss in (20) proceeds analogously to that of (6). We further observe that the terms in (20), denoted by $\mathcal{L}_{\text{DDS}}$, can be estimated using a set of generated trajectories $\{x_{0:T}^k\}_{k=0}^{K-1} \overset{\text{iid}}{\sim} p_\theta$

$$\begin{aligned}\mathcal{L}_{\text{DDS,Single}}(\{X_{0:T}^k\}_{k=0}^{K-1}) \approx &\sum_{t=1}^{T-1}\frac{1}{K}\sum_{k=0}^{K-1}\mathbb{E}_{X_t\sim p_k(\cdot)}\left[\frac{2\sigma^2\lambda_t^2}{\alpha_t}\|s_\theta(t, X_t)\|^2\right] \\ &+ \frac{1}{K}\sum_{k=0}^{K-1}\mathbb{E}_{X_T\sim p_k(\cdot)}\left[\log\frac{\mathcal{N}(X_T;0,\sigma^2 I)}{\pi_T(X_T)}\right].\end{aligned}$$

We now apply importance sampling once more and express the expectation with respect to the trajectory m. We then approximate this expectation using a single-sample estimator

$$\begin{aligned}&\mathcal{L}_{\text{DDS}}(\{X_{0:T}^k\}_{k=0}^{K-1})_m \approx \sum_{t=1}^{T-1}\frac{1}{K}\sum_{k=0}^{K-1}\mathbb{E}_{X_t\sim p_m(\cdot)}\left[w_{mk}(X_t)\ \frac{2\sigma^2\lambda_t^2}{\alpha_t}\|s_\theta(t, X_t)\|^2\right] \\ &+ \frac{1}{K}\sum_{k=0}^{K-1}\mathbb{E}_{X_T\sim p_m(\cdot)}\left[w_{mk}(X_T)\ \log\frac{\mathcal{N}(X_T;0,\sigma^2 I)}{\pi_T(X_T)}\right] \\ &\approx \sum_{t=1}^{T-1}\frac{1}{K}\sum_{k=0}^{K-1} w_{mk}(x_t^m)\ \frac{2\sigma^2\lambda_t^2}{\alpha_t}\|s_\theta(t, x_t^m)\|^2 + \frac{1}{K}\sum_{k=0}^{K-1} w_{mk}(x_T^m)\ \log\frac{\mathcal{N}(X_T^m;0,\sigma^2 I)}{\pi_T(x_T^m)}.\end{aligned}$$

Finally, we average over all trajectories m

$$\begin{aligned}\mathcal{L}_{\text{DDS, Multi}}(\{X_{0:T}^k\}_{k=0}^{K-1}) \approx &\sum_{t=1}^{T-1}\frac{1}{K^2}\sum_{m,k=0}^{K-1} w_{mk}(x_t^m)\ \frac{2\sigma^2\lambda_t^2}{\alpha_t}\|s_\theta(t, x_t^m)\|^2 \\ &+ \frac{1}{K^2}\sum_{m,k=0}^{K-1} w_{mk}(x_T^m)\ \log\frac{\mathcal{N}(x_T^m;0,\sigma^2 I)}{\pi_T(x_T^m)}\end{aligned}$$

C Experimental Details

C.1 Runtime and Memory Cost of ITIS

ITIS reuses the batch of trajectories that the sampler generates anyway, so it requires no additional forward passes and adds no additional energy evaluations per iteration. Its only additional work is the evaluation of the K^2 pairwise transition densities per diffusion step, together with the corresponding entries in the backward graph. Since the cost of a training iteration is otherwise dominated by the T sequential network evaluations, the relevant question is how the quadratic term compares to this cost in the regimes in which ITIS is used. Table 3 reports per-iteration training time and peak GPU memory for DDS on GMM40 in 50d, measured on a single Nvidia A100 40GB with batch size $B = 2000$. Per-iteration time is the slope of total wall-clock time between runs of 100 and 300 iterations, which removes fixed costs such as JIT compilation and evaluation. Peak memory is the largest per-process usage reported by nvidia-smi (polled once per second) during a 300-iteration run with XLA preallocation disabled.

Table 3: Per-iteration training time and peak GPU memory for DDS on GMM40 ($d = 50$), as a function of the number of diffusion steps T and the chunk size K, where $K = 1$ is the single-sample loss.

	Time per iteration (ms)				Peak GPU memory (MiB)			
T	K=1	K=5	K=50	K=200	K=1	K=5	K=50	K=200
5	25.2	34.7	28.4	37.7	322	322	386	386
20	39.4	31.2	32.7	36.8	514	514	514	770
50	46.5	56.4	77.3	77.0	770	770	1282	1282
128	66.6	72.1	118.6	148.8	2306	2306	4354	15264

Both quantities are governed primarily by T, and the influence of K grows with it. Up to $T = 20$ the chunk size is essentially free: raising K from 1 to 200 leaves the iteration time unchanged to within the measurement scatter, which is of the order of the difference between $K = 1$ and $K = 5$ at $T = 20$, and increases peak memory by a single allocation step. Memory is in fact identical for $K \leq 50$ at every $T \leq 50$, so the pairwise quantities remain below the granularity of the memory allocator here. Only at $T = 128$ does the quadratic term become clearly visible, since the backward graph retains per-pair quantities at every step, so that memory grows with TK^2.

The configurations used in our experiments lie in the inexpensive part of this grid. The DDS runs in Sec. 5 use $T \leq 35$ throughout, with $T = 10$ and $K = 100$ on GMM40 in 50d, where the measured cost is comparable to that of the single-sample loss at the same T and below that of the single-sample loss at the $T = 128$ commonly used in this benchmark suite.

C.2 Hyperparameter cross-over between $K = 1$ and $K = 400$

Because each K in Fig. 1 is tuned separately, the comparison across K is one of equal tuning budget rather than of identical hyperparameters. To probe how much of the effect is attributable to the estimator and how much to its operating point, we take the tuned pairs of the two extreme chunk sizes from the 50d panel, $(\text{noise scale}, \eta) = (19.152,\ 6.16 \cdot 10^{-6})$ for $K = 1$ and $(7.937,\ 3.975 \cdot 10^{-5})$ for $K = 400$, and evaluate both estimators under both settings. Note that $K = 1$ under the multi-sample objective is exactly the single-sample loss, since the sole weight $w_{mm} = 1$.

Table 4: Cross-over study on GMM40 ($d = 50$, prior scale 25, DDS). Rows select the estimator, columns the tuned hyperparameter pair. Sinkhorn distance (SD, ↓) and entropic mode coverage (EMC, ↑), mean ± standard error over 5 seeds. Diagonal entries are reproduced from Figure 1.

	K=1's (noise, η)		K=400's (noise, η)	
Estimator	SD	EMC	SD	EMC
$K = 1$	$\mathbf{4418 \pm 249}$	$\mathbf{0.9905 \pm 0.0016}$	9093 ± 124	0.9516 ± 0.0017
$K = 400$	6236 ± 196	0.9855 ± 0.0008	$\mathbf{3139 \pm 107}$	$\mathbf{0.9964 \pm 0.0002}$

Each estimator performs best at its own tuned operating point, and transplanting the other's hyperparameters roughly doubles the Sinkhorn distance in both directions ($4418 \to 9093$ and $3139 \to 6236$). This result confirms that the observed performance difference between $K = 1$ and $K = 400$ cannot be simply explained via the other hyperparameters.

C.3 Mode-Discovery Study: Tuning Protocol

Fixed configuration. All runs use DDS with $T = 10$ diffusion steps, a cosine β schedule, a constant learning-rate schedule, Langevin guidance, no importance-weight clipping, and 10,000 training iterations. The two panels differ only in the target dimension and the prior scale σ_0, which is deliberately reduced relative to the standard benchmark so that the prior does not already cover the modes: $d = 50$ uses $\sigma_0 = 25$ (rather than 40), batch size 2000 and β-offset 10^{-3}; $d = 2$ uses $\sigma_0 = 10$, batch size 512 and β-offset $5 \cdot 10^{-2}$. The component centres $\{\mu_i\}$ are held fixed across all runs via a shared random seed.

Search space and budget. For every chunk size K independently, we tune the two parameters that govern the operating point of the sampler, i.e. the noise scale and the learning rate η, with an equal budget per K, so that no value of K is favoured by receiving more tuning. We use a Tree-structured Parzen Estimator sampler (Akiba et al., 2019) in two stages:

1. **Broad stage.** Log-uniform search over noise scale $\in [0.5, 20]$ and $\eta \in [10^{-6}, 5 \cdot 10^{-3}]$ for $d = 50$, and noise scale $\in [3, 30]$ and $\eta \in [10^{-6}, 10^{-2}]$ for $d = 2$. The same box is used for every K within a panel.
2. **Zoom stage.** A second search of equal budget, restricted to $[\frac{1}{2}s^\star,\ 2s^\star]$ for the noise scale and $[\frac{1}{4}\eta^\star,\ 4\eta^\star]$ for the learning rate, where $(s^\star, \eta^\star)$ is the best configuration of the broad stage, clipped to the broad box.

The budget is 15 trials per stage for $d = 2$ and 22 trials per stage for $d = 50$; the larger budget in 50d reflects its rougher objective landscape. Trials are scored by the Sinkhorn distance at a single fixed seed (0). The best configuration over both stages is then retained.

Reporting. The selected configuration is then re-run over 5 seeds (0–4). Performance criteria are computed from 2000 samples at 21 evaluation checkpoints spread over training, and a moving average of length 5 (MA-5) is applied over consecutive checkpoints. We report the *minimum* of the MA-5 Sinkhorn distance over training, and the MA-5 entropic mode coverage evaluated at the same checkpoint at which that minimum is attained. Cells report the mean $\pm$ standard error over the 5 seeds.

Selected hyperparameters. Table 5 lists the configuration selected by the search discussed above. Only two quantities are tuned: the noise scale and the learning rate η. They are tuned separately for each chunk size K and each dimensionality. Everything else is held fixed at the values in Table 6.

Table 5: Selected hyperparameters for each value in Fig. 1. The noise scale and learning rate η are the only tuned quantities; each (K, panel) pair is tuned independently with an equal budget.

	$d = 50$, prior scale 25		$d = 2$, prior scale 10	
K	noise scale	η	noise scale	η
1	19.1520	$6.16000 \cdot 10^{-6}$	6.86654	$4.88890 \cdot 10^{-5}$
2	19.1427	$6.06865 \cdot 10^{-6}$	3.58788	$6.87997 \cdot 10^{-5}$
5	19.8226	$5.16117 \cdot 10^{-6}$	26.9843	$3.33233 \cdot 10^{-4}$
10	19.7443	$6.45739 \cdot 10^{-6}$	32.5264	$3.93246 \cdot 10^{-4}$
50	9.11581	$2.75830 \cdot 10^{-5}$	7.34062	$1.30351 \cdot 10^{-4}$
100	19.8118	$1.88211 \cdot 10^{-5}$	12.6362	$1.01174 \cdot 10^{-4}$
200	8.33657	$4.44143 \cdot 10^{-5}$	28.3579	$7.59446 \cdot 10^{-5}$
400	7.93660	$3.97520 \cdot 10^{-5}$	29.5126	$2.33460 \cdot 10^{-4}$

Table 6: Settings held fixed. The two panels differ only in the target dimension, the prior scale, the batch size and the β offset.

	$d = 50$, prior scale 25	$d = 2$, prior scale 10
Target	GMM40 (40 components, fixed centres)	
Sampler	DDS	
Target dimension d	50	2
Initial / prior scale σ_0	25.0	10.0
Batch size B	2000	512
β offset	10^{-3}	$5 \cdot 10^{-2}$
Diffusion steps T	10	
β schedule	cosine	
Training iterations	10,000	
Learning-rate schedule	constant	
Evaluation samples	2000	
Evaluation checkpoints	21, evenly spaced over training	
Seeds	$0, 1, 2, 3, 4$	

C.4 Training & Evaluation Specifics

Here, we offer an overview of the main set of hyperparameters tuned and the evaluation protocol used across for our training runs.

Hyperparameter optimization. The multi-sample objective introduces a minimal hyperparameter footprint relative to the gains it provides. The only parameters tuned across experiments are: (i) chunk size K, (ii) maximum diffusion, (iii) learning rate with optional cosine decay (Loshchilov & Hutter, 2017), (iv) number of diffusion steps T, and (v) importance weights clip to the interval $[\exp(-n), \exp(+n)]$ – beneficial for numerical stability for specific experimental settings. Complete method-specific configurations and search spaces are covered in Section C.7.

Evaluation protocol. To ensure a fair and consistent comparison against all baselines, our evaluation protocol follows Blessing et al. (2024) and Chen et al. (2025). For each algorithm, all performance criteria are evaluated 100 times over the course of training; a moving average of length 5 (MA-5) is applied over consecutive evaluations to reduce within-run noise and yield stable estimates. Our training runs use $40,000$ iterations unless stated otherwise. The Sinkhorn values are reported when ground-truth samples are available, while the MMD values are also reported for targets, if available. We use 2000 samples to compute the performance criteria. All results are reported over 4 independent random seeds unless stated otherwise. The best-performing checkpoint is selected per seed prior to averaging.

Metrics. When access to the target distribution is available, Probability Metrics (PMs) can be used to measure the approximation quality (Arenz et al., 2018). We follow Blessing et al. (2024), which identifies the best suited PMs to assess the fidelity and employ them to gauge the mode coverage: Sinkhorn distance (Cuturi, 2013; Peyré & Cuturi, 2019) and the Maximum Mean Discrepancy (MMD) (Gretton et al., 2012). The former, is rooted in OT and computes the minimum displacement cost to morph the empirical measure into the target. As also discussed further in Blessing et al. (2024), low OT distances correspond to better mode coverage (i.e., avoidance of mode collapse). The latter characterizes distributional divergence by measuring the distance between mean embeddings within a Reproducing Kernel Hilbert Space (RKHS) (Gretton et al., 2012). We refer the reader to Section C.5 for specifics regarding these PMs.

C.5 Probability Metrics

Maximum Mean Discrepancy. The Maximum Mean Discrepancy (MMD) (Gretton et al., 2012) is a kernel-based measure of discrepancy between probability distributions, defined as the distance between their mean embeddings in a reproducing kernel Hilbert space (RKHS) (Aronszajn, 1950) associated with a kernel k. Given the RKHS $\mathcal{H}_k$, the MMD between the model distribution q_θ and the target π is the integral probability metric quantifying the dissimilarity between the two

distributions:

$$\mathrm{MMD}_k(q_\theta, \pi) = \sup_{f \in \mathcal{H}_k:\, \|f\|_{\mathcal{H}_k} \leq 1} \left(\mathbb{E}_{x\sim q_\theta}[f(x)] - \mathbb{E}_{y\sim\pi}[f(y)]\right), \tag{21}$$

which satisfies $\mathrm{MMD}_k(q_\theta, \pi) \geq 0$, with equality if and only if $q_\theta = \pi$. Given sample sets $X = \{x_i\}_{i=1}^n \sim q_\theta$ and $Y = \{y_j\}_{j=1}^m \sim \pi$, the minimum-variance unbiased estimator of MMD_k is:

$$\widehat{\mathrm{MMD}}_k(q_\theta, \pi) = \left(\frac{1}{n(n-1)} \sum_{i \neq j} k(x_i, x_j) + \frac{1}{m(m-1)} \sum_{i \neq j} k(y_i, y_j) - \frac{2}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} k(x_i, y_j) \right)^{1/2}. \tag{22}$$

In all experiments, we adopt a squared exponential kernel $k(x, y) = \exp\left(-\|x - y\|_2^2/\alpha\right)$, where the bandwidth α is selected via the median heuristic (Gretton et al., 2012).

Entropic Optimal Transport Distance. The 2-Wasserstein distance between q_θ and π is defined as:

$$W_2(q_\theta, \pi) = \left(\inf_{\xi \in \Gamma(q_\theta, \pi)} \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x, y)\, \xi(x, y)\, \mathrm{d}x\, \mathrm{d}y \right)^{1/2}, \tag{23}$$

where $\Gamma(q_\theta, \pi)$ denotes the set of couplings with marginals q_θ and π, and $c(x, y) = \|x - y\|^2$ is the metric, i.e. the squared Euclidean cost. To obtain a computationally tractable objective, we adopt the entropy-regularised formulation (Cuturi, 2013):

$$W_{2,\varepsilon}(q_\theta, \pi) = \left(\inf_{\xi \in \Gamma(q_\theta, \pi)} \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x, y)\, \xi(x, y)\, \mathrm{d}x\, \mathrm{d}y - \varepsilon\, \mathcal{H}(\xi) \right)^{1/2}, \tag{24}$$

where $\mathcal{H}(\xi) = -\int_{\mathbb{R}^d \times \mathbb{R}^d} \xi(x, y) \log \xi(x, y)\, \mathrm{d}x\, \mathrm{d}y$ is the differential entropy of the coupling. We set $\varepsilon = 10^{-3}$, analogous to setup as (Blessing et al., 2024). Our implementation also uses the `ott-jax` library (Cuturi et al., 2022) to compute $W_{2,\varepsilon}$.

C.5.1 Particle potential metrics

The target Boltzmann distributions for our benchmark particle systems (DW-4, LJ-13, and LJ-55) are inherently invariant under the Euclidean group. This comprises of translations, rotations, and reflections and symmetric groups governing particle permutations. Directly evaluating raw coordinates in ambient space fails to account for these geometric and combinatorial degeneracies, incorrectly penalizing physically indistinguishable configurations. To ensure a rigorous comparison, we assess all methods using symmetry-aware metrics that are explicitly invariant under the corresponding joint group actions.

Geometric Wasserstein distance. The energy and diffusion process are invariant under rotation and permutation symmetries. We measure the distance between generated and ground-truth point clouds with a symmetry-aware Wasserstein-2 metric. For empirical distributions $\hat{\nu}$ (generated) and ν (ground truth),

$$\mathcal{W}_2(\mu, \nu) = \left(\inf_{\pi} \int \pi(x, y)\, d(x, y)^2\, \mathrm{d}x\, \mathrm{d}y \right)^{\frac{1}{2}} \tag{25}$$

The metric $d(x, y)$ defined in eq. (25) takes into account all point-cloud symmetries, and is obtained by minimizing the squared Euclidean distance over all possible combinations of symmetries such as rotations, reflections ($O(\mathrm{d})$)), and permutation symmetries ($\mathbb{S}(k)$) for k particles of spatial dimension d.

$$d(x, y) = \min_{R \in O(d),\, P \in \mathbb{S}(k)} \left\| x - (R \otimes P)\, y \right\|_2. \tag{26}$$

Since eq. (26) is computationally intractable, thus we opt for the same methodology described in (Havens et al., 2025; Köhler et al., 2020) which approximates the minimizer by performing a sequential search:

$$d(x, y) = \min_{R \in SO(\mathrm{d})} ||x - (R \otimes \tilde{P})y||_2^2\,, \tilde{P} = \operatorname*{arg\,min}_{P \in \mathbb{S}(k)} ||x - Py||_2^2\,. \tag{27}$$

Energy Wasserstein distance ($E(.)\mathcal{W}_2$) Moreover, we also report the Wasserstein-2 distance on the 1-dimensional energy distributions.This boils down to the standard euclidean $\mathcal{W}_2$ metric between the energy distributions defined as (Havens et al., 2025):

$$\mathcal{W}_2(E_\mu, E_\nu) = \left(\inf_{\pi} \pi(x,y)|x-y|^2 \, dx \, dy \right)^{\frac{1}{2}}. \tag{28}$$

C.6 Continuous energy functions

C.6.1 Synthetic energy functions

We evaluate the performance on the following target densities exhibiting varying dimensionalities and multimodalities: i) 5-dimensional 32-mode Many-Well task (MW54) [$d = 5, m = 5,$ and $\delta = 4$], ii) Robot1 and Robot4 (Arenz et al., 2018) tasks constituting robotics control problems over a configuration space of a planar 10-joint robotic arm. These pose a 10-dimensional sampling problem with multiple well-separated, including sharp modes that challenge both global exploration and local precision, (iii) the famous 10-dimensional Funnel distribution (Neal, 2003) and (iv) the following mixture models: a 40-mode Gaussian mixture model in 50d and a 10-mode Student-t mixture model (MoS) in 50d. These mixture benchmarks have the advantages of having well-defined, known numbers of modes and in addition, they offer straightforward ways of obtaining ground-truth samples allowing a robust evaluation of the generated samples that is sensitive to mode collapse.

For all synthetic benchmarks, ground-truth samples are available and used for evaluation. Here, for the sake of completeness we describe our targets in detail. These synthetic targets utilized in our work are taken from (Blessing et al., 2024; Chen et al., 2025) and we also direct the reader to their works and the references therein for detailed expositions.

Many-Well (MW). The Many-Well potential is a prototypical target from molecular dynamics, corresponding to the stationary distribution of Langevin dynamics under a multi-modal energy landscape. The unnormalised target density is:

$$\rho_{\text{target}}(x) = \exp\left(-\sum_{i=1}^{m} (x_i^2 - \delta)^2 - \frac{1}{2} \sum_{i=m+1}^{d} x_i^2 \right). \tag{29}$$

Following Berner et al. (2024), we set $d = 5$, $m = 5$, and $\delta = 4$, yielding $2^m = 32$ well-separated modes that stress-test the capacity of samplers to maintain broad distributional coverage.

Robot. The Robot targets (Arenz et al., 2020) (Robot1, Robot4) require learning joint configurations of a 10-degrees-of-freedom planar robot, parameterised by $\alpha = (\alpha_1, \ldots, \alpha_{10})$, such that the end-effector reaches a desired goal position while maintaining smooth configurations. The target density factorises as:

$$\rho_{\text{target}}(\alpha) = p_{\text{conf}}(\alpha)\, p_{\text{cart}}(\alpha), \tag{30}$$

where p_{conf} enforces configuration smoothness and p_{cart} penalises deviations from the goal. Specifically, p_{conf} is a zero-mean Gaussian with diagonal covariance, assigning variance 1 to the first joint angle α_1 and variance 4×10^{-2} to the remaining angles $\alpha_2, \ldots, \alpha_{10}$. The Cartesian coordinates of the i-th joint are defined by the forward kinematics map:

$$x_i(\alpha) = \sum_{j=1}^{i} \cos(\alpha_j), \qquad y_i(\alpha) = \sum_{j=1}^{i} \sin(\alpha_j), \qquad i = 0, \ldots, 10. \tag{31}$$

In the **Robot1** task, a single goal is placed at $(7, 0)$, and p_{cart} is defined as:

$$p_{\text{cart}}(\alpha) = \mathcal{N}\left(\begin{pmatrix} x_{10}(\alpha) \\ y_{10}(\alpha) \end{pmatrix}; \begin{pmatrix} 7 \\ 0 \end{pmatrix}, 10^{-4} I \right). \tag{32}$$

In the **Robot4** task, four goals are placed at $(\pm 7, 0)$ and $(0, \pm 7)$, and p_{cart} is given by the maximum over the four corresponding Gaussian distributions as above, up to a normalising constant. Groundtruth samples are obtained via long-run slice sampling (Neal, 2003) and sourced from the repository of Arenz et al. (2020).

Mixture models (GMM and MoS). For the mixture distribution benchmarks, the target takes the form:

$$p_{\text{target}} = \frac{1}{m}\sum_{i=1}^{m} p_i. \tag{33}$$

The **Gaussian Mixture Model (GMM)** which we use from the Blessing et al. (2024) comprises $m = 40$ components with $p_i = \mathcal{N}(\mu_i, I)$, where $\mu_i \sim \mathcal{U}_d(-40, 40)$ and $\mathcal{U}_d(l, u)$ denotes the uniform distribution on $[l, u]^d$. We set $d = 50$ in all main experiments. The **Mixture of Student's t-distributions (MoS)** (Blessing et al., 2024) comprises $m = 10$ Student's t-distributions with 2 degrees of freedom, defined as $p_i = t_2 + \mu_i$ with $\mu_i \sim \mathcal{U}_d(-10, 10)$, where μ_i denotes the translation of each component. We set $d = 50$ throughout. For both GMM and MoS, the component centres $\{\mu_i\}$ are held fixed across all experiments via a shared random seed.

Funnel. The Funnel distribution (Neal, 2003) is a canonical benchmark for samplers operating in hierarchically structured geometries. It is defined as:

$$p_{\text{target}}(x) = \mathcal{N}(x_1;\, 0,\, \sigma^2)\,\mathcal{N}(x_2, \ldots, x_d;\, 0,\, e^{x_1} I), \tag{34}$$

with $\sigma^2 = 9$, for any $d \geq 2$. Although, the funnel distribution is not strictly multimodal in nature, the funnel geometry induces a strongly varying posterior curvature that challenges samplers relying on fixed-step-size dynamics. We set $d = 10$ in all main experiments.

C.6.2 Particle energy functions

DW-4 Potential. Originally introduced in (Köhler et al., 2020), the DW4 system models a four-particle double-well potential in two-dimensional space ($d = 8$). The collective state vector is defined as $x = [x_1, x_2, x_3, x_4] \in \mathbb{R}^d$, where each particle position satisfies $x_i \in \mathbb{R}^2$. The unnormalized energy function for the double-well potential reads:

$$E(x) = \frac{1}{2\tau}\sum_{i<j}\left[a(d_{ij} - d_0) + b(d_{ij} - d_0)^2 + c(d_{ij} - d_0)^4\right], \tag{35}$$

where $d_{ij} = \|x_i - x_j\|_2$ denotes the pairwise Euclidean distance between particles i and j. Following standard experimental configurations, we set the hyperparameters to $a = 0$, $b = -4$, $c = 0.9$, $d_0 = 1$, and temperature $\tau = 1$.

LJ-13 and LJ-55 Clusters. We further scale our evaluation to higher dimensions using the well-known Lennard-Jones (LJ) potentials, which are standard classical models for atomic interactions. Consider n particles in a three-dimensional space with the state vector $x = [x_1, \ldots, x_n] \in \mathbb{R}^{3n}$ and positions $x_i \in \mathbb{R}^3$. The nomenclature "LJ-13" and "LJ-55" indicates the respective particle counts of $n = 13$ ($d = 39$) and $n = 55$ ($d = 165$).

To constrain the clusters and prevent translational drift, the unnormalized energy incorporates an additional center-of-mass harmonic confinement term:

$$E(x) = \frac{\epsilon}{2\tau}\sum_{i\neq j}\left[\left(\frac{r_m}{d_{ij}}\right)^{12} - 2\left(\frac{r_m}{d_{ij}}\right)^{6}\right] + \frac{c}{2}\sum_{i=1}^{n}\|x_i - C(x)\|^2, \tag{36}$$

where $d_{ij} = \|x_i - x_j\|_2$ is the inter-particle Euclidean distance, and $C(x) = \frac{1}{n}\sum_{i=1}^{n} x_i$ defines the center of mass. In accordance with prior description in the literature (Liu et al., 2025b; Havens et al., 2025), we adopt the same parameter configurations $r_m = 1$, $\epsilon = 1$, $c = 0.5$, and $\tau = 1$.

C.7 Experimental Configurations: Diffusion Samplers and Target Benchmarks

To ensure reproducibility and facilitate a rigorous comparative analysis, we detail the specific hyperparameter configurations for our ITIS-based multi-sample objective across different diffusion samplers (DDS, DIS, and PIS) and target densities.

DDS.

i) *MW54* ($d = 5$): The diffusion steps used were: $T = 35$, with an initial learning-rate of $= 1.5$e-6 and a noise scale: 3, initial-scale: 1.0, multi-sample chunk size: 125.

ii) *Robot1* ($d = 10$): The number of diffusion steps utilized are: $T = 25$, with a learning-rate = 3e-3 and noise scale: 1.8, initial-scale: 1.0, multi-sample chunk size: 100

iii) *Robot4* ($d = 10$): The number of diffusion steps utilized are: $T = 10$, with a learning-rate $= 1.5$e-6 and noise scale: 7.5, initial-scale: 0.15, multi-sample chunk size: 115

iv) *Funnel (*$d = 10$*)*: The number of diffusion steps is $T = 20$, the learning rate is $\eta = 2$e-3, initial scale: 5.0, and noise-scale = 1.0. For numerical reasons it was found beneficial in these experiments to clip the importance weights to the range of $[\exp(-5), \exp(+5)] = [0.007, 148]$.

v) *GMM40* ($d = 50$): The number of diffusion steps utilized are: $T = 10$, noise-scale: 6.0, learning rate = 1e-5, initial scale: 40.0, and a multi-sample chunk size: $K = 100$.

vi) *MoS* ($d = 50$): The number of diffusion steps utilized are: $T = 20$, noise-scale: 1.0, learning rate : 5e-3, initial scale: 8.0, and a multi-sample chunk size: $K = 5$.

vii) *DW-4* ($d = 8$) The number of diffusion steps utilized are: $T = 80$, noise-scale: 9.9, learning rate : 9e-3, initial scale: 1.6, and a multi-sample chunk size: $K = 120$.

viii) *LJ-13* ($d = 39$) The number of diffusion steps utilized are: $T = 70$, noise-scale: 4.5, learning rate : 7e-3, initial scale: 0.6, and a multi-sample chunk size: $K = 50$.

ix) *LJ-55* ($d = 165$) The number of diffusion steps utilized are: $T = 80$, noise-scale: 4.3, learning rate : 8e-3, initial scale: 2.6, and a multi-sample chunk size: $K = 60$.

DIS.

i) *Funnel (*$d = 10$*).* The multi-sample chunk size is $K = 100$, number of diffusion steps is $T = 7$, the learning rate is $\eta = 3$e-6, maximum diffusion: 0.25, and initial-scale: 2.1.

ii) *GMM40* ($d = 50$): The multi-sample chunksize is $K = 60$, the number of diffusion steps utilized are: $T = 25$, with a learning-rate of : 4e-4 and a maximum diffusion: 25 and initial-scale : 1.9.

iii) *MoS* ($d = 50$): The multi-sample chunk-size used in the experiment is $K = 50$, the number of diffusion steps utilized are: $T = 35$. To stabilize training and prevent MMD divergence in later iterations, we apply cosine learning rate decay (Loshchilov & Hutter, 2017) with an initial learning rate of $\eta = 1$e-5. We use a maximum diffusion: 0.4 and initial-scale: 5.

PIS.

The ideal multi-sample chunk size for the PIS diffusion sampler experiments is: `multi_batch_chunk_size` : $K = 5$

i) *MW54* ($d = 5$): For the 32-mode Many-Well task (MW54) in 5d, the diffusion steps used were: $T = 15$, with an initial learning-rate of : 5e-5 and a maximum diffusion:200.

ii) *Robot 1* ($d = 10$): For the robot with 1 goal in 10d, the number of diffusion steps utilized are: $T = 35$, with a initial learning-rate of : 9e-6 annealed down to a final value of 4.5e-7 (decay factor 0.05) and a maximum diffusion: 0.058. We apply importance weights clipping in the range of $[\exp(-2), \exp(+2)] = [0.135, 7.38]$

iii) *Robot 4* ($d = 10$): For the robot with 4 goals in 10d, the number of diffusion steps utilized are: $T = 35$, with an initial learning-rate of : 8.6e-6 annealed down to a final value of 8e-7 (decay factor 0.1) and a maximum diffusion: 0.04. We apply importance weights clipping in the range of $[\exp(-2), \exp(+2)] = [0.135, 7.38]$

iv) *Funnel (*$d = 10$*).* The number of diffusion steps is $T = 25$, the learning rate is $\eta = 5$e-4 (constant), and maximum diffusion $= 7.5$.

v) *GMM40* ($d = 50$): For the GMM 40 mode target in 50d, the number of diffusion steps utilized are: $T = 15$, with a learning-rate of : 1.6e-4 and a maximum diffusion: 200.

vi) *MoS* ($d = 50$): For the student mixture model in 50d, the number of diffusion steps utilized are: $T = 20$. We start with an initial learning rate of $\eta = 1$e-4, annealed to a final value of 5e-5 (decay factor 0.5). We use a maximum diffusion: 25. We apply importance weights clipping in the range of $[\exp(-3), \exp(+3)] = [0.049, 20.08]$

Table of hyperparameter choices for Multi-sample objective. The Table 7 enumerates the optimized configurations for our proposed interaction-based objective across six benchmark targets. Parameters include number of diffusion steps (T), interaction chunk size (K), and learning rate (η) strategies. PISGradNet is the shared architecture across all three diffusion samplers (DDS, DIS, PIS) and it applies to all targets. Note that initial scale (σ_0) and maximum diffusion are sampler-dependent. Importance weight clipping intervals are specified where utilized to prevent gradient instability in high-variance regimes.

Table 7: **Hyperparameter suite for multi-sample objectives.** The table summarizes the hyperparameter values used for our ITIS-based training. The dashed line in DIS indicate that not baseline values were made available for these targets in Chen et al. (2025).

	MW54 (5d)	**Robot1 (10d)**	**Robot4 (10d)**	**Funnel (10d)**	**GMM40 (50d)**	**MoS (50d)**
DDS						
Initial Scale (σ_0)	1.0	1.0	0.15	5.0	40.0	8.0
Noise Scale	3.0	1.8	7.5	1.0	6.0	1.0
Chunk size (K)	125	100	115	100	100	5
Diffusion steps (T)	35	25	10	20	10	20
LR (η)	1.5e-6	3e-3	1.5e-6	2e-3	1e-5	5e-3
Importance weight clip	—	—	—	[e^{-5}, e^{5}]	—	—
DIS						
Maximum diffusion	—	—	—	0.25	25.0	0.4
Initial scale (σ_0)	—	—	—	2.1	1.9	5.0
Chunk size (K)	—	—	—	100	60	50
Diffusion steps (T)	—	—	—	7	25	35
LR (η)	—	—	—	3e-6	4e-4	1e-5
LR Schedule	—	—	—	Constant	Constant	Cosine
PIS						
Maximum diffusion	200.0	0.058	0.04	7.5	200.0	25.0
Chunk size (K)	5	5	5	5	5	5
Diffusion steps (T)	15	35	35	25	15	20
LR (η)	5e-5	9e-6	8.6e-6	5e-4	1.6e-4	1e-4
LR Decay Factor	—	0.05	0.1	—	—	0.5
Importance weight clip	—	[e^{-2}, e^{2}]	[e^{-2}, e^{2}]	—	—	[e^{-3}, e^{3}]

Table 8: **Maximum Mean Discrepancy (MMD ↓) across Varied Target Densities.** Comparison of our ITIS-based multi-sample interaction objective (*Multi*) against baseline single-sample diffusion samplers. The MMD values corresponding to Table 2 is reported. best MA-5 results are indicated in **bold**.

Method	**Funnel (10d)**	**GMM40 (50d)**	**MoS (50d)**
DDS (Single)	$0.172_{\pm 0.031}$	$0.046_{\pm 0.001}$	$0.131_{\pm 0.001}$
DDS (Multi)	$0.202_{\pm 0.011}$	$\mathbf{0.032_{\pm 0.000}}$	$0.107_{\pm 0.004}$
DIS (Single)	$0.159_{\pm 0.036}$	$0.034_{\pm 0.003}$	$0.155_{\pm 0.001}$
DIS (Multi)	$\mathbf{0.158_{\pm 0.031}}$	$0.032_{\pm 0.001}$	$0.142_{\pm 0.003}$
PIS (Single)	$0.210_{\pm 0.001}$	$0.083_{\pm 0.000}$	$0.218_{\pm 0.001}$
PIS (Multi)	$0.178_{\pm 0.001}$	$0.050_{\pm 0.000}$	$\mathbf{0.078_{\pm 0.001}}$

D MISCELLANEOUS

Sinkhorn distance mismatch. We identify a notable mismatch in the Sinkhorn distance metrics reported in prior work regarding the Many Well (**MW54**) benchmark. Following the evaluation protocol established by Chen et al. (2025), performance should be reported as a moving average over five

consecutive samples (MA-5). However, upon replicating the experiments using the official repository: https://github.com/anonymous3141/SCLD and the parameters specified by Chen et al. (2025), we observed that the values reported in their primary results correspond to the absolute minimum Sinkhorn distance rather than the MA-5 metric. To ensure a rigorous and standardized comparison, we re-evaluated the **MW54** benchmark across four independent seeds. Consistent with the evaluation criteria in Blessing et al. (2024) and Chen et al. (2025), we report the corrected MA-5 minimum value averaged over these seeds in Table 2. This corrected value is denoted with an asterisk (0.52 vs. the previously reported 0.42) to maintain transparency and facilitate fair cross-methodological assessment. We mainly use Nvidia A100 40GB GPUs to train the compared methods.