ERROR-CORRECTED INFERENCE-TIME SCALING FOR IMPERFECT DIFFUSION MODELS

Anonymous authors
Paper under double-blind review

ABSTRACT

Inference-time scaling adapts pretrained diffusion models to new sampling tasks without additional training. Existing methods rely primarily on Monte Carlo sampling with more particles, yet are premised on the pretrained model being exact. In practice, data and training limitations make the model imperfect, and these methods inherit its error. More particles reduce Monte Carlo error but cannot remove the mismatch between the endpoint and the desired target or the error in tracking the prescribed probability path. We introduce the *Energy-based Feynman–Kac Corrector (EBFKC)*, a framework for energy-based diffusion models that corrects these errors on the fly given a reference energy. We first derive Feynman–Kac dynamics that track a prescribed path exactly in the continuous-time population limit even when the model is imperfect, and approximate these dynamics using sequential Monte Carlo with variance-controlling guidance. To remove the endpoint mismatch, we use the pretrained energy as a surrogate along the diffusion path and progressively incorporate the discrepancy between the learned and target terminal energies. Experiments on Gaussian mixture models, particle systems, alanine dipeptide, and alanine tetrapeptide show that our method closely matches target distributions and molecular free-energy profiles under annealing and reward tilting, whereas standard inference-time scaling baselines retain substantial sampling errors.

1 INTRODUCTION

Diffusion models (Sohl-Dickstein et al., 2015; Ho et al., 2020; Song et al., 2021) have become a leading paradigm in generative modeling, achieving remarkable performance across high-dimensional applications ranging from visual synthesis (Peebles & Xie, 2023; Rombach et al., 2022; Karras et al., 2022) to scientific discovery (Watson et al., 2023; Duan et al., 2023; Xu et al., 2022). By learning a time-dependent score function, a diffusion model defines a transport from a tractable prior distribution to a complex target distribution. However, training such models is expensive, particularly as modern datasets and architectures continue to grow. It is therefore often impractical to train a separate diffusion model for every new target distribution of interest.

One natural strategy is to perform additional training after pretraining, commonly referred to as fine-tuning (Domingo-Enrich et al., 2025; Han et al., 2025; Uehara et al., 2024; Liu et al., 2026). However, fine-tuning remains computationally expensive. This limitation has motivated inference-time scaling, which adapts a pretrained diffusion model to a new sampling task by allocating additional computation at inference time. One class of existing approaches consists of guidance-based methods, which modify the reverse dynamics using classifier guidance (Dhariwal & Nichol, 2021), classifier-free guidance (Ho & Salimans, 2021), or many variants (He et al., 2024; Guo et al., 2024; Li et al., 2026a). More recently, probabilistic inference methods have gained attention as a principled approach to inference-time scaling by casting the problem within Bayesian inference and Monte Carlo sampling frameworks (Chung et al., 2022; Bruna & Han, 2024), with particular emphasis on Sequential Monte Carlo (SMC) methods (Wu et al., 2023; Ren et al., 2026; Wang et al., 2026). In parallel, energy-based diffusion models have been developed (Yu et al., 2026; OuYang et al., 2026; Guth et al., 2026), whose energy representations also support compositional generation and control with Monte Carlo methods (Du et al., 2023; Thornton et al., 2025).

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.

Despite their theoretical grounding, these methods commonly assume the pretrained diffusion model is perfect (Ren et al., 2026; He et al., 2026; Wang et al., 2026). In practice, finite data, limited capacity, and imperfect optimization make the learned score only approximate, so the model's marginals deviate from the true ones. Because existing methods treat these approximate marginals as perfect, they inherit this error: adding particles reduces Monte Carlo error but cannot remove the pretrained model error itself. The error is especially pronounced in out-of-distribution regions, where the learned score is less reliable yet the model is often deployed (Lewis et al., 2025). Tellingly, Soiffer et al. (2026) show that in this regime it is score estimation error, not the vanilla inference-time approximation error, that dominates. Inference-time reweighting derived under an exact-model assumption can therefore leave substantial pretrained model error uncorrected, motivating our goal of correcting this error.

In this work, we introduce error-corrected inference-time scaling, realized by the *Energy-based Feynman–Kac Corrector (EBFKC)*, a framework that corrects pretrained model error on the fly given a reference energy. The reference energy specifies which distribution to sample but not how to explore it: on high-dimensional, multimodal targets, Monte Carlo sampling from the energy alone mixes slowly across modes, whereas a pretrained diffusion model already encodes a learned transport that reaches them efficiently. We first derive the full Feynman–Kac reweighting potential for any prescribed measure transport path, without assuming that the underlying score satisfies the base diffusion equation. We approximate the resulting dynamics with SMC and use variance-controlling guidance to mitigate particle-weight degeneracy. Then we construct an energy-aligned path that uses a pretrained energy-based diffusion model as a surrogate along the diffusion path while progressively correcting the discrepancy between the learned and target energies to remove the endpoint mismatch. To our knowledge, EBFKC is the *first* inference-time scaling method that corrects the path-tracking error of an imperfect pretrained model and, treating the pretrained energy as an approximate surrogate, provably realigns the terminal distribution to the target energy, rather than composing or reweighting several assumed-exact models.

Our Contributions. Building on this insight, our work makes the following contributions:

• **Energy-based Feynman–Kac Corrector (EBFKC).** We derive the Feynman–Kac–type Fokker–Planck equation for a general prescribed probability path, freeing it from the model's own marginals, and adapt the Variance-Controlling SMC (Ren et al., 2026) to track it while controlling particle-weight variance. Unlike the Feynman–Kac Corrector (Skreta et al., 2025), EBFKC tracks the path exactly at the population level even when the pretrained model is imperfect.

• **Progressive target-energy correction.** Instantiating this freedom, we keep the imperfect pretrained energy as a surrogate proposal while using the reference energy to pin the terminal distribution to the desired energy-aligned target (EAT), removing the endpoint mismatch that increasing the number of particles alone cannot eliminate. We call the resulting method *EBFKC on EAT*.

• **Error correction across benchmarks.** Across Gaussian mixture models, particle systems, alanine dipeptide, and alanine tetrapeptide, we show that a moderately imperfect pretrained model can severely bias existing inference-time scaling methods, and that EBFKC on EAT corrects this bias at inference time, recovering target distributions and molecular free-energy profiles where these baselines fail.

2 Background and Problem Formulation

We review score- and energy-based diffusion models, then formulate inference-time scaling and explain how pretrained model error affects both the terminal target and path tracking.

2.1 Score- and Energy-Based Diffusion Models

Diffusion models (Sohl-Dickstein et al., 2015; Ho et al., 2020; Song et al., 2021) connect the data distribution $p^\star$ to a tractable reference ρ through a forward noising diffusion

$$\mathrm{d}\mathbf{y}_s = u_s(\mathbf{y}_s)\,\mathrm{d}s + \beta_s\,\mathrm{d}\mathbf{w}_s, \qquad \mathbf{y}_0 \sim p^\star, \qquad s \in [0, T]. \tag{1}$$

The coefficients u_s, β_s are chosen so that the process carries $\mathbf{y}_0 \sim p^\star$ to $\mathbf{y}_T \sim \rho$ (typically very close to a standard Gaussian); we take the drift linear, $u_s(\mathbf{y}) = -F_s\mathbf{y}$, so the forward transition

kernels are Gaussian in closed form. Generation runs a reverse clock $t = T - s$; let $p_t^\star$ denote the law of $\mathbf{y}_{T-t}$, with $\widetilde{u}_t := u_{T-t}$ and $\sigma_t := \beta_{T-t}$ (so $p_0^\star = \rho$ and $p_T^\star = p^\star$). The exact reverse-time drift generating marginals $(p_t^\star)_{t\in[0,T]}$ is

$$v_t^\star(\mathbf{x}) = -\widetilde{u}_t(\mathbf{x}) + \sigma_t^2 s_t^\star(\mathbf{x}), \qquad s_t^\star(\mathbf{x}) := \nabla_{\mathbf{x}} \log p_t^\star(\mathbf{x}). \tag{2}$$

Since $p_t^\star$ is unavailable, diffusion models learn a score model $s_\theta(\mathbf{x}, t) \approx s_t^\star(\mathbf{x})$ from samples of (1) (Ho et al., 2020; Song et al., 2021); we assume such a pretrained model is available. Substituting s_θ for $s_t^\star$ gives the learned drift $v_t^\theta := -\widetilde{u}_t + \sigma_t^2 s_\theta$ and the generative diffusion

$$\mathrm{d}\mathbf{x}_t = v_t^\theta(\mathbf{x}_t)\,\mathrm{d}t + \sigma_t\,\mathrm{d}\mathbf{w}_t, \qquad \mathbf{x}_0 \sim \rho, \qquad t \in [0, T], \tag{3}$$

whose law we denote μ_t. If $s_\theta = s_t^\star$ for all $(\mathbf{x}, t)$, the generative and exact reverse diffusions share the same drift and initial law ρ, so $\mu_t = p_t^\star$ for all t and in particular $\mu_T = p^\star$. When s_θ approximates $s_t^\star$, μ_T is only an approximation of $p^\star$.

The score model s_θ is often parameterized directly, as in standard score-based models. We instead adopt an *energy-based* parameterization (Salimans & Ho, 2021): a scalar energy $E_\theta(\mathbf{x}, t)$ defines $p_t^\theta(\mathbf{x}) = \exp(-E_\theta(\mathbf{x}, t))/Z_t^\theta$ (with $Z_t^\theta := \int \exp(-E_\theta(\mathbf{x}, t))\,\mathrm{d}\mathbf{x}$), and the score follows as $s_\theta = -\nabla_{\mathbf{x}} E_\theta = \nabla \log p_t^\theta$. Unlike a directly parameterized score, which is a general vector field with no potential and hence no associated density, this yields explicit intermediate densities $(p_t^\theta)_{t\in[0,T]}$; they need not coincide with the reverse-SDE law μ_t, since fixing the spatial score does not enforce the correct evolution in time, nor with the truth $p_t^\star$. They further underlie energy-based diffusion's strength in compositional generation (Du et al., 2023; Thornton et al., 2025), unnormalized density evaluation (Yu et al., 2025; Guth et al., 2026), and downstream reweighting (OuYang et al., 2026).

Starting point: energy-based diffusion. Its explicit densities p_t^θ, unavailable from a bare score, later let us realign the target and reweight along the prescribed path.

2.2 Inference-Time Scaling with Imperfect Models

Inference-time scaling adapts a pretrained diffusion model to a new sampling target without retraining. In full generality, the goal is to sample from

$$q_T^{\mathrm{ref}}(\mathbf{x}) \propto \left(p^{\mathrm{ref}}(\mathbf{x})\right)^\gamma \exp\left(r(\mathbf{x})\right), \qquad p^{\mathrm{ref}}(\mathbf{x}) \propto \exp\left(-E^{\mathrm{ref}}(\mathbf{x})\right), \tag{4}$$

for an annealing exponent $\gamma > 0$, a reward r, and a time-independent *reference energy* E^{ref} ($r = 0$ gives the *annealing* task, $\gamma = 1$ the *reward-tilting* task). The pretrained model enters because its densities $(p_t^\theta)_{t\in[0,T]}$ are expected to be close to the reference, making it a far better starting point than sampling from E^{ref} directly. The canonical choice is $p^{\mathrm{ref}} = p^\star$, recovering

$$q_T^\star(\mathbf{x}) \propto (p^\star(\mathbf{x}))^\gamma \exp\left(r(\mathbf{x})\right). \tag{5}$$

Since each p_t^θ approximates $p_t^\star$, a natural construction interpolates through the model's own densities (Skreta et al., 2025; Ren et al., 2026; Wang et al., 2026; He et al., 2026).

Definition 1 (Model-Induced Target)**.** *The* model-induced target (MIT) *is*

$$q_T^{\mathrm{MIT}}(\mathbf{x}) \propto (p_T^\theta(\mathbf{x}))^\gamma \exp(r(\mathbf{x})),$$

which is (4) *with* $E^{\mathrm{ref}} = E_\theta(\cdot, T)$*. It is the terminal distribution of the model-induced path*

$$q_t^{\mathrm{MIT}}(\mathbf{x}) \propto (p_t^\theta(\mathbf{x}))^\gamma \exp(r_t(\mathbf{x})),$$

where the intermediate reward r_t *interpolates between* $r_0 = 0$ *and* $r_T = r$*.*

This construction, however, is undermined by an imperfect pretrained model in two distinct ways.

Two failure modes under an imperfect model. *Endpoint mismatch:* the MIT can differ from the desired q_T^{ref}, even if its path is tracked exactly. *Path-tracking error:* existing reweighting presumes a consistent density p_t^θ that need not even exist, so it fails to track even the model-induced path.

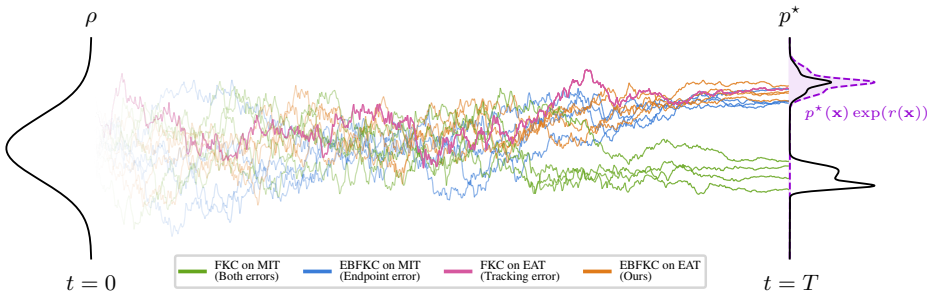


Figure 1: SMC trajectory lineages for the four combinations of reweighting scheme (FKC or EBFKC) and prescribed path (MIT or EAT). The target $p^\star$ is a 64-dimensional mixture of six isotropic Gaussians in two groups (right edge). The figure displays one-dimensional trajectories for four lineages from a population of 2048 particles. The reward r favors the weak mode, with the desired terminal distribution shown in dashed purple. FKC (Skreta et al., 2025) on MIT suffers from both errors and ends in the wrong modes. EBFKC on MIT tracks its path but inherits the endpoint mismatch, while FKC on EAT has the correct endpoint but its path-tracking error makes the weights degenerate, so that only one lineage survives. Only EBFKC on EAT, which removes both errors, reaches the desired target.

Endpoint mismatch. Whether this gap matters depends on the reference energy. In data-driven settings such as image generation, the goal is to steer the model's own distribution, so $E^{\mathrm{ref}} = E_\theta(\cdot, T)$ and the MIT is exactly the desired target. In most scientific computing tasks, a physical model supplies an external $E^\star$ and the quantities of interest are defined with respect to $p^\star \propto \exp(-E^\star)$, so the target is $E^{\mathrm{ref}} = E^\star$; here the mismatch must be corrected, yet existing steering methods for pretrained diffusion models are not designed to use a known $E^\star$ for this purpose. More broadly, the same holds whenever the target is defined with the corresponding reference energy $E^{\mathrm{ref}} \neq E_\theta(\cdot, T)$, for which the MIT differs from the desired q_T^{ref}.

Path-tracking error. Concretely, to obtain efficient reweighting formulas that use only the learned score, most existing methods implicitly assume that such a density p_t^θ exists and satisfies the Fokker–Planck equation associated with (1) (Skreta et al., 2025). This assumption generally fails for a learned, imperfect score, so the resulting reweighting formulas are biased and no longer track q_t^{MIT}, even at the population level.

3 Method

Both issues above point to the same remedy: decouple the sampling path from the model's own marginals. We therefore track a prescribed path with a general time-dependent potential Φ_t,

$$q_t(\mathbf{x}) = \exp\left(-\Phi_t(\mathbf{x})\right)/Z_t, \qquad t \in [0, T], \tag{6}$$

requiring only that the path be normalizable, that q_0 admit easy sampling, and that Φ_t and its required derivatives be evaluable; Z_t need not be known. The model-induced path is recovered with $\Phi_t^{\mathrm{MIT}}(\mathbf{x}) = \gamma E_\theta(\mathbf{x}, t) - r_t(\mathbf{x})$.

We then resolve both failure modes with two ingredients, whose effect Fig. 1 illustrates:

> *Energy-based Feynman–Kac Corrector (EBFKC)* (Sec. 3.1) tracks any prescribed path exactly at the population level, removing the path-tracking error. *Energy-aligned target (EAT)* (Sec. 3.2) pins the terminal distribution to q_T^{ref} for any E^{ref}, removing the endpoint mismatch.

Sec. 3.3 further relates our method to annealed SMC and energy-based modeling.

3.1 Fokker–Planck Equation for a Prescribed Path

For the prescribed path in Eq. (6), we derive tracking dynamics directly from its potential Φ_t, without requiring diffusion consistency. The result applies to any sufficiently smooth, strictly positive, normalizable path, including paths to the learned terminal density or to an externally specified target.

Theorem 2 (Fokker–Planck Equation for a Prescribed Path). *Assume that Φ_t is C^1 in t and C^2 in $\mathbf{x}$, that $Z_t < \infty$, and that $\widetilde{u}_t$ is C^1 in $\mathbf{x}$. For the prescribed path q_t in Eq. (6), we define the proposal drift $v_t(\mathbf{x}) = -\widetilde{u}_t(\mathbf{x}) - \sigma_t^2 \nabla \Phi_t(\mathbf{x})$, then q_t satisfies the Feynman–Kac-type Fokker–Planck equation*

$$\partial_t q_t = -\nabla \cdot (q_t v_t) + \frac{\sigma_t^2}{2} \Delta q_t + q_t g_t, \tag{7}$$

where the centered reweighting potential is $g_t(\mathbf{x}) = G_t(\mathbf{x}) - \mathbb{E}_{q_t}[G_t]$, and the corresponding uncentered potential is

$$G_t(\mathbf{x}) = -\partial_t \Phi_t(\mathbf{x}) - \nabla \cdot \widetilde{u}_t(\mathbf{x}) + \widetilde{u}_t(\mathbf{x})^\top \nabla \Phi_t(\mathbf{x}) - \frac{\sigma_t^2}{2} \Delta \Phi_t(\mathbf{x}) + \frac{\sigma_t^2}{2} \left\| \nabla \Phi_t(\mathbf{x}) \right\|^2 . \tag{8}$$

The derivation is provided in App. B.1. The state-dependent potential G_t corrects the discrepancy between the proposal diffusion and the prescribed evolution of q_t, while centering G_t by its expectation under q_t preserves normalization and eliminates the unknown derivative $\partial_t \log Z_t$.

We simulate Eq. (7) with Sequential Monte Carlo (SMC). Let $\{(\mathbf{x}_t^{(i)}, w_t^{(i)})\}_{i=1}^N$ denote a system of normalized weighted particles, initialized independently as $\mathbf{x}_0^{(i)} \sim q_0 \propto \exp(-\Phi_0)$ with $w_0^{(i)} = 1/N$. The particles then evolve according to the continuous-time dynamics

$$\begin{cases} \mathrm{d}\mathbf{x}_t^{(i)} = v_t(\mathbf{x}_t^{(i)})\mathrm{d}t + \sigma_t \mathrm{d}\mathbf{w}_t^{(i)}, & i \in [N], \\ \mathrm{d}\log w_t^{(i)} = \hat{g}_t(\mathbf{x}_t^{(i)})\mathrm{d}t, & i \in [N], \\ \hat{g}_t(\mathbf{x}_t^{(i)}) := G_t(\mathbf{x}_t^{(i)}) - \sum_{j=1}^N w_t^{(j)} G_t(\mathbf{x}_t^{(j)}), & i \in [N], \end{cases} \tag{9}$$

In this way, SMC approximates this normalized Feynman–Kac flow using a weighted particle system, with optional resampling to control weight degeneracy. Simulating Eq. (9) from $t = 0$ to $t = T$ therefore yields a weighted particle approximation of the target distribution q_T. To further mitigate particle weight degeneracy during SMC simulation, we also incorporate Variance-Controlling Guidance (VCG) (Ren et al., 2026), which modifies the particle drift in a principled manner; details are in App. B.2. We call this framework the **Energy-based Feynman–Kac Corrector (EBFKC)**. We remark that FKC (Skreta et al., 2025) follows the same weighted-particle scheme as Eq. (9) using only the learned score, but its reweighting potential is valid only when p_t^θ is consistent with the base diffusion (satisfies its Fokker–Planck equation). Our EBFKC instead evaluates the full potential G_t, which requires the energy E_θ rather than the score alone; in exchange it makes no such assumption and tracks the path q_t exactly in the continuous-time population limit even when the pretrained model is imperfect. We also extend FKC to general time-dependent rewards in App. B.5.

3.2 From Model-Induced to Energy-Aligned Target

The prescribed-path dynamics provided by EBFKC separate the choice of target from the assumption that the model follows an exact diffusion. Applying EBFKC to the model-induced path in Definition 1 gives **EBFKC on MIT**. We now construct a path ending at a general target q_T^{ref}, with the model-induced path recovered with the special choice $E^{\mathrm{ref}} = E_\theta(\cdot, T)$.

Energy-Aligned Target. To reach the reference target rather than the model-induced target, we constrain the terminal distribution of the path.

Definition 3 (Energy-Aligned Target). *The* energy-aligned target (EAT) *is* $q_T^{\mathrm{EAT}}(\mathbf{x}) \propto \exp\left(-\Phi_T(\mathbf{x})\right)$ *with*

$$\Phi_T(\mathbf{x}) = \gamma E^{\mathrm{ref}}(\mathbf{x}) - r(\mathbf{x}) + c_T,$$

where c_T is independent of $\mathbf{x}$.

Energy-aligned target constrains only the terminal distribution: the intermediate distributions can be constructed flexibly, provided that the resulting path terminates at q_T^{ref}. We retain the model-induced path and progressively incorporate the terminal energy correction $\Delta E^{\mathrm{ref}}(\mathbf{x}) = E^{\mathrm{ref}}(\mathbf{x}) - E_\theta(\mathbf{x}, T)$. For a smooth correction schedule κ_t with $\kappa_0 = 0$ and $\kappa_T = 1$, we instantiate the energy-aligned

path through the potential

$$\Phi_t^{\mathrm{EAT}}(\mathbf{x}) = \Phi_t^{\mathrm{MIT}}(\mathbf{x}) + \gamma\kappa_t \Delta E^{\mathrm{ref}}(\mathbf{x}) = \gamma E_\theta(\mathbf{x}, t) - r_t(\mathbf{x}) + \gamma\kappa_t \Delta E^{\mathrm{ref}}(\mathbf{x}).$$

When $E^{\mathrm{ref}} = E_\theta(\cdot, T)$, the correction $\Delta E^{\mathrm{ref}}(\mathbf{x})$ vanishes and EAT reduces to the MIT; and when $E^{\mathrm{ref}} = E^\star$, the correction closes the terminal energy mismatch and the path terminates at $q_T^\star$. We also introduce a look-ahead map to stabilize target-energy evaluation, with details in App. B.3. Building on the energy-aligned path, we denote the resulting method **EBFKC on EAT** and summarize it in Alg. 1.

3.3 Connections to Annealed SMC and Energy-Based Modeling

Our method can be viewed as annealed SMC along a diffusion-based probability path. Traditional annealed SMC typically uses a geometric tempering path between the initial and target densities. Its intermediate densities can be evaluated up to normalization, but transporting particles between separated modes can remain difficult (Máté & Fleuret, 2023). We describe the geometric tempering SMC baseline in App. B.6 and compare it with EBFKC on EAT in Lennard-Jones systems (Sec. 4.2).

Existing work replaces the tempering path with a diffusion path, whose progressive Gaussian smoothing eases transport between modes, and uses it to build Boltzmann Generators, i.e., samplers for $p^\star(\mathbf{x}) \propto \exp(-E^\star(\mathbf{x}))$ with $E^\star$ given (Phillips et al., 2024; Grenioux & Noble, 2026). These methods use a reweighting strategy similar to that in Thm. 2, specialized to $\Phi_t = E_\theta(\cdot, t)$, where the architecture of E_θ enforces $E_\theta(\cdot, T) = E^\star$ by construction. We generalize this to an arbitrary reference energy, sampling from q_T^{ref} rather than from $p^\star$ alone, and without requiring the terminal constraint to be baked into the architecture.

This perspective also motivates our choice of building energy-based rather than score-based diffusion models. Reweighting requires evaluating the density along the path, which an energy model provides directly. With only a score, the density must be recovered by integrating the score for every particle at every step, which is costly and adds numerical error (Sjöberg et al., 2026; Lam et al., 2026), making efficient and reliable reweighting impractical.

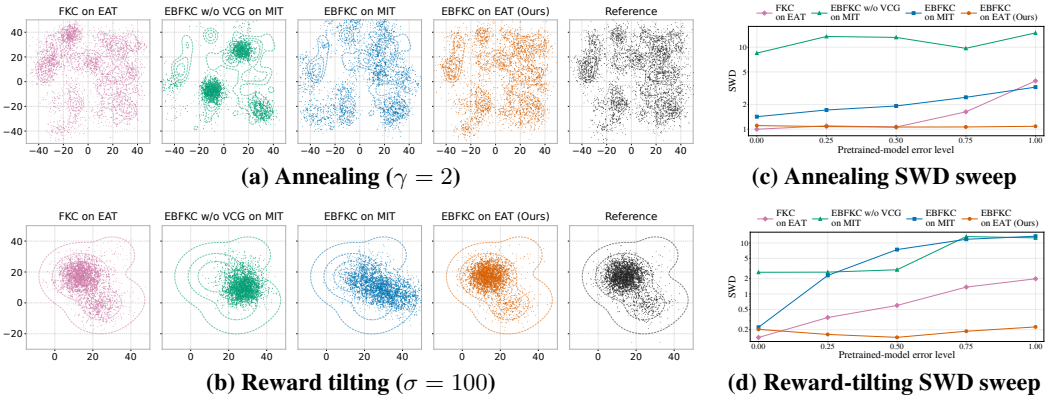
4 Experiments

We evaluate whether EBFKC on EAT recovers the reference target despite pretrained model error, on annealing and reward-tilting tasks for a high-dimensional Gaussian mixture model, particle systems, alanine dipeptide, and alanine tetrapeptide. We take $E^{\mathrm{ref}} = E^\star$, the scientific setting that motivates our method, for which reference samples are available independently of the pretrained model. We compare against two baselines that each remove only one failure mode: (1) EBFKC on MIT tracks its path correctly but terminates at the MIT, isolating the endpoint mismatch; (2) FKC on EAT runs FKC with the reward redefined as $\hat{r} = r - \gamma\Delta E^{\mathrm{ref}}$, which moves its endpoint to $q_T^\star$ (App. B.5), isolating the path-tracking error. Across all systems, only EBFKC on EAT recovers the reference, showing that both failure modes arise in practice and must be addressed. To quantify the computational cost of these accuracy gains, we compare runtimes at the reported sampling budgets. EBFKC on EAT takes approximately 1.1–3.6$\times$ the runtime of EBFKC on MIT in these comparisons, owing to additional energy and derivative evaluations (App. D.1–D.5). Unless marked "w/o VCG", all methods use VCG (Ren et al., 2026), which we found essential for stable weight dynamics in both FKC and EBFKC. Energy models for particle and molecular systems are trained with Diffusive Classification (OuYang et al., 2026) (App. A).

4.1 Gaussian Mixture Model

We use a 30-dimensional Gaussian mixture model (GMM) to isolate the effect of pretrained model error in a fully controlled setting. The marginals $p_t^\star(\mathbf{x})$ and their energies $E_t^\star(\mathbf{x}) = -\log p_t^\star(\mathbf{x})$ are available in closed form, which allows us to emulate an imperfect pretrained diffusion model analytically by perturbing the component means and variances. Varying the perturbation magnitudes allows us to systematically evaluate how sampling quality depends on pretrained model error. Complete definitions are given in App. C.1.

Both failure modes caused by the imperfect pretrained model appear in the annealing and reward-tilting tasks (Fig. 2(a,b)). Although the EAT path terminates at the desired target, FKC on EAT



(a) **Annealing** ($\gamma = 2$) (c) **Annealing SWD sweep**

(b) **Reward tilting** ($\sigma = 100$) (d) **Reward-tilting SWD sweep**

Figure 2: GMM results under controlled pretrained model error. (a–b) Sample distributions and (c–d) SWD under increasing pretrained model error for annealing and reward tilting.

Table 1: Quantitative comparison on the 30-dimensional GMM under the controlled-error setting $\sigma_\mu = 10$ and $\rho = 0.1$. Results are mean $\pm$ standard deviation over 5 runs.

Task	Method	$\lvert\Delta\text{NLL}\rvert \downarrow$	MMD $\downarrow$	SWD $\downarrow$	Mean $L_2 \downarrow$	Cov $F \downarrow$
Annealing	FKC on EAT	1.043 ± 0.196	0.076 ± 0.007	4.105 ± 0.375	21.218 ± 1.462	2437.100 ± 358.582
	EBFKC w/o VCG on MIT	59.355 ± 9.014	0.239 ± 0.035	11.913 ± 1.374	63.396 ± 13.044	6770.604 ± 1682.698
	EBFKC on MIT	59.204 ± 1.041	0.092 ± 0.002	3.359 ± 0.694	9.890 ± 0.385	1879.349 ± 129.679
	EBFKC on EAT (Ours)	$\mathbf{0.133 \pm 0.095}$	$\mathbf{0.022 \pm 0.001}$	$\mathbf{1.110 \pm 0.331}$	$\mathbf{5.493 \pm 0.774}$	$\mathbf{639.193 \pm 27.982}$
Reward tilting	FKC on EAT	0.928 ± 0.374	0.061 ± 0.019	1.313 ± 1.191	8.661 ± 7.779	507.355 ± 680.777
	EBFKC w/o VCG on MIT	21.575 ± 10.209	0.435 ± 0.040	10.136 ± 6.044	64.321 ± 29.173	604.058 ± 878.509
	EBFKC on MIT	22.875 ± 6.175	0.383 ± 0.048	10.287 ± 5.792	63.261 ± 25.670	1027.872 ± 1435.608
	EBFKC on EAT (Ours)	$\mathbf{0.113 \pm 0.016}$	$\mathbf{0.018 \pm 0.008}$	$\mathbf{0.194 \pm 0.102}$	$\mathbf{1.000 \pm 0.684}$	$\mathbf{50.068 \pm 32.847}$

remains affected by path-tracking error and fails to sample accurately from that target. EBFKC on MIT corrects path tracking, but its samples still deviate from the reference because of endpoint mismatch, and without VCG it also suffers from severe weight degeneracy. Switching from MIT to EAT further removes the endpoint mismatch, allowing EBFKC on EAT to sample accurately from the desired target.

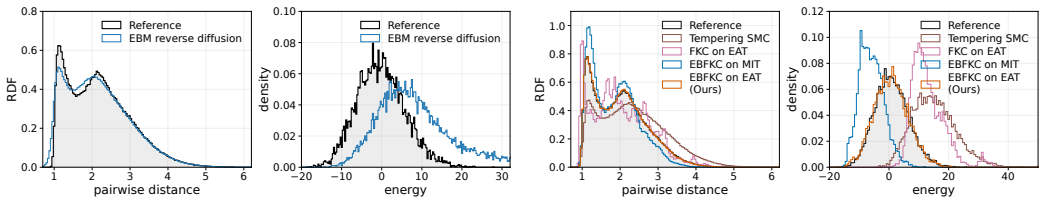
We further evaluate sampling accuracy as the pretrained model error increases. Fig. 2(c,d) reports the sliced-Wasserstein distance (SWD) for annealing and reward tilting, respectively. EBFKC on EAT maintains a low SWD across all error levels in both tasks, whereas EBFKC on MIT and FKC on EAT deviate increasingly from the reference as the model error grows. As shown in App. D.1, FKC w/o VCG on EAT suffers from severe weight degeneracy, indicating the importance of VCG in stabilizing FKC. Table 1 reports the quantitative comparison at a fixed error level, where EBFKC on EAT has the lowest mean error on all five reported metrics for both tasks.

4.2 Particle Systems

We next test whether the correction observed with analytically perturbed GMMs extends to energy models trained on finite data. We consider the Double-Well (DW) and Lennard-Jones (LJ) systems, whose time-dependent energies are no longer available in closed form. We train energy-based diffusion models with limited training steps to obtain imperfect models and evaluate their inference-time adaptation to annealing and reward tilting. Detailed experimental settings are provided in App. C.2.

Double-Well system. We first study a 2D 4-particle Double-Well system. Figures 9 and 10 in App. D.2 compare the radial distribution functions (RDFs) and energy distributions for annealing and reward tilting. In both tasks, FKC on EAT and EBFKC on MIT produce visibly shifted RDFs and energy distributions, indicating that these methods retain pretrained model error. In contrast, EBFKC on EAT closely matches the reference pairwise-distance and energy distributions under both sampling objectives.

Lennard-Jones system. We next consider the more challenging 3D 13-particle LJ system, whose potential has a rugged energy landscape and exhibits singular short-range behavior. As shown in



(a) Pretrained model reverse diffusion results **(b) Reward tilting task results**

Figure 3: LJ results comparing pairwise-distance and energy distributions with the reference. (a) Samples from the pretrained model. (b) Reward-tilting samples from the inference-time methods.

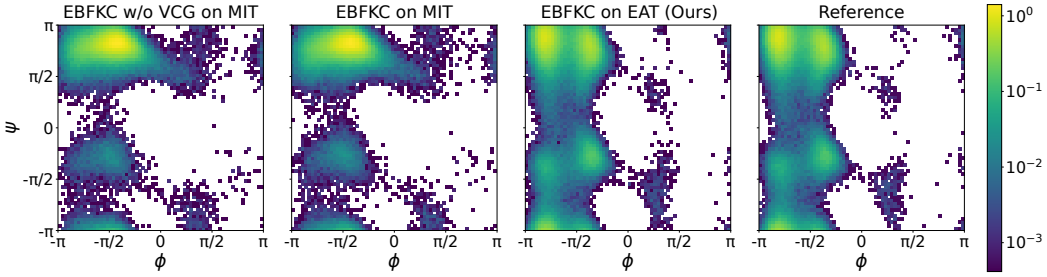


Figure 4: Ramachandran plots of the backbone dihedral angles (ϕ, ψ) for annealing ALDP from 800 K to 300 K. Reference samples at 300 K are compared with samples generated by EBFKC on MIT with/without VCG and EBFKC on EAT using the same pretrained diffusion model at 800 K.

Fig. 3(a), the pretrained model visibly deviates from the reference. For reward tilting (Fig. 3(b)), FKC on EAT and EBFKC on MIT inherit this error, whereas EBFKC on EAT closely matches the reference distributions. Geometric tempering SMC (App. B.6), a standard SMC method working without the pretrained model, also fails to sample the target accurately with the same number of particles, illustrating the benefit of the diffusion path (Sec. 3.3). App. D.3 reports the annealing results. Table 7 further confirms that EBFKC on EAT corrects the pretrained model error.

4.3 Alanine Dipeptide Annealing

We next evaluate our method on alanine dipeptide (ALDP) annealing tasks, a standard benchmark for molecular conformational sampling. Its slow conformational degrees of freedom are characterized by the backbone dihedral angles ϕ and ψ. We train an energy-based diffusion model at 800 K (Fig. 12 in App. D.4) and anneal it to 300 K via inference-time scaling.

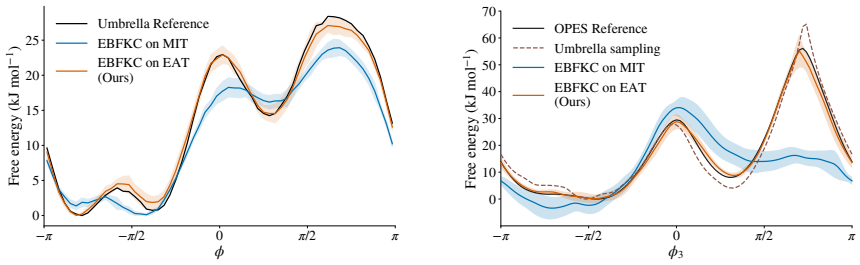
Fig. 4 compares the Ramachandran plots of the annealed samples. EBFKC on EAT recovers both the conformational basins and their relative populations, whereas EBFKC on MIT, with or without VCG, deviates substantially from the 300 K reference because it anneals toward the MIT rather than the 300 K Boltzmann distribution. The marginal distributions of ϕ and ψ in Fig. 13 support this observation. Table 2 quantifies the comparison: despite the error of the model pretrained at 800 K, EBFKC on EAT closely matches the reference across all metrics. FKC was numerically unstable on both paths in our ALDP runs, so we omit it. The look-ahead ablation study in App. D.4 shows that removing this map leads to numerical instability, demonstrating its importance in difficult tasks.

Table 2: Quantitative ALDP annealing results from 800 K to 300 K; the first row evaluates the pretrained model at 800 K against the 800 K reference. Results are mean ± std over 5 runs.

Methods	Temperature (K)	(ϕ, ψ) TVD ↓	Basin-pop. error ↓	Energy TVD ↓	Distance TVD ↓
Pretrained reverse diffusion	800	0.3005 ± 0.0022	0.1699 ± 0.0016	0.6730 ± 0.0013	0.0253 ± 0.0001
EBFKC w/o VCG on MIT	300	0.5089 ± 0.0163	0.1643 ± 0.0037	0.4460 ± 0.0130	0.0366 ± 0.0008
EBFKC on MIT	300	0.5063 ± 0.0035	0.1657 ± 0.0026	0.4659 ± 0.0095	0.0381 ± 0.0006
EBFKC on EAT (Ours)	300	$\mathbf{0.0855 \pm 0.0031}$	$\mathbf{0.0047 \pm 0.0041}$	$\mathbf{0.0240 \pm 0.0028}$	$\mathbf{0.0030 \pm 0.0002}$

4.4 Free-Energy Estimation

Free-energy profiles characterize the relative thermodynamic stability of molecular conformations along selected reaction coordinates. Their accurate estimation remains a rare-event sampling problem because low-probability conformations are infrequently observed even when a pretrained equilibrium sampler can already generate independent configurations.



(a) ALDP free-energy profile **(b) Ala-4 free-energy profile**

Figure 5: One-dimensional molecular free-energy profiles at 300 K. (a) ALDP along ϕ. (b) Ala-4 along ϕ_3. The references are umbrella sampling for ALDP and OPES for Ala-4. Diffusion-based results show the mean and standard deviation (shaded) over five independent runs.

Motivated by this challenge, we follow UmbrellaDiff (Xie et al., 2026) to estimate one-dimensional free-energy profiles at 300 K for ALDP along ϕ and for the 42-atom alanine tetrapeptide (Ala-4) along its third backbone dihedral ϕ_3, starting from pretrained diffusion equilibrium samplers (App. D.5). Each umbrella window defines a reward-tilting target with $E^{\mathrm{ref}} = E^{\star}$, which we sample with EBFKC on EAT or EBFKC on MIT, and the weighted particles from all windows are combined with the weighted multistate Bennett acceptance ratio (MBAR) (Shirts & Chodera, 2008; Xie et al., 2026).

Alanine dipeptide. As shown in Fig. 5(a), EBFKC on MIT underestimates the principal barriers because its window-wise samples retain pretrained model error. In contrast, EBFKC on EAT closely follows the umbrella sampling reference (Torrie & Valleau, 1977) and reproduces its principal barriers and basin structure, indicating that it samples the reward-tilted target in each umbrella window accurately.

Alanine tetrapeptide. Ala-4 poses a high-barrier sampling problem: its three slow backbone dihedrals partition the conformational space into eight metastable basins (settings in App. C.4). Fig. 5(b) uses as reference OPES (Invernizzi & Parrinello, 2020), which biases all three slow dihedrals ϕ_1, ϕ_2, and ϕ_3. Traditional umbrella sampling, which biases only ϕ_3, becomes trapped along the hidden slow variables and yields an inaccurate profile. EBFKC on EAT, although it also biases only ϕ_3, closely reconstructs the OPES basin-and-barrier structure, whereas EBFKC on MIT retains the pretrained model error and severely distorts the barrier heights. Without VCG, the profile of EBFKC on MIT is distorted even further (App. D.5).

5 Conclusion

We introduced the *Energy-based Feynman–Kac Corrector (EBFKC)*, a framework for inference-time scaling with imperfect pretrained energy-based diffusion models given a reference energy. EBFKC derives Feynman–Kac dynamics that track any prescribed path exactly at the population level and simulates them with SMC, using VCG to mitigate weight degeneracy, thereby removing the path-tracking error. The energy-aligned target (EAT) uses the pretrained energy as a surrogate along the path and progressively incorporates the reference energy, removing the endpoint mismatch. Across Gaussian mixture models, particle systems, and molecular systems, both ingredients prove necessary, and EBFKC on EAT recovers the target distributions under annealing and reward tilting and yields accurate molecular free-energy profiles. Future directions include adapting the correction schedules automatically, which we currently design by hand for the high-barrier alanine tetrapeptide system, reducing derivative-evaluation costs, and correcting large pretrained equilibrium samplers for more reliable inference-time scaling in challenging real-world applications (Lewis et al., 2025).

AI USE STATEMENT

In this work, we have not used generative AI tools for any tasks with required disclosure. We used generative AI tools to polish the writing and to assist with coding. We have reviewed all AI-assisted work: text edits were checked by the authors, and AI-assisted code was reviewed and tested for correctness. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

We provide the necessary details to facilitate reproduction of our results. The proposed Energy-based Feynman-Kac Corrector is described in detail in the main text, with complete algorithmic specifications and implementation details provided in the appendix. The assumptions and derivations underlying our theoretical results, together with complete proofs, are also included in the appendix. For all experiments, we report the experimental setups, model architectures, training procedures, sampling configurations, evaluation metrics, and hyperparameters, as well as details of the datasets and target distributions used. The code will be released later.

REFERENCES

Joan Bruna and Jiequn Han. Provable posterior sampling with denoising oracles via tilted transport. *Advances in Neural Information Processing Systems*, 37:82863–82894, 2024.

Hyungjin Chung, Jeongsol Kim, Michael Thompson Mccann, Marc Louis Klasky, and Jong Chul Ye. Diffusion posterior sampling for general noisy inverse problems. In *The Eleventh International Conference on Learning Representations*, 2022.

Prafulla Dhariwal and Alexander Nichol. Diffusion models beat gans on image synthesis. *Advances in neural information processing systems*, 34:8780–8794, 2021.

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 *The Thirteenth International Conference on Learning Representations*, 2025. URL .

Yilun Du, Conor Durkan, Robin Strudel, Joshua B Tenenbaum, Sander Dieleman, Rob Fergus, Jascha Sohl-Dickstein, Arnaud Doucet, and Will Sussman Grathwohl. Reduce, reuse, recycle: Compositional generation with energy-based diffusion models and MCMC. In *International conference on machine learning*, pp. 8489–8510. PMLR, 2023.

Chenru Duan, Yuanqi Du, Haojun Jia, and Heather J Kulik. Accurate transition state generation with an object-aware equivariant elementary reaction diffusion model. *Nature computational science*, 3(12):1045–1055, 2023.

Louis Grenioux and Maxence Noble. Diffusion-based annealed boltzmann generators : benefits, pitfalls and hopes. *Transactions on Machine Learning Research*, 2026. ISSN 2835-8856. URL .

Yingqing Guo, Hui Yuan, Yukang Yang, Minshuo Chen, and Mengdi Wang. Gradient guidance for diffusion models: An optimization perspective. *Advances in neural information processing systems*, 37:90736–90770, 2024.

Florentin Guth, Zahra Kadkhodaie, and Eero Simoncelli. Learning normalized image densities via dual score matching. *Advances in Neural Information Processing Systems*, 38:89796–89826, 2026.

Yinbin Han, Meisam Razaviyayn, and Renyuan Xu. Stochastic control for fine-tuning diffusion models: Optimality, regularity, and convergence. In *Forty-second International Conference on Machine Learning*, 2025. URL .

Jiajun He, José Miguel Hernández-Lobato, Yuanqi Du, and Francisco Vargas. RNE: plug-and-play diffusion inference-time control and energy-based training. In *The Fourteenth International Conference on Learning Representations*, 2026. URL .

Yutong He, Naoki Murata, Chieh-Hsin Lai, Yuhta Takida, Toshimitsu Uesaka, Dongjun Kim, WeiHsiang Liao, Yuki Mitsufuji, Zico Kolter, Ruslan Salakhutdinov, et al. Manifold preserving guided diffusion. In *International Conference on Learning Representations*, volume 2024, pp. 44819–44850, 2024.

Jonathan Ho and Tim Salimans. Classifier-free diffusion guidance. In *NeurIPS 2021 Workshop on Deep Generative Models and Downstream Applications*, 2021. URL .

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

Michael F Hutchinson. A stochastic estimator of the trace of the influence matrix for Laplacian smoothing splines. *Communications in Statistics-Simulation and Computation*, 18(3):1059–1076, 1989.

Michele Invernizzi and Michele Parrinello. Rethinking metadynamics: from bias potentials to probability distributions. *The journal of physical chemistry letters*, 11(7):2731–2736, 2020.

Michele Invernizzi, Andreas Krämer, Cecilia Clementi, and Frank Noé. Skipping the replica exchange ladder with normalizing flows. *The Journal of Physical Chemistry Letters*, 13(50):11643–11649, 2022. doi: 10.1021/acs.jpclett.2c03327.

Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-based generative models. *Advances in neural information processing systems*, 35:26565–26577, 2022.

Kevin H Lam, Tyler Farghly, Christopher Williams, Jun Yang, Yee Whye Teh, and Arnaud Doucet. Metropolis-adjusted diffusion models. *arXiv preprint arXiv:2605.09654*, 2026.

Sarah Lewis, Tim Hempel, José Jiménez-Luna, Michael Gastegger, Yu Xie, Andrew YK Foong, Victor García Satorras, Osama Abdin, Bastiaan S Veeling, Iryna Zaporozhets, et al. Scalable emulation of protein equilibrium ensembles with generative deep learning. *Science*, 389(6761): eadv9817, 2025.

Minhuan Li, Jiequn Han, Pilar Cossio, and Luhuan Wu. Robust inference-time steering of protein diffusion models via embedding optimization. *arXiv preprint arXiv:2602.05285*, 2026a. URL .

Tiancheng Li, Wentao Li, Anyang Peng, Jianming Xue, Linfeng Zhang, Duo Zhang, and Han Wang. Dpa4: Pushing the accuracy-cost frontier of interatomic potentials with emfa so (2) convolution. *arXiv preprint arXiv:2606.02419*, 2026b.

Jie Liu, Gongye Liu, Jiajun Liang, Yangguang Li, Jiaheng Liu, Xintao Wang, Pengfei Wan, Di Zhang, and Wanli Ouyang. Flow-grpo: Training flow matching models via online rl. *Advances in neural information processing systems*, 38:40783–40818, 2026.

Bálint Máté and François Fleuret. Learning interpolations between boltzmann densities. *Transactions on Machine Learning Research*, 2023. ISSN 2835-8856. URL .

Raphael A Meyer, Cameron Musco, Christopher Musco, and David P Woodruff. Hutch++: Optimal stochastic trace estimation. In *Symposium on Simplicity in Algorithms (SOSA)*, pp. 142–155. SIAM, 2021.

RuiKang OuYang, Louis Grenioux, and José Miguel Hernández-Lobato. A diffusive classification loss for learning energy-based generative models. In *Forty-third International Conference on Machine Learning*, 2026. URL .

William Peebles and Saining Xie. Scalable diffusion models with transformers. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 4195–4205, 2023.

Angus Phillips, Hai-Dang Dau, Michael John Hutchinson, Valentin De Bortoli, George Deligiannidis, and Arnaud Doucet. Particle denoising diffusion sampler. volume 235 of *Proceedings of Machine Learning Research*, pp. 40688–40724. PMLR, 21–27 Jul 2024.

Yinuo Ren, Wenhao Gao, Lexing Ying, Grant M. Rotskoff, and Jiequn Han. Driftlite: Lightweight drift control for inference-time scaling of diffusion models. In *International Conference on Learning Representations (ICLR)*, 2026.

Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 10684–10695, 2022.

Amirmojtaba Sabour, Michael S Albergo, Carles Domingo-Enrich, Nicholas M Boffi, Sanja Fidler, Karsten Kreis, and Eric Vanden-Eijnden. Test-time scaling of diffusions with flow maps. *arXiv preprint arXiv:2511.22688*, 2025.

Tim Salimans and Jonathan Ho. Should EBMs model the energy or the score? In *Energy Based Models Workshop-ICLR 2021*, 2021.

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

Anders Sjöberg, Jakob Lindqvist, Magnus Önnheim, Mats Jirstrand, and Lennart Svensson. Mcmc correction of score-based diffusion models for model composition. *Entropy*, 28(3), 2026. ISSN 1099-4300. doi: 10.3390/e28030351. URL .

Marta Skreta, Tara Akhound-Sadegh, Viktor Ohanesian, Roberto Bondesan, Alan Aspuru-Guzik, Arnaud Doucet, Rob Brekelmans, Alexander Tong, and Kirill Neklyudov. Feynman-kac correctors in diffusion: Annealing, guidance, and product of experts. In *Forty-second International Conference on Machine Learning*, 2025. URL .

Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *International conference on machine learning*, pp. 2256–2265. pmlr, 2015.

Duncan Soiffer, Chandler Squires, Yuan Guan, Jason Hartford, and Pradeep Ravikumar. Catastrophic compositional generation: Why vanilla diffusion models fail to extrapolate. *arXiv preprint arXiv:2606.23920*, 2026.

Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *The Ninth International Conference on Learning Representations*, 2021. URL .

Charlie B. Tan, Joey Bose, Chen Lin, Leon Klein, Michael M. Bronstein, and Alexander Tong. Scalable equilibrium sampling with sequential Boltzmann generators. In *Forty-second International Conference on Machine Learning*, 2025.

James Thornton, Louis Béthune, Ruixiang ZHANG, Arwen Bradley, Preetum Nakkiran, and Shuangfei Zhai. Composition and control with distilled energy diffusion models and sequential monte carlo. In Yingzhen Li, Stephan Mandt, Shipra Agrawal, and Emtiyaz Khan (eds.), *Proceedings of The 28th International Conference on Artificial Intelligence and Statistics*, volume 258 of *Proceedings of Machine Learning Research*, pp. 3259–3267. PMLR, 03–05 May 2025. URL .

Glenn M Torrie and John P Valleau. Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling. *Journal of computational physics*, 23(2):187–199, 1977.

Masatoshi Uehara, Yulai Zhao, Kevin Black, Ehsan Hajiramezanali, Gabriele Scalia, Nathaniel Lee Diamant, Alex M Tseng, Tommaso Biancalani, and Sergey Levine. Fine-tuning of continuous-time diffusion models as entropy-regularized control. *arXiv preprint arXiv:2402.15194*, 2024.

Chenyang Wang, Weizhong Wang, Yinuo Ren, Jose Blanchet, and Yiping Lu. Simple approximation and derivative free inference-time scaling for diffusion models via sequential Monte Carlo on path measures. In *Forty-third International Conference on Machine Learning*, 2026.

Joseph L Watson, David Juergens, Nathaniel R Bennett, Brian L Trippe, Jason Yim, Helen E Eisenach, Woody Ahern, Andrew J Borst, Robert J Ragotte, Lukas F Milles, et al. De novo design of protein structure and function with rfdiffusion. *Nature*, 620(7976):1089–1100, 2023.

Li K. Wenliang and Heishiro Kanagawa. Blindness of score-based methods to isolated components and mixing proportions, 2021.

Luhuan Wu, Brian Trippe, Christian Naesseth, David Blei, and John P Cunningham. Practical and asymptotically exact conditional sampling in diffusion models. *Advances in Neural Information Processing Systems*, 36:31372–31403, 2023.

Yu Xie, Ludwig Winkler, Lixin Sun, Sarah Lewis, Adam Foster, Jose Jimenez-Luna, Tim Hempel, Michael Gastegger, Yaoyi Chen, Iryna Zaporozhets, Cecilia Clementi, Christopher M. Bishop, and Frank Noe. Enhanced diffusion sampling: Efficient rare event sampling and free energy calculation with diffusion models. In *ICML 2026 Workshop on Structured Probabilistic Inference & Generative Modeling*, 2026. URL .

Minkai Xu, Lantao Yu, Yang Song, Chence Shi, Stefano Ermon, and Jian Tang. Geodiff: A geometric diffusion model for molecular conformation generation. In *International Conference on Learning Representations*, 2022. URL .

Hanlin Yu, Arto Klami, Aapo Hyvarinen, Anna Korba, and Omar Chehab. Density ratio estimation with conditional probability paths. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267, pp. 73146–73174. PMLR, 13–19 Jul 2025.

Hanlin Yu, RuiKang OuYang, Partha Kaushik, Arto Klami, Michael U Gutmann, and Omar Chehab. Learning energy-based models from stochastic interpolants using spatiotemporal differences. *arXiv preprint arXiv:2605.26850*, 2026.

Robert W. Zwanzig. High-temperature equation of state by a perturbation method. i. nonpolar gases. *The Journal of Chemical Physics*, 22(8):1420–1426, 08 1954. ISSN 0021-9606.

A Energy-Based Diffusion Model Training

Following Diffusive Classification (OuYang et al., 2026), we train the time-dependent energy model E_θ in up to two stages: Stage 1 relies on denoising score matching, and an optional Stage 2 further calibrates the model through noise-level classification.

The forward SDE (1) admits the closed-form solution

$$\mathbf{y}_s = \underbrace{\exp\left(-\int_0^s F_u \mathrm{d}u\right)}_{=:S(s)} \mathbf{y}_0 + \underbrace{\exp\left(-\int_0^s F_u \mathrm{d}u\right) \sqrt{\int_0^s \beta_u^2 \exp\left(2\int_0^u F_r \mathrm{d}r\right) \mathrm{d}u}}_{=:\nu(s)} \mathbf{z},$$

where $\mathbf{y}_0 \sim p^\star$ and $\mathbf{z} \sim \mathcal{N}(0, \mathrm{I}_d)$. We write $p^\star_{t|T}(\cdot \mid \mathbf{x}_T) = \mathcal{N}\big(S(T-t)\mathbf{x}_T,\, \nu(T-t)^2 \mathrm{I}_d\big)$ for the resulting law of $\mathbf{x}_t$ conditioned on $\mathbf{x}_T$.

Tweedie's formula then relates the score of the marginal to the score of this conditional,

$$\nabla \log p^\star_t(\mathbf{x}) = \mathbb{E}\left[\nabla_{\mathbf{x}} \log p^\star_{t|T}(\mathbf{x} \mid \mathbf{x}_T) \,\middle|\, \mathbf{x}_t = \mathbf{x}\right], \tag{10}$$

which motivates the Stage 1 denoising score matching objective

$$\mathcal{L}_{\mathrm{DSM}}(\theta) = \mathbb{E}\left[\lambda(t) \left\| -\nabla_{\mathbf{x}} E_\theta(\mathbf{x}_t, t) - \nabla_{\mathbf{x}_t} \log p^\star_{t|T}(\mathbf{x}_t \mid \mathbf{x}_T) \right\|^2\right],$$

training the model to match the spatial gradient of the energy at each noise level. Tweedie's formula (10) also gives a direct link between the denoiser and the score,

$$\begin{aligned} D^\star_t(\mathbf{x}) &= \mathbb{E}[\mathbf{x}_T \mid \mathbf{x}_t = \mathbf{x}] = \frac{\mathbf{x} + \nu(T-t)^2 \nabla_{\mathbf{x}} \log p^\star_t(\mathbf{x})}{S(T-t)}, \\ D_t(\mathbf{x}) &= \frac{\mathbf{x} - \nu(T-t)^2 \nabla_{\mathbf{x}} E_\theta(\mathbf{x}, t)}{S(T-t)}, \end{aligned} \tag{11}$$

where D_t is the model denoiser obtained by replacing the true score with $-\nabla_{\mathbf{x}} E_\theta$.

Stage 1 alone, however, only matches local properties of the energy landscape: it misses the relative proportions between disjoint regions of the energy (Wenliang & Kanagawa, 2021; OuYang et al., 2026, Appendix C). Stage 2 corrects for this through noise-level classification: we select noise levels $t_{1:K}$ and train the model to identify the level from which a perturbed sample was drawn. The resulting classifier is

$$p_\theta(c = i \mid \mathbf{x}) = \frac{\exp\left(-E_\theta(\mathbf{x}, t_i)\right)}{\sum_{j=1}^{K} \exp\left(-E_\theta(\mathbf{x}, t_j)\right)},$$

trained with the classification loss

$$\mathcal{L}_{\mathrm{DiffCLF}}(\theta) = -\frac{1}{K} \sum_{i=1}^{K} \mathbb{E}_{\mathbf{x}_{t_i} \sim p_{t_i}} \left[\log p_\theta(c = i \mid \mathbf{x}_{t_i})\right].$$

The complete Stage 2 training objective is then

$$\mathcal{L}_{\mathrm{EBM}}(\theta) = \mathcal{L}_{\mathrm{DSM}}(\theta) + \lambda_{\mathrm{clf}}\, \mathcal{L}_{\mathrm{DiffCLF}}(\theta).$$

Preconditioned parameterization. We follow the preconditioning scheme of OuYang et al. (2026); Thornton et al. (2025). Let $\Gamma(t) = \nu(T-t)/S(T-t)$ denote the noise-to-signal ratio, and let μ_{data} and σ^2_{data} denote the empirical mean and scalar variance of the training data. We define the input scale

$$\alpha_{\mathrm{in}}(t) = \sqrt{S(T-t)^2 \sigma^2_{\mathrm{data}} + \nu(T-t)^2}$$

and input center $\beta_{\mathrm{in}}(t) = S(T-t)\mu_{\mathrm{data}}$, and parameterize the energy as

$$\begin{aligned} E_\theta(\mathbf{y}, t) = \frac{\|\mathbf{y}\|^2}{2\alpha_{\mathrm{in}}(t)^2} - \frac{\sigma_{\mathrm{data}}}{\Gamma(t)} U_t^\theta\left(\frac{\mathbf{y} - \beta_{\mathrm{in}}(t)}{\alpha_{\mathrm{in}}(t)}\right) \\ - \frac{S(T-t)\mu_{\mathrm{data}}^\top \mathbf{y}}{\alpha_{\mathrm{in}}(t)^2} + \frac{1}{2}\frac{S(T-t)^2}{\alpha_{\mathrm{in}}(t)^2}\|\mu_{\mathrm{data}}\|^2 \\ + \frac{d}{2}\log\left(2\pi\alpha_{\mathrm{in}}(t)^2\right). \end{aligned}$$

With this parameterization, the network only needs to learn a non-Gaussian residual in standardized coordinates, while the dominant Gaussian scale and normalization constant are represented analytically. Writing the standardized input as $\mathbf{z} = (\mathbf{y} - \beta_{\mathrm{in}}(t))/\alpha_{\mathrm{in}}(t)$, we construct the scalar residual U_t^θ from a vector-valued network NN_t^θ via

$$U_t^\theta(\mathbf{z}) = \mathbf{z}^\top \mathrm{NN}_t^\theta(\mathbf{z}).$$

We parameterize NN_t^θ with an $E(n)$-equivariant graph neural network (EGNN), using sinusoidal time embeddings together with learnable bond- and atom-type embeddings.

B Algorithm Details

B.1 Fokker–Planck Equation for a Prescribed Path

We prove Thm. 2 here. Suppose that q_t satisfies the Feynman–Kac-type Fokker–Planck equation

$$\partial_t q_t(\mathbf{x}) = -\nabla \cdot [q_t(\mathbf{x}) v_t(\mathbf{x})] + \frac{\sigma_t^2}{2} \Delta q_t(\mathbf{x}) + q_t(\mathbf{x}) g_t(\mathbf{x}). \tag{12}$$

Dividing Equation (12) by $q_t(\mathbf{x})$ and using $\nabla \cdot (q_t v_t)/q_t = \nabla \cdot v_t + v_t^\top \nabla \log q_t$ and $\Delta q_t / q_t = \Delta \log q_t + \|\nabla \log q_t\|^2$ gives the reweighting potential

$$\begin{aligned} g_t(\mathbf{x}) = \partial_t \log q_t(\mathbf{x}) + \nabla \cdot v_t(\mathbf{x}) + v_t(\mathbf{x})^\top \nabla \log q_t(\mathbf{x}) \\ - \frac{\sigma_t^2}{2}\left[\Delta \log q_t(\mathbf{x}) + \|\nabla \log q_t(\mathbf{x})\|^2\right]. \end{aligned} \tag{13}$$

For the prescribed path $q_t(\mathbf{x}) = Z_t^{-1} \exp(-\Phi_t(\mathbf{x}))$, its log-density derivatives satisfy $\partial_t \log q_t = -\partial_t \Phi_t - \partial_t \log Z_t$, $\nabla \log q_t = -\nabla \Phi_t$, and $\Delta \log q_t = -\Delta \Phi_t$. Because $\partial_t \log Z_t$ is state independent, define the uncentered potential as

$$G_t = g_t + \partial_t \log Z_t.$$

We then substitute the proposal drift $v_t = -\widetilde{u}_t(\mathbf{x}) - \sigma_t^2 \nabla \Phi_t(\mathbf{x})$ into Equation (13). Expanding the divergence and inner-product terms yields

$$\begin{aligned} G_t(\mathbf{x}) = & -\partial_t \Phi_t(\mathbf{x}) - \nabla \cdot \widetilde{u}_t(\mathbf{x}) + \widetilde{u}_t(\mathbf{x})^\top \nabla \Phi_t(\mathbf{x}) \\ & - \frac{\sigma_t^2}{2} \Delta \Phi_t(\mathbf{x}) + \frac{\sigma_t^2}{2} \left\| \nabla \Phi_t(\mathbf{x}) \right\|^2 . \end{aligned}$$

This is the uncentered potential in Equation (8).

It remains to determine the state-independent normalization term. Assuming the boundary flux vanishes, integrating Equation (12) over $\mathbf{x}$ and using $\int q_t(\mathbf{x})\, \mathrm{d}\mathbf{x} = 1$ gives $\mathbb{E}_{q_t}[g_t] = 0$. Since $g_t = G_t - \partial_t \log Z_t$, the normalization term is $\partial_t \log Z_t = \mathbb{E}_{q_t}[G_t]$. Consequently, the normalized reweighting potential is

$$g_t(\mathbf{x}) = G_t(\mathbf{x}) - \mathbb{E}_{q_t}[G_t],$$

which proves Thm. 2.

B.2 Variance-Controlling Guidance

The Feynman–Kac dynamics in Eq. (7) preserve the prescribed path at the population level, but a finite-particle SMC approximation is prone to weight degeneracy when the reweighting potential has high variance. Variance-Controlling Guidance (VCG) (Ren et al., 2026) exploits the degree of freedom in the Fokker–Planck equation to actively reduce this variance. Adding a control drift $\mathbf{b}_t$ gives an equivalent Fokker–Planck equation:

$$\partial_t q_t(\mathbf{x}) = -\nabla \cdot \left[\left(v_t(\mathbf{x}) + \mathbf{b}_t(\mathbf{x}) \right) q_t(\mathbf{x}) \right] + \frac{\sigma_t^2}{2} \Delta q_t(\mathbf{x}) + q_t(\mathbf{x}) \phi_t(\mathbf{x}),$$

where the residual potential is $\phi_t(\mathbf{x}) = g_t(\mathbf{x}) + h_t(\mathbf{x}; \mathbf{b}_t)$, and the control potential is $h_t(\mathbf{x}; \mathbf{b}_t) = -\nabla \Phi_t(\mathbf{x}) \cdot \mathbf{b}_t(\mathbf{x}) + \nabla \cdot \mathbf{b}_t(\mathbf{x})$. VCG selects the control drift by directly minimizing the variance of the residual potential:

$$\min_{\mathbf{b}_t} \ \mathrm{Var}_{\mathbf{x} \sim q_t} \left[\phi_t(\mathbf{x}) \right] = \mathrm{Var}_{\mathbf{x} \sim q_t} \left[g_t(\mathbf{x}) + h_t(\mathbf{x}; \mathbf{b}_t) \right]. \tag{14}$$

To obtain a computationally lightweight control, we restrict the drift to the linear span of n problem-informed basis: $\mathbf{b}_t(\mathbf{x}) = \sum_{i=1}^n \lambda_t^i \boldsymbol{\psi}_i(\mathbf{x}, t)$. Let $h_t^i(\mathbf{x}) = h_t(\mathbf{x}; \boldsymbol{\psi}_i)$. Under this linear ansatz, Eq. (14) reduces to the $n \times n$ system $\mathbf{A}_t \boldsymbol{\lambda}_t = \mathbf{c}_t$, where $A_{ij} = \mathbb{E}_{q_t}[h_t^i h_t^j]$ and $c_i = -\mathbb{E}_{q_t}[g_t h_t^i]$. Solving this system at each time step yields the control drift that minimizes the residual potential variance within the chosen linear span. In practice, $\mathbf{A}_t$ may be ill-conditioned, so we solve the regularized system

$$(\mathbf{A}_t + \lambda_{\mathrm{VCG}} \mathbf{I}_n) \boldsymbol{\lambda}_t = \mathbf{c}_t. \tag{15}$$

For the EAT path, we use the following problem-informed basis:

$$\boldsymbol{\psi}_1 = \mathbf{x}, \ \ \boldsymbol{\psi}_2 = \nabla_{\mathbf{x}} E_\theta(\mathbf{x}, t), \ \ \boldsymbol{\psi}_3 = \nabla_{\mathbf{x}} [E^\star \left(\mathcal{D}_t \left(\mathbf{x} \right) \right) - E_\theta(\mathcal{D}_t \left(\mathbf{x} \right), T)], \ \ \boldsymbol{\psi}_4 = \nabla_{\mathbf{x}} r \left(\mathcal{D}_t \left(\mathbf{x} \right) \right).$$

The identity basis $\boldsymbol{\psi}_1$ and learned-energy basis $\boldsymbol{\psi}_2$ are always active. We activate the target-energy correction basis $\boldsymbol{\psi}_3$ when $\kappa_t \neq 0$, and use the reward basis $\boldsymbol{\psi}_4$ for reward-tilting tasks. Compared with DriftLite (Ren et al., 2026), we add the target-energy correction basis $\boldsymbol{\psi}_3$, which is significantly important in control drift estimation. At each time, we retain the active basis, evaluate $h_t^i = h_t(\cdot; \boldsymbol{\psi}_i)$ on the current weighted particles, form weighted empirical estimates of $\mathbf{A}_t$ and $\mathbf{c}_t$, and solve the empirical version of the regularized system in Eq. (15).

B.3 LOOK-AHEAD MAP

The EAT path evaluates the reference and learned terminal energies, while both MIT and EAT paths evaluate the reward, at the current state $\mathbf{x}_t$. At high noise levels, these states can lie far from the data manifold, where evaluating $E^\star$ and its gradient can be numerically unstable. Therefore, in the challenging LJ-13 and molecular sampling tasks we evaluate these quantities at a look-ahead point defined by the map

$$\mathcal{D}_t(\mathbf{x}) = (1-\omega_t)\,\mathbf{x} + \omega_t\, D_t(\mathbf{x}), \tag{16}$$

which interpolates between $\mathbf{x}$ and the denoiser prediction $D_t(\mathbf{x})$ of Eq. (11) with strength $\omega_t \in [0,1]$. Flow maps (Sabour et al., 2025) could play the same role, but the denoiser is directly available from the learned energy. Two schedules govern the correction: the look-ahead strength ω_t and the correction schedule κ_t. Since the denoiser is weakly informative early in the reverse process, we activate the correction only later; over its active window ω_t decreases from the denoiser toward the identity, and $\omega_T = 0$ gives $\mathcal{D}_T(\mathbf{x}) = \mathbf{x}$. Replacing the terminal-time evaluations by their look-ahead counterparts gives the practical potentials

$$\Phi_t^{\mathrm{MIT,LA}}(\mathbf{x}) = \gamma E_\theta(\mathbf{x},t) - r_t(\mathcal{D}_t(\mathbf{x})),$$
$$\Phi_t^{\mathrm{EAT,LA}}(\mathbf{x}) = \gamma E_\theta(\mathbf{x},t) - r_t(\mathcal{D}_t(\mathbf{x})) + \gamma\kappa_t\,\Delta E^{\mathrm{ref}}(\mathcal{D}_t(\mathbf{x})).$$

Because $\omega_T = 0$ and $\kappa_T = 1$, the terminal potential is still $\gamma E^{\mathrm{ref}}(\mathbf{x}) - r(\mathbf{x})$, so the look-ahead map alters only the intermediate marginals, not the endpoint. An imperfect denoiser can still send a few particles to non-physical configurations, so we apply the alive filter (App. B.7) for numerical stability.

B.4 DERIVATIVE COMPUTATION AND STOCHASTIC LAPLACIAN ESTIMATION

Hutchinson trace estimation. We use Hutchinson trace estimator (Hutchinson, 1989) to estimate the Laplacian term. Let $H_t(\mathbf{x}) = \nabla_{\mathbf{x}}^2 \Phi_t(\mathbf{x})$, and let $\boldsymbol{\epsilon} \in \mathbb{R}^d$ be a random probe satisfying $\mathbb{E}[\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\top] = \mathbf{I}_d$. Since $\Delta\Phi_t = \mathrm{tr}(H_t)$, the Hutchinson identity gives

$$\Delta\Phi_t(\mathbf{x}) = \mathbb{E}_{\boldsymbol{\epsilon}}\left[\boldsymbol{\epsilon}^\top H_t(\mathbf{x})\boldsymbol{\epsilon}\right].$$

Given R independent probes, we evaluate $\ell_t^{(r)}(\mathbf{x}) = (\boldsymbol{\epsilon}^{(r)})^\top H_t(\mathbf{x})\boldsymbol{\epsilon}^{(r)}$ and average them as $\widehat{L}_t(\mathbf{x}) = R^{-1}\sum_{r=1}^R \ell_t^{(r)}(\mathbf{x})$. Both Rademacher and standard Gaussian probes satisfy the required second-moment condition. For the symmetric Hessian $H_t(\mathbf{x})$, the Gaussian estimator additionally contains the contribution from the Hessian diagonal:

$$\mathrm{Var}_{\mathrm{Gauss}}\left[\widehat{L}_t(\mathbf{x})\right] - \mathrm{Var}_{\mathrm{Rad}}\left[\widehat{L}_t(\mathbf{x})\right] = \frac{2}{R}\sum_i [H_t(\mathbf{x})]_{ii}^2.$$

Thus, Rademacher probes have no greater variance than Gaussian probes, and the inequality is strict whenever the Hessian has a nonzero diagonal entry. We therefore use Rademacher probes by default for the stochastic estimation of all Laplacian terms.

Derivatives and matrix-free Hessian–vector products. We first detail how to compute derivatives of functions composed with the look-ahead map in Eq. (16). For any differentiable function $f : \mathbb{R}^d \to \mathbb{R}$, define the look-ahead-composed function

$$F_t(\mathbf{x}) = f(\mathcal{D}_t(\mathbf{x})).$$

Let $J_{\mathcal{D}_t}(\mathbf{x})$ denote the Jacobian of $\mathcal{D}_t$ with respect to $\mathbf{x}$. The spatial gradient and time derivative of F_t follow from the chain rule:

$$\nabla_{\mathbf{x}} F_t(\mathbf{x}) = J_{\mathcal{D}_t}(\mathbf{x})^\top \nabla f(\mathcal{D}_t(\mathbf{x})),$$
$$\partial_t F_t(\mathbf{x}) = \nabla f(\mathcal{D}_t(\mathbf{x}))^\top \partial_t \mathcal{D}_t(\mathbf{x}).$$

For the look-ahead map in Eq. (16), these two derivatives of the map are

$$J_{\mathcal{D}_t}(\mathbf{x}) = (1-\omega_t)\,\mathbf{I}_d + \omega_t J_{D_t}(\mathbf{x}),$$
$$\partial_t \mathcal{D}_t(\mathbf{x}) = \dot{\omega}_t\,(D_t(\mathbf{x}) - \mathbf{x}) + \omega_t \partial_t D_t(\mathbf{x}).$$

The Hessian of the composite function contains contributions from both f and the look-ahead map:

$$\nabla^2_{\mathbf{x}} F_t(\mathbf{x}) = J_{\mathcal{D}_t}(\mathbf{x})^\top \nabla^2 f\left(\mathcal{D}_t\left(\mathbf{x}\right)\right) J_{\mathcal{D}_t}(\mathbf{x}) + \sum_{j=1}^{d} \left[\nabla f\left(\mathcal{D}_t\left(\mathbf{x}\right)\right)\right]_j \nabla^2_{\mathbf{x}} \left[\mathcal{D}_t(\mathbf{x})\right]_j .$$

We evaluate the required Hessian-vector product (HVP) without explicitly forming the Hessian. For the learned intermediate energy, second-order automatic differentiation directly computes

$$\nabla^2_{\mathbf{x}} E_\theta(\mathbf{x}, t)\boldsymbol{\epsilon} = \nabla_{\mathbf{x}} \left[\left(\nabla_{\mathbf{x}} E_\theta(\mathbf{x}, t)\right)^\top \boldsymbol{\epsilon}\right].$$

For a look-ahead-composed term F_t, we instead approximate the HVP by applying a centered finite difference to the complete spatial gradient:

$$\begin{aligned}\nabla^2_{\mathbf{x}} F_t(\mathbf{x})\boldsymbol{\epsilon} &\approx \frac{\nabla_{\mathbf{x}} F_t(\mathbf{x} + \delta\boldsymbol{\epsilon}) - \nabla_{\mathbf{x}} F_t(\mathbf{x} - \delta\boldsymbol{\epsilon})}{2\delta} \\ &= \frac{1}{2\delta}\Big[J_{\mathcal{D}_t}(\mathbf{x} + \delta\boldsymbol{\epsilon})^\top \nabla f\left(\mathcal{D}_t(\mathbf{x} + \delta\boldsymbol{\epsilon})\right) \\ &\qquad - J_{\mathcal{D}_t}(\mathbf{x} - \delta\boldsymbol{\epsilon})^\top \nabla f\left(\mathcal{D}_t(\mathbf{x} - \delta\boldsymbol{\epsilon})\right)\Big], \end{aligned} \tag{17}$$

where $\delta > 0$ is the finite-difference step size. By applying the finite difference to the complete composite gradient, this approximation captures the second-order contributions from both f and $\mathcal{D}_t$, without explicitly forming $\nabla^2 f$ or the componentwise Hessians of $\mathcal{D}_t$. In particular, we use this approximation with $f = E^*$ because OpenMM provides only first-order derivatives of the target energy.

In implementation, we evaluate the composite gradient without materializing either $J_{\mathcal{D}_t}$ or J_{D_t}. For an arbitrary evaluation point $\mathbf{z}$, let

$$\begin{cases} \mathbf{x}_t^{\mathrm{LA}}(\mathbf{z}) = \mathcal{D}_t(\mathbf{z}), \\ \quad \mathbf{g}_{\mathbf{z}} = \nabla f\big(\mathbf{x}_t^{\mathrm{LA}}(\mathbf{z})\big) . \end{cases}$$

We obtain $\mathbf{g}_{\mathbf{z}}$ by automatic differentiation for differentiable neural potentials. For $f = E^*$, we instead obtain this energy gradient from the force returned by OpenMM, with the sign reversed. Using $\mathbf{g}_{\mathbf{z}}$ as the upstream gradient at the output of $\mathcal{D}_t$, backpropagation computes the vector-Jacobian product

$$\mathbf{G}_t(\mathbf{z}) = J_{\mathcal{D}_t}(\mathbf{z})^\top \mathbf{g}_{\mathbf{z}} = \nabla_{\mathbf{z}} F_t(\mathbf{z}).$$

Because D_t is evaluated as part of $\mathcal{D}_t$, backpropagation also applies $J_{D_t}^\top$ internally without explicitly constructing either Jacobian. Finally, setting $\mathbf{z}_\pm = \mathbf{x} \pm \delta\boldsymbol{\epsilon}$, we approximate the composite HVP by

$$\nabla^2_{\mathbf{x}} F_t(\mathbf{x})\boldsymbol{\epsilon} \approx \frac{\mathbf{G}_t(\mathbf{z}_+) - \mathbf{G}_t(\mathbf{z}_-)}{2\delta},$$

which is Eq. (17). Thus, the desired second-order derivative can be computed by HVP, whereas each composite gradient in its finite-difference approximation is evaluated by a first-order VJP. The same calculation applies to other look-ahead-composed terms for which only first-order derivatives are available. However, the OpenMM finite-difference approximation is time-consuming and cannot be parallelized across particles. Using Machine-learned interatomic potential (MLIP) such as DPA4 (Li et al., 2026b) to evaluate target energy can help with this problem, and we leave this for future work.

Variance correction for stochastic log-weights. An unbiased stochastic Laplacian generally does not yield an unbiased incremental importance weight, because SMC exponentiates the estimated log-weight increment. Consider the EDM forward process (Karras et al., 2022) with $u_s = 0$ and $\beta_s^2 = 2s$, so that $\widetilde{u}_t = 0$ and $\sigma_t^2 = \beta_{T-t}^2 = 2(T-t)$; the potential in Eq. (8) then reduces to

$$G_t(\mathbf{x}) = -\partial_t \Phi_t(\mathbf{x}) - (T-t)\Delta\Phi_t(\mathbf{x}) + (T-t)\left\|\nabla\Phi_t(\mathbf{x})\right\|^2 .$$

For a fixed particle $\mathbf{x}$ at time t_k, let $h_k = t_{k+1} - t_k$ and define

$$D_k = -\partial_t \Phi_{t_k}(\mathbf{x}) + (T - t_k) \left\|\nabla \Phi_{t_k}(\mathbf{x})\right\|^2,$$
$$L_k = \Delta \Phi_{t_k}(\mathbf{x}), \qquad c_k = h_k(T - t_k).$$

Up to a particle-independent centering term, which cancels once the weights are normalized, the corresponding Euler log-weight increment is

$$\Delta \log w_k = h_k G_{t_k}(\mathbf{x}) = h_k D_k - c_k L_k.$$

Replacing L_k by its stochastic R-probe estimate $\widehat{L}_k = \widehat{L}_{t_k}(\mathbf{x})$ gives

$$\widehat{\Delta \log w}_k = \Delta \log w_k + \eta_k, \qquad \eta_k = -c_k(\widehat{L}_k - L_k).$$

By the Hutchinson identity, $\mathbb{E}_{\boldsymbol{\epsilon}}[\eta_k \mid \mathbf{x}] = 0$, so the estimated log-weight increment is unbiased conditional on the particle, but the incremental weight itself is not: exponentiating gives

$$\mathbb{E}_{\boldsymbol{\epsilon}}\left[\exp\left(\widehat{\Delta \log w}_k\right) \mid \mathbf{x}\right] = \exp(\Delta \log w_k)\mathbb{E}_{\boldsymbol{\epsilon}}[\exp(\eta_k) \mid \mathbf{x}] \geq \exp(\Delta \log w_k),$$

where the inequality follows from Jensen's inequality and is strict whenever the trace noise is non-degenerate. Since the trace-estimation variance generally depends on $\mathbf{x}$, this upward bias varies across particles and thereby distorts their normalized relative weights. Under the Gaussian approximation $\eta_k \mid \mathbf{x} \sim \mathcal{N}(0, v_k)$, the moment-generating function gives

$$\mathbb{E}_{\boldsymbol{\epsilon}}\left[\exp\left(\widehat{\Delta \log w}_k\right) \mid \mathbf{x}\right] = \exp\left(\Delta \log w_k + \frac{v_k}{2}\right).$$

Subtracting $v_k/2$ from the noisy log-weight increment therefore cancels the multiplicative factor introduced by the additional trace noise, following the classical cumulant-based bias correction of Zwanzig (1954) for exponentiated stochastic quantities. For $R \geq 2$, we estimate the conditional variance of the mean trace estimator by

$$\widehat{\sigma}^2_{L,k} = \frac{1}{R(R-1)} \sum_{r=1}^{R} \left(\ell^{(r)}_{t_k}(\mathbf{x}) - \widehat{L}_k\right)^2.$$

Since the Laplacian estimate enters the log-weight increment with coefficient $-c_k$, its additional variance at step k is estimated by

$$\widehat{v}_k = c_k^2 \widehat{\sigma}^2_{L,k}.$$

At every SMC step, we subtract half of this estimated variance from the noisy log-weight increment:

$$\begin{aligned}\widehat{\Delta \log w}_k^{\,\mathrm{corr}} &= h_k D_k - c_k \widehat{L}_k - \frac{1}{2}\widehat{v}_k \\ &= h_k D_k - c_k \widehat{L}_k - \frac{1}{2} c_k^2 \widehat{\sigma}^2_{L,k}. \end{aligned} \tag{18}$$

In our implementation, we apply the variance correction in Eq. (18) only to the Ala-4 free-energy estimation experiment. In the other experiments, we did not observe a noticeable effect of this correction on sampling performance.

Hutch++ estimator. We can also estimate $\Delta \Phi_t(\mathbf{x})$ using advanced Hutch++ (Meyer et al., 2021). Its key idea is to separate the Hessian into a data-dependent low-rank component and an orthogonal residual. The trace on the sketched subspace is evaluated directly, while stochastic probes are applied only to the residual. Thus, when the sketch captures the dominant spectral components of the Hessian, the randomly estimated residual has lower variance than a Hutchinson estimate applied to the full Hessian. Concretely, given an m-column Gaussian sketch $S \in \mathbb{R}^{d \times m}$, let $Q = \mathrm{orth}(H_t(\mathbf{x})S)$ and $P_Q = \mathbf{I}_d - QQ^\top$. We then draw independent residual probes $\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(R)} \sim \mathcal{N}(0, \mathbf{I}_d)$. For each probe, define the single-probe Hutch++ estimate

$$\ell_t^{++,(r)}(\mathbf{x}) = \mathrm{tr}\left(Q^\top H_t(\mathbf{x}) Q\right) + \left(\mathbf{g}^{(r)}\right)^\top P_Q H_t(\mathbf{x}) P_Q \mathbf{g}^{(r)}.$$

Their average is $\widehat{L}_t^{++} = R^{-1} \sum_{r=1}^{R} \ell_t^{++,(r)}$. The same low-rank term appears in every single-probe estimate and therefore cancels when these estimates are centered. For Hutch++, we can also adopt the variance correction method and apply Eq. (18) to reduce the corresponding additional variance. Hutch++ is not used in our implementation and we leave it for future work.

B.5 FKC WITH A TIME-DEPENDENT REWARD

Skreta et al. (2025, Proposition 3.4) considers tilts of the form $\beta_t r$. We extend it to an arbitrary deterministic, time-dependent reward R_t, which covers both the MIT and EAT paths.

Time-dependent reward tilting. Let $p_t > 0$ satisfy the base diffusion equation

$$\partial_t p_t = -\nabla \cdot \left[p_t \left(-\widetilde{u}_t + \sigma_t^2 \nabla \log p_t\right)\right] + \frac{\sigma_t^2}{2} \Delta p_t,$$

and let R_t be C^1 in time and C^2 in space, with $Z_t = \int p_t(x)\, e^{R_t(x)}\, dx \in (0, \infty)$. We claim that, under the usual integrability and vanishing-boundary-flux conditions, the tilted path $q_t = Z_t^{-1} p_t\, e^{R_t}$ is represented by the weighted process

$$dX_t = \left[-\widetilde{u}_t(X_t) + \sigma_t^2 \nabla \log p_t(X_t) + \frac{\sigma_t^2}{2} \nabla R_t(X_t)\right] dt + \sigma_t\, dW_t, \tag{19}$$

$$d\ell_t = G_t^R(X_t)\, dt, \qquad G_t^R = \partial_t R_t + \left\langle \nabla R_t,\, \frac{\sigma_t^2}{2} \nabla \log p_t - \widetilde{u}_t \right\rangle, \tag{20}$$

initialized with $X_0 \sim q_0$ and $\ell_0 = 0$, with weights proportional to e^{ℓ_t}. The representation is exact in the sense that the normalized weighted marginals equal q_t.

To verify this claim, define $a_t = -\widetilde{u}_t + \frac{\sigma_t^2}{2} \nabla \log p_t$. Since $\nabla p_t = p_t \nabla \log p_t$, the diffusion term in the base equation partially cancels the score term in the drift, leaving the continuity equation $\partial_t p_t = -\nabla \cdot (p_t a_t)$. Differentiating $q_t = Z_t^{-1} p_t e^{R_t}$ and using $\partial_t \log Z_t = \mathbb{E}_{q_t}[G_t^R]$, which follows by integrating over x under the boundary conditions, gives

$$\partial_t q_t = -\nabla \cdot (q_t a_t) + q_t \left(G_t^R - \mathbb{E}_{q_t}[G_t^R]\right).$$

Since $\nabla q_t = q_t \nabla \log q_t$, the terms $-\nabla \cdot \left(q_t\, \sigma_t^2 \nabla \log q_t / 2\right)$ and $\sigma_t^2 \Delta q_t / 2$ cancel. Adding the former to the drift and the latter as a diffusion term therefore leaves the equation unchanged. With $\nabla \log q_t = \nabla \log p_t + \nabla R_t$, the resulting drift is exactly (19), and the reaction term is realized by the weights (20). Taking $R_t = \beta_t r$ recovers Proposition 3.4 of Skreta et al. (2025).

Instantiation on MIT and EAT. Both paths follow directly from (19)–(20) with a suitable choice of reward. For the MIT path, we simply take $R_t = r_t$. For the EAT path, we take

$$R_t(\mathbf{x}) = r_t(\mathbf{x}) - \gamma \kappa_t\, \Delta E^{\mathrm{ref}}(\mathbf{x}).$$

Since $r_T = r$ and $\kappa_T = 1$, the terminal reward is $R_T = r - \gamma \Delta E^{\mathrm{ref}}$, and the endpoint of the tilted path is

$$\begin{aligned} q_T(\mathbf{x}) &\propto \left(p_T^\theta(\mathbf{x})\right)^\gamma \exp\left(R_T(\mathbf{x})\right) \\ &\propto \exp\left(-\gamma E_\theta(\mathbf{x}, T)\right) \exp\left(r(\mathbf{x}) - \gamma E^{\mathrm{ref}}(\mathbf{x}) + \gamma E_\theta(\mathbf{x}, T)\right) \\ &= \exp\left(-\gamma E^{\mathrm{ref}}(\mathbf{x}) + r(\mathbf{x})\right) \\ &\propto q_T^{\mathrm{ref}}(\mathbf{x}). \end{aligned}$$

FKC on EAT is therefore FKC on the model-induced construction with the redefined reward $\hat{r} = R_T$, under which the MIT coincides with q_T^{ref} and the endpoint mismatch vanishes by construction. Any error of FKC on EAT at the population level therefore stems from path tracking.

Scope of the extension. These constructions extend FKC to both paths, but they still require p_t^θ to satisfy the base diffusion equation. For an imperfect model this condition generally fails, so the extended FKC formulas need not track the prescribed path, even at the population level. EBFKC avoids this limitation: it evaluates the full prescribed-path potential of Thm. 2 without imposing diffusion consistency.

B.6 TEMPERING MODEL-FREE SMC

Tempering path. As a model-free baseline, we use traditional SMC along a geometric tempering path with a scheduled reward. Let $\rho = \mathcal{N}(0, \Sigma_0)$ with $\Sigma_0 \succ 0$ and $\Phi_G(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \Sigma_0^{-1}\mathbf{x}$. For a smooth nondecreasing schedule $\lambda(t)$ satisfying $\lambda(0) = 0$ and $\lambda(T) = 1$, we define the tempering path as

$$\Phi_t^{\text{temp}}(\mathbf{x}) = (1 - \lambda(t))\Phi_G(\mathbf{x}) + \lambda(t)\gamma E^*(\mathbf{x}) - \alpha_t r(\mathbf{x}), \tag{21}$$
$$q_t^{\text{temp}}(\mathbf{x}) \propto \rho(\mathbf{x})^{1-\lambda(t)}\big(p^\star(\mathbf{x})\big)^{\gamma\lambda(t)} \exp\big(\alpha_t r(\mathbf{x})\big).$$

With $\alpha_0 = 0$ and $\alpha_T = 1$, the path starts at $q_0^{\text{temp}} = \rho$ and terminates at the desired target $q_T^{\text{temp}} \propto \exp(-\gamma E^* + r)$.

Prescribed-path Langevin Feynman–Kac equation. For a prescribed path $q_t = Z_t^{-1}\exp(-\Phi_t)$, this baseline uses an annealed Langevin proposal. We choose a space-independent diffusivity $\varepsilon_t \geq 0$, with drift $v_t = -\varepsilon_t \nabla\Phi_t$ and diffusion coefficient $\sigma_t = \sqrt{2\varepsilon_t}$. Assume Φ_t is C^1 in time and C^2 in space, $0 < Z_t < \infty$, differentiation under the integral is valid, and boundary fluxes vanish. Then

$$\partial_t q_t = -\nabla\cdot(q_t v_t) + \varepsilon_t \Delta q_t + q_t g_t,$$
$$G_t(\mathbf{x}) = -\partial_t \Phi_t(\mathbf{x}), \qquad g_t(\mathbf{x}) = G_t(\mathbf{x}) - \mathbb{E}_{q_t}[G_t].$$

Indeed, $\nabla q_t = -q_t \nabla\Phi_t$ makes the drift and diffusion terms cancel, while differentiating q_t gives $\partial_t q_t = q_t g_t$. For the tempering path in Eq. (21), the uncentered reweighting potential is $G_t(\mathbf{x}) = \dot\lambda(t)[\Phi_G(\mathbf{x}) - \gamma E^*(\mathbf{x})] + \dot\alpha_t r(\mathbf{x})$.

Annealed Langevin dynamics with reweighting. With $\Phi_t = \Phi_t^{\text{temp}}$, the corresponding weighted particle dynamics between resampling events are

$$\mathrm{d}\mathbf{x}_t^{(i)} = -\varepsilon_t \nabla\Phi_t(\mathbf{x}_t^{(i)})\mathrm{d}t + \sqrt{2\varepsilon_t}\,\mathrm{d}\mathbf{w}_t^{(i)},$$
$$\mathrm{d}\log w_t^{(i)} = \left[G_t(\mathbf{x}_t^{(i)}) - \sum_{j=1}^N w_t^{(j)} G_t(\mathbf{x}_t^{(j)})\right]\mathrm{d}t.$$

Starting from $\mathbf{x}_0^{(i)} \overset{\text{iid}}{\sim} q_0^{\text{temp}}$ with $w_0^{(i)} = 1/N$, we discretize the particle SDE using Euler–Maruyama and apply left-point Euler updates to the log-weights. We then normalize the weights and resample when the ESS falls below the prescribed threshold, resetting the weights to $1/N$.

B.7 ALIVE FILTER FOR NUMERICAL STABILITY

During the look-ahead step of SMC inference, a small number of noisy particles can be mapped to severely non-physical configurations, producing non-finite or extremely large target energies and gradients, which will destroy the following VCG and SMC reweighting. Following the numerical stabilization strategy used in (Tan et al., 2025), we apply an energy-based *alive filter* to prevent numerically unstable, extremely high-energy configurations from propagating through SMC dynamics.

Specifically, at time t_k, we first use look-ahead map to predict each particle's terminal position $\mathbf{x}_{t_k}^{\text{LA},(i)} = \mathcal{D}_{t_k}\left(\mathbf{x}_{t_k}^{(i)}\right)$. In molecular experiments, we use the dimensionless reduced potential energy $u(\mathbf{x}) = U_{\text{OpenMM}}(\mathbf{x})/(RT_{\text{target}})$ at the target temperature T_{target}, where U_{OpenMM} is the molar potential energy and R is the molar gas constant. Given an energy threshold $c_{\text{alive}}(t_k)$, particle i is declared alive if both $u\left(\mathbf{x}_{t_k}^{\text{LA},(i)}\right)$ and its gradient are finite, and $u\left(\mathbf{x}_{t_k}^{\text{LA},(i)}\right) < c_{\text{alive}}(t_k)$.

Let $\mathcal{S}$ be the surviving set, we renormalize the current SMC weights over the surviving particles,

$$\bar w_i = \frac{w_{t_k}^{(i)}}{\sum_{j\in\mathcal{S}} w_{t_k}^{(j)}}, \qquad i \in \mathcal{S},$$

Algorithm 1: Energy-based Feynman–Kac Corrector on Energy-Aligned Target (EBFKC on EAT)

Input: Learned energy E_θ, target energy E^*, path schedule (α, κ); time grid $0 = t_0 < \cdots < t_K = T$; particle count N; ESS threshold τ.

Construct the EAT path $q_t \propto \exp(-\Phi_t)$; Initialize $\mathbf{x}_0^{(i)} \sim q_0$ and $w_0^{(i)} \leftarrow 1/N$ for $i \in [N]$

for $k \leftarrow 0$ **to** $K-1$ **do**

  Select the active EAT basis and solve the weighted VCG system in Eq. (15) for $\boldsymbol{\lambda}_{t_k}$

  Form the control drift $\mathbf{b}_{t_k} = \sum_{j \in \mathcal{I}(t_k)} \lambda_{t_k}^j \boldsymbol{\psi}_j$ and residual potential $\phi_{t_k} = \hat{g}_{t_k} + h_{t_k}(\cdot; \mathbf{b}_{t_k})$

  Compute $\delta\ell_k^{(i)} \leftarrow \Delta t_k \, \phi_{t_k}(\mathbf{x}_{t_k}^{(i)})$ at the current particle positions

  Propagate particles to t_{k+1} by Euler–Maruyama with drift $v_{t_k} + \mathbf{b}_{t_k}$ and diffusion coefficient σ_{t_k}

  Reweight $w_{t_{k+1}}^{(i)} \propto w_{t_k}^{(i)} \exp(\delta\ell_k^{(i)})$ and normalize

  if $\mathrm{ESS}(w_{t_{k+1}}^{(i)}) < \tau$ **then**

    Resample particles using their weights and reset $w_{t_{k+1}}^{(i)} \leftarrow 1/N$

Output: Weighted particles $\{(\mathbf{x}_T^{(i)}, w_T^{(i)})\}_{i=1}^N$ approximating $q_T^\star \propto \exp(-\gamma E^* + r)$.

and use systematic resampling from $\{\bar{w}_i\}_{i \in \mathcal{S}}$ to draw $N - |\mathcal{S}|$ parent particles whose copies replace the filtered particles. The post-filter weights are

$$\begin{cases} w_{t_k}'^{(i)} = \dfrac{|\mathcal{S}|}{N} \bar{w}_i, & i \in \mathcal{S}, \\ w_{t_k}'^{(j)} = \dfrac{1}{N}, & j \notin \mathcal{S}. \end{cases}$$

Thus, the alive filter preserves both the particle count and the total weight since $\sum_i w_{t_k}'^{(i)} = 1$. The replacement particles are sampled from the surviving set in proportion to the renormalized SMC weights $\{\bar{w}_i\}_{i \in \mathcal{S}}$. The alive filter samples from a truncated distribution by removing the extreme high-energy tail. In practice, this truncation removes only a few particles, and we observe no degradation in algorithm performance.

C Experimental Details

C.1 GMM settings

Target distribution. We consider an equally weighted Gaussian mixture model (GMM) with $K = 40$ components in $d = 30$ dimensions:

$$p^*(\mathbf{x}) = \frac{1}{K} \sum_{k=1}^{K} \mathcal{N}\left(\mathbf{x}; \boldsymbol{\mu}_k, 50\mathbf{I}_d\right),$$

where the component means are sampled independently from $\mathrm{Unif}([-40, 40]^d)$. We apply the variance exploding (VE) forward process, whose samples at noise level s are given by $\mathbf{y}_s = \mathbf{y}_0 + s\boldsymbol{\epsilon}$ for $s \in [0, T]$, with $\mathbf{y}_0 \sim p^\star$ and $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \mathbf{I}_d)$ independently. The resulting forward marginals are

$$r_s^\star(\mathbf{y}) = \frac{1}{K} \sum_{k=1}^{K} \mathcal{N}\left(\mathbf{y}; \boldsymbol{\mu}_k, (50 + s^2)\mathbf{I}_d\right).$$

Under the generative-time convention $t = T - s$, these marginals give $p_t^\star(\mathbf{x}) = r_{T-t}^\star(\mathbf{x})$. Consequently, both the intermediate energy $E_t^\star(\mathbf{x}) = -\log p_t^\star(\mathbf{x})$ and its score function $s_t^*(\mathbf{x}) = -\nabla_{\mathbf{x}} E_t^\star(\mathbf{x})$ are available in closed form. This allows us to emulate pretrained model error through explicit perturbations to the component means and variance, with direct control over the perturbation magnitudes.

Controlled pretrained model error. We emulate an imperfect pretrained model by independently perturbing the component means and variance. For each component, we add an independent Gaussian perturbation to its mean:

$$\widehat{\boldsymbol{\mu}}_k = \boldsymbol{\mu}_k + \boldsymbol{\xi}_k,$$

where $\boldsymbol{\xi}_k \sim \mathcal{N}(0, \sigma_\mu^2 \mathbf{I}_d)$. We also perturb the reference variance $\sigma_s^2 = 50 + s^2$ by rescaling the noise contribution by a factor of $1 + \rho$ to obtain the perturbed variance

$$\widehat{\sigma}_s^2 = 50 + (1 + \rho)s^2.$$

These perturbations define the marginals

$$r_s^\theta(\mathbf{y}) = \frac{1}{K} \sum_{k=1}^{K} \mathcal{N}\left(\mathbf{y}; \widehat{\boldsymbol{\mu}}_k, \widehat{\sigma}_s^2 \mathbf{I}_d\right),$$

with corresponding generative-time marginals $p_t^\theta(\mathbf{x}) = r_{T-t}^\theta(\mathbf{x})$ and intermediate energy $\widehat{E}(\mathbf{x}, t) = -\log p_t^\theta(\mathbf{x})$.

For the comparisons in Table 1, we set $\sigma_\mu = 10$ and $\rho = 0.1$. To evaluate robustness across perturbation magnitudes, we sweep each perturbation parameter over five evenly spaced values: $\sigma_\mu \in \{0, 2.5, 5, 7.5, 10\}$ and $\rho \in \{0, 0.025, 0.05, 0.075, 0.1\}$.

Inference-time targets. Following Eq. (5), both tasks use the unified terminal target

$$q_T^\star(\mathbf{x}) \propto (p^*(\mathbf{x}))^\gamma \exp\left(r(\mathbf{x})\right).$$

For annealing, we set $r(\mathbf{x}) = 0$ and $\gamma = 2$ to sharpen the reference distribution. For reward tilting, we set $\gamma = 1$ and use a quadratic reward centered at $\boldsymbol{\mu} \sim \mathcal{N}(\mathbf{0}, 100\mathbf{I}_d)$, with $\boldsymbol{\Sigma} = \sigma \mathbf{I}_d$ and $\sigma = 100$:

$$r(\mathbf{x}) = -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}).$$

This reward favors samples near $\boldsymbol{\mu}$, allowing us to test whether inference-time sampling can redirect probability mass toward this region despite pretrained model error.

C.2 Particle system settings

Following the setup of DriftLite (Ren et al., 2026), we consider two particle systems: Double-Well (DW-4) and Lennard-Jones (LJ-13). For a configuration $\mathbf{x} = (\mathbf{r}_1, \ldots, \mathbf{r}_n)$ with a system-specific interaction energy $H(\mathbf{x})$, the equilibrium distribution at temperature T_{phys} is

$$p^\star(\mathbf{x}) \propto \exp\left\{ -\frac{1}{T_{\text{phys}}} \left[H(\mathbf{x}) + \frac{\lambda}{2} \sum_{i=1}^{n} \|\mathbf{r}_i - \overline{\mathbf{r}}\|^2 \right] \right\}, \tag{22}$$

where λ controls harmonic confinement around configuration center $\overline{\mathbf{r}}$. For both particle systems, we obtain ground truth reference samples from underdamped Langevin dynamics simulations using the BAOAB splitting scheme.

For reward tilting, we use the following quadratic reward to favor spatially compact configurations centered at $\overline{\mathbf{r}}$:

$$r(\mathbf{x}) = -\frac{\lambda'}{2} \sum_{i=1}^{n} \|\mathbf{r}_i - \overline{\mathbf{r}}\|^2. \tag{23}$$

Combining Eqs. (22) and (23) gives

$$p^\star(\mathbf{x}) \exp\left(r(\mathbf{x})\right) \propto \exp\left\{ -\frac{1}{T_{\text{phys}}} \left[H(\mathbf{x}) + \frac{\lambda + T_{\text{phys}}\lambda'}{2} \sum_{i=1}^{n} \|\mathbf{r}_i - \overline{\mathbf{r}}\|^2 \right] \right\}.$$

Thus, the reward preserves the interaction potential and temperature while increasing the effective confinement strength from λ to $\lambda + T_{\text{phys}}\lambda'$. The default reward strength is $\lambda' = 0.5$ for both particle systems.

Double-Well system. The DW-4 benchmark consists of four particles in two dimensions, so $\mathbf{x} \in \mathbb{R}^{4\times 2}$. Let $d_{ij} = \|\mathbf{r}_i - \mathbf{r}_j\|$ denote the distance between a pair of particles. The interaction energy is

$$H_{\text{DW}}(\mathbf{x}) = \sum_{i<j} \left[a(d_{ij} - d_0) + b(d_{ij} - d_0)^2 + c(d_{ij} - d_0)^4 \right].$$

We set $a = 0$, $b = -4$, $c = 0.9$, and $d_0 = 4$. The resulting pair potential has two preferred separations, $d_{ij} = 4 \pm \frac{2}{3}\sqrt{5}$, which induce the characteristic double-well structure. The harmonic-confinement strength in Eq. (22) is $\lambda = 0.05$.

Lennard-Jones system. The LJ-13 benchmark contains thirteen particles in three dimensions, with $\mathbf{x} \in \mathbb{R}^{13\times 3}$. Its interaction energy is

$$H_{\mathrm{LJ}}(\mathbf{x}) = \frac{\epsilon}{2} \sum_{i<j} \left[\left(\frac{\sigma}{d_{ij}} \right)^{12} - 2 \left(\frac{\sigma}{d_{ij}} \right)^{6} \right].$$

The inverse-power terms produce strong short-range repulsion and longer-range attraction, yielding a rugged energy landscape for the thirteen-particle cluster. We use $\sigma = 1$, $\epsilon = 2$, and harmonic confinement strength $\lambda = 1$.

C.3 Alanine Dipeptide settings

Molecular system and reference data. Following the setup of RNE (He et al., 2026), we consider alanine dipeptide (ALDP), a 22-atom molecule represented by $\mathbf{x} \in \mathbb{R}^{22\times 3}$. Its potential energy is specified by the AMBER ff96 force field with a generalized Born implicit-solvent model.

We use the configurations provided by RNE: 210,384 at 300 K and 150,000 at 800 K. The 300 K reference ensemble was generated by a 5 μs molecular dynamics simulation using a Langevin middle integrator with a friction coefficient of $1\,\mathrm{ps}^{-1}$ and a time step of 2 fs. We center each configuration before training.

Free-energy estimation. We estimate the one-dimensional free-energy profile at 300 K along the backbone dihedral angle ϕ. The periodic reaction coordinate is $\xi_{\mathrm{ALDP}}(\mathbf{x}) = \phi(\mathbf{x})$, computed from the atom quadruplet $(4, 6, 8, 14)$. To obtain the reference profile, we perform umbrella sampling with 25 windows whose centers are uniformly spaced over the periodic domain:

$$\phi_i = -\pi + \frac{2\pi i}{25}, \qquad i = 0, \ldots, 24.$$

In each window, we use a Langevin middle integrator with a friction coefficient of $1\,\mathrm{ps}^{-1}$ and a time step of 2 fs. Each trajectory consists of 50,000 equilibration steps (100 ps) followed by 1,000,000 production steps (2 ns). We save one configuration every 1,000 production steps (2 ps), yielding 1,000 configurations per window.

C.4 Alanine Tetrapeptide settings

Molecular system and reference data. Following Invernizzi et al. (2022), we consider a 42-atom alanine tetrapeptide (Ala-4) system. The original benchmark constrains covalent bonds involving hydrogen atoms and represents configurations using 98 internal coordinates. We instead impose no bond-length constraints (`constraints=None` in OpenMM) and use Cartesian coordinates $\mathbf{x} \in \mathbb{R}^{42\times 3}$, retaining the stretching degrees of freedom of bonds involving hydrogen atoms. The backbone has three pairs of Ramachandran angles, $(\phi_i, \psi_i)_{i=1}^{3}$. The three ϕ_i angles are slow variables, and their sign combinations define eight metastable basins.

To generate reference data for the implicit-solvent equilibrium ensemble at 300 K, we run molecular dynamics with on-the-fly probability enhanced sampling (OPES) (Invernizzi & Parrinello, 2020), applying the bias along all three slow dihedrals ϕ_1, ϕ_2, and ϕ_3. We use a friction coefficient of $1\,\mathrm{ps}^{-1}$ and a time step of 1 fs. The simulation consists of 500,000 equilibration steps (0.5 ns) without OPES bias, followed by 100,000,000 production steps (100 ns) with the bias updated every 1,000 steps (1 ps). The resulting dataset contains 100,000 configurations with associated reweighting weights. We use these weights to estimate the equilibrium marginal of ϕ_3 and obtain the reference free-energy profile.

Free-energy estimation. We estimate the one-dimensional free-energy profile at 300 K along $\xi_{\mathrm{Ala4}}(\mathbf{x}) = \phi_3(\mathbf{x}) \in [-\pi, \pi)$, whose range is covered by overlapping umbrella windows. For the traditional umbrella sampling baseline, we use the ALDP settings in App. C.3, with the bias applied to ϕ_3. For umbrella diffusion sampling, we combine the terminal weighted particles from all windows using weighted MBAR, as described in App. C.5. All diffusion-based reward-tilting methods use the same window definitions, per-window sampling budgets, and angular bins. Each resulting free-energy profile is shifted so that its minimum is zero.

C.5 MOLECULAR FREE-ENERGY ESTIMATION METHODS

We estimate the potential of mean force (PMF) along a periodic reaction coordinate $\xi(\mathbf{x}) \in [-\pi, \pi)$. At temperature T_{phys}, the target equilibrium distribution is $p^*(\mathbf{x}) \propto \exp[-\beta_{\mathrm{th}} U(\mathbf{x})]$, where U is the molecular potential energy and $\beta_{\mathrm{th}} = 1/(k_{\mathrm{B}} T_{\mathrm{phys}})$.

Traditional umbrella sampling. Traditional umbrella sampling covers the reaction-coordinate range with K overlapping biased ensembles. Because a dihedral angle is periodic, we measure displacement from a window center c by $d_{\mathbb{T}}(\zeta, c) = \mathrm{atan2}(\sin(\zeta - c), \cos(\zeta - c))$. Window k uses the restraint potential $B_k(\mathbf{x}) = \kappa_k^{\mathrm{umb}} d_{\mathbb{T}}(\xi(\mathbf{x}), c_k)^2/2$ and its dimensionless counterpart $b_k(\mathbf{x}) = \beta_{\mathrm{th}} B_k(\mathbf{x})$. The desired biased ensemble in each window is therefore

$$q_k^*(\mathbf{x}) = \frac{1}{Z_k} p^*(\mathbf{x}) \exp\left[-b_k(\mathbf{x})\right]. \tag{24}$$

Here, Z_k is the normalization constant of window k. In traditional umbrella sampling, molecular dynamics under $U + B_k$ generates samples for each window. MBAR then combines the generally correlated samples across windows to account for the bias potentials and estimate equilibrium statistics under p^*.

Umbrella diffusion sampling. Following UmbrellaDiff (Xie et al., 2026), we formulate sampling within each umbrella window as a reward-tilting task with $\gamma = 1$ and $r_k(\mathbf{x}) = -b_k(\mathbf{x})$, so that the desired terminal distribution is q_k^* in Eq. (24). For each window, we initialize particles independently from the Gaussian prior and run the SMC sampler based on the pretrained diffusion models, obtaining terminal weighted particles $\{(\mathbf{x}_{k,i}, w_{k,i})\}_{i=1}^{N_k}$.

Weighted MBAR. We combine weighted particles across umbrella windows using the weighted multistate Bennett acceptance ratio (weighted MBAR) estimator of Xie et al. (2026). For window k, we compute the effective sample size as $\mathrm{ESS}_k = (\sum_i w_{k,i})^2 / \sum_i w_{k,i}^2$. We normalize the relative weights within each window as $\widetilde{w}_{k,i} = w_{k,i} / \sum_{j=1}^{N_k} w_{k,j}$ and assign each sample the effective mass $\alpha_{k,i} = \mathrm{ESS}_k \, \widetilde{w}_{k,i}$. The total effective mass of window k is then $M_k = \sum_i \alpha_{k,i} = \mathrm{ESS}_k$. This construction preserves the relative particle weights within each window and sets its total effective mass to ESS_k.

Pooling all windows gives $\{(\mathbf{x}_n, \alpha_n)\}_{n=1}^N$, where $N = \sum_{k=1}^K N_k$. We estimate the reduced free energies $\{\widehat{f}_k\}_{k=1}^K$ of the biased ensembles by solving the weighted MBAR self-consistency equations, for $k = 1, \ldots, K$:

$$\exp(-\widehat{f}_k) = \sum_{n=1}^{N} \frac{\alpha_n \exp\left[-b_k(\mathbf{x}_n)\right]}{\sum_{\ell=1}^K M_\ell \exp\left[\widehat{f}_\ell - b_\ell(\mathbf{x}_n)\right]}. \tag{25}$$

These equations determine the free energies up to a common additive constant, which does not affect normalized observables.

Given these reduced free energies, we reweight the pooled particles to estimate observables in a target state $a \in \{0, 1, \ldots, K\}$. The unnormalized weight of sample $\mathbf{x}_n$ in this state is

$$W_n^{(a)} = \frac{\alpha_n \exp\left[-b_a(\mathbf{x}_n)\right]}{\sum_{\ell=1}^K M_\ell \exp\left[\widehat{f}_\ell - b_\ell(\mathbf{x}_n)\right]}.$$

Here, $a = 1, \ldots, K$ index the sampled biased states, whereas $a = 0$ denotes the unbiased equilibrium state, for which $b_0 \equiv 0$. Because no samples are drawn directly from the unbiased state, it does not appear in the denominator of Eq. (25). We denote its unnormalized weights by $W_n \equiv W_n^{(0)}$ and normalize them as $\overline{W}_n = W_n / \sum_m W_m$. When the window particles consistently approximate q_k^*, the expectation of an observable $o(\mathbf{x})$ under the unbiased ensemble p^* is estimated using these normalized equilibrium weights:

$$\mathbb{E}_{p^*}[o(\mathbf{x})] \approx \sum_{n=1}^{N} \overline{W}_n o(\mathbf{x}_n).$$

Table 3: EGNN architectures and EBM training details.

System	Layers	Hidden dim.	EBM training epochs
DW-4	4	32	Stage 1: 125 epochs
LJ-13	4	32	Stage 1: 300 epochs
ALDP (800 K)	4	32	Stage 1: 200 epochs; Stage 2: 200 epochs
ALDP (300 K)	4	32	Stage 1: 200 epochs; Stage 2: 200 epochs
Ala-4	6	128	Stage 1: 500 epochs

Free-energy profile estimation. We use the normalized equilibrium weights $\overline{W}_n$ to estimate the marginal density along ξ and construct the PMF for ALDP and Ala-4. The equilibrium marginal density is $p_\xi(\zeta) = \int p^*(\mathbf{x})\delta(\xi(\mathbf{x}) - \zeta)\,\mathrm{d}\mathbf{x}$. For a partition $\{I_j\}$ of $[-\pi, \pi)$, we estimate the probability mass in bin j using the weighted samples:

$$\widehat{m}_j = \sum_n \overline{W}_n \mathbf{1}\{\xi(\mathbf{x}_n) \in I_j\}\,.$$

The corresponding density estimate is $\widehat{p}_j = \widehat{m}_j/|I_j|$, which gives the discrete PMF

$$\widehat{F}_j = -k_\mathrm{B}T_\mathrm{phys}\log\widehat{p}_j + C,$$

where C is an arbitrary additive constant. We choose C so that $\min_j \widehat{F}_j = 0$ and plot each estimate at the corresponding bin center.

C.6 Training and inference details

Training configuration. We train the energy-based diffusion models using the Adam optimizer with a learning rate of 5×10^{-4}. The batch size is 1024 for Stage 1 and 512 for Stage 2 by default. The loss weighting function is

$$\lambda(\sigma) = \frac{\sigma^2 + \sigma^2_\mathrm{data}}{\sigma^2\sigma^2_\mathrm{data}},$$

where σ is the noise level and σ^2_data is the empirical scalar variance of the training data. Table 3 lists the architecture and number of training epochs for each system.

Diffusion schedule. During inference, we use the following EDM-style noise schedule:

$$\sigma_i = \left[\sigma_\mathrm{max}^{1/7} + \frac{i}{K}\left(\sigma_\mathrm{min}^{1/7} - \sigma_\mathrm{max}^{1/7}\right)\right]^7, \qquad i = 0, \ldots, K,$$

where K denotes the number of inference steps and σ_i is the noise level at step i. For numerical stability, we truncate integration at $t_K = T - \sigma_\mathrm{min}$, corresponding to a small positive terminal noise level σ_min. For molecular systems, we use $(K, \sigma_\mathrm{min}, \sigma_\mathrm{max}) = (1000, 0.002, 80)$. The corresponding settings are $(1000, 0.005, 50)$ for particle systems and $(500, 0.005, 50)$ for GMM, with $T = \sigma_\mathrm{max}$ throughout. For the pretrained reverse-diffusion evaluations, we use the second-order Heun sampler with 512 steps. For annealing experiments, we initialize particles from the annealed Gaussian prior $q_0(\mathbf{x}) \propto \rho(\mathbf{x})^\gamma$. For $\rho = \mathcal{N}(0, \sigma^2_\mathrm{max}\mathbf{I})$, this corresponds to sampling from $\mathcal{N}(0, \frac{\sigma^2_\mathrm{max}}{\gamma}\mathbf{I})$ with uniform initial weights.

Reward and correction schedules. We instantiate the intermediate reward as $r_t(\mathbf{x}) = \alpha_t r(\mathbf{x})$. The reward schedule α_t and correction schedule κ_t satisfy

$$\alpha_0 = \kappa_0 = 0, \qquad \alpha_T = \kappa_T = 1.$$

We define the linear normalized progress as

$$p(t) = \mathrm{clip}\left(\frac{t}{T - \sigma_\mathrm{min}}, 0, 1\right) \in [0, 1].$$

At the inference grid points $t_i = T - \sigma_i$, this equals $(\sigma_{\max} - \sigma_i)/(\sigma_{\max} - \sigma_{\min})$. To control when each schedule activates and how it increases, we construct the delayed-power smootherstep non-linear progress based on the quintic polynomial $S(u) = 6u^5 - 15u^4 + 10u^3$:

$$A(p; s, q) = S\left(\left[\mathrm{clip}\left(\frac{p-s}{1-s}, 0, 1\right)\right]^q\right), \qquad s \in [0, 1), \quad q > 0.$$

Then the reward and correction schedules can be configured independently as

$$\alpha_t = A(p(t); s_\alpha, q_\alpha)\,, \qquad \kappa_t = A(p(t); s_\kappa, q_\kappa)\,.$$

Here, s sets the activation point, while q controls the growth shape. Separate parameter pairs specify the reward and correction schedules. Table 4 lists the settings for each experiment.

For Ala-4 umbrella diffusion sampling, a single pair of schedules does not serve all windows well, so we choose the activation order per window; since both orders share the same terminal values $\alpha_T = \kappa_T = 1$, this does not change the window target. For windows centered at $\phi_3 < 0$, where the pretrained model already provides adequate particle support, we use a *reward-first* path with $s_\alpha < s_\kappa$, so that the reward is activated before the target correction. For windows centered at $\phi_3 > 0$, a low-probability region that the pretrained model covers poorly, we use a *target-first* path with $s_\alpha > s_\kappa$, which first corrects the pretrained model error to improve particle coverage and then concentrates the corrected particles within the window.

Table 4: Reward and target-correction schedule parameters.

Experiment	(s_κ, q_κ)	(s_α, q_α)
DW-4	$(0.999, 3)$	(0.999,3)
LJ-13	$(0.999, 4)$	$(0.999, 2)$
ALDP annealing	$(0.999, 3)$	-
ALDP umbrella diffusion sampling	$(0.999, 3)$	$(0.9988, 1)$
Ala-4 umbrella diffusion sampling	$(0.999, 3)$	Reward-first: $(0.9988, 1)$ Target-first: $(0.9994, 1)$

Look-ahead schedule. We parameterize the look-ahead strength by the noise level so that the look-ahead map transitions from the denoiser prediction to the identity as sampling approaches the endpoint. We define

$$\omega_\sigma = S\left(\mathrm{clip}\left(\frac{\sigma - \sigma_{\min}}{\sigma_{\mathrm{LA}} - \sigma_{\min}}, 0, 1\right)\right),$$

For $\sigma_{\mathrm{LA}} > \sigma_{\min}$, the schedule equals one for $\sigma \geq \sigma_{\mathrm{LA}}$ and decreases to zero at $\sigma_{\min}$. We set $\sigma_{\mathrm{LA}} = 4\sigma_{\min}$ by default.

Inference configuration. We use $N = 8192$ particles for GMM, DW-4, and LJ-13, and $N = 1000$ particles per batch for ALDP annealing. For free-energy estimation, we use $N = 200$ particles per window for ALDP and $N = 12$ for Ala-4. For the Ramachandran plots, we pool terminal particles from 250 batches of $1{,}000$ particles each, yielding $250{,}000$ samples.

Across all SMC experiments, we use systematic resampling with an ESS threshold of $0.9N$. For methods using VCG, the regularization parameter is $\lambda_{\mathrm{VCG}} = 10^{-6}$. Unless otherwise stated, experiments use five random seeds.

For Hutchinson trace estimation, we use $R = 2$ probes for DW-4, $R = 4$ for LJ-13 and ALDP, and $R = 8$ for Ala-4. When finite differences are required, we use a default step size of $\delta = 10^{-4}$.

For the alive filter, we set the energy cutoff at time t_k to $c_{\mathrm{alive}}(t_k) = \frac{10}{\kappa_{t_k}}$ for ALDP and $c_{\mathrm{alive}}(t_k) = \frac{10^{20}}{\kappa_{t_k}}$ for Ala-4. For ALDP, the terminal cutoff of 10 exceeds the maximum reduced potential energy of -8.35 observed among $125{,}000$ reference configurations at $300\,\mathrm{K}$. Thus, none of these reference configurations exceeds the energy cutoff. The alive filter is used to exclude numerically problematic configurations encountered during SMC inference.

MALA rejuvenation. To improve particle mixing in molecular experiments, we apply Metropolis-adjusted Langevin algorithm (MALA) rejuvenation during SMC inference. For ALDP annealing, we enable MALA rejuvenation when $\kappa_t > 10^{-4}$ and perform 2 MALA steps every 25 SMC steps. ALDP umbrella diffusion sampling uses the same rule, with 5 additional MALA steps at the endpoint. For Ala-4 umbrella diffusion sampling, we instead enable MALA from the 500th SMC step and perform 2 MALA steps every 10 SMC steps, followed by 5 final MALA steps. In all molecular experiments, we additionally perform four MALA steps immediately after each resampling event.

Tempering SMC configuration. For the tempering SMC baseline, we use $N = 8192$ for LJ-13 and a uniform time grid with 1000 steps,

$$t_k = \frac{k}{1000}, \quad k = 0, \dots, 1000, \qquad \Delta t = 0.001.$$

We use the DABG (Grenioux & Noble, 2026) exponential tempering schedule

$$\lambda(t) = \frac{\delta_{\text{temp}}^{1-t} - \delta_{\text{temp}}}{1 - \delta_{\text{temp}}}, \qquad \delta_{\text{temp}} = 10^{-5},$$

and choose the diffusivity schedule

$$\varepsilon(t) = \varepsilon_{\text{final}} + \big(\varepsilon_{\text{initial}} - \varepsilon_{\text{final}}\big)(1-t)^2, \qquad t \in [0,1],$$

where $\varepsilon_{\text{initial}} = 1500$ and $\varepsilon_{\text{final}} = 1$.

C.7 Evaluation metrics

We describe the metrics and evaluation protocols used in our experiments. Throughout this section, P and Q denote the generated and reference empirical distributions with normalized sample weights. $E(\mathbf{x}) = -\gamma \log p^\star(\mathbf{x}) - r(\mathbf{x})$ denotes the target energy for the inference-time scaling task.

GMM evaluation. Following DriftLite (Ren et al., 2026), we evaluate GMM results using the following metrics. Let the generated samples be weighted samples $\{(w_i, \mathbf{x}_i)\}_{i=1}^N$ and the reference samples be uniform-weight samples $\{(1/M, \mathbf{y}_j)\}_{j=1}^M$.

- *Negative log-likelihood (NLL) difference:* $|\Delta\text{NLL}| = |\mathbb{E}_P[E] - \mathbb{E}_Q[E]|$ measures the log-likelihood difference between the generated and reference distributions.
- *Maximum mean discrepancy (MMD):* $\text{MMD} = \|\mathbb{E}_P[k(\mathbf{x}, \cdot)] - \mathbb{E}_Q[k(\mathbf{x}, \cdot)]\|_{\mathcal{H}}$ measures the gap between kernel mean embeddings. We use the RBF kernel $k(\mathbf{x}, \mathbf{y}) = \exp(-\|\mathbf{x} - \mathbf{y}\|^2/(2\sigma_k^2))$ with bandwidth $\sigma_k = 20$ for GMM and $\sigma_k = 5$ for the particle systems. In all experiments, we approximate the kernel using 2048 random Fourier features.
- *Sliced-Wasserstein distance (SWD):* We report the computationally efficient sliced-Wasserstein distance $\text{SWD}_2^2(P, Q) \approx L^{-1} \sum_{\ell=1}^L W_2^2(\pi_{\theta_\ell} P, \pi_{\theta_\ell} Q)$, where π_θ projects configurations onto direction θ. In all experiments, we use $L = 10$ projection directions sampled uniformly from the unit sphere.
- *Mean L_2:* $\text{Mean } L_2 = \|\hat{\boldsymbol{\mu}}_x - \hat{\boldsymbol{\mu}}_y\|_2$ measures the Euclidean distance between the global sample means, where $\hat{\boldsymbol{\mu}}_x = \sum_{i=1}^N w_i \mathbf{x}_i$ and $\hat{\boldsymbol{\mu}}_y = M^{-1} \sum_{j=1}^M \mathbf{y}_j$.
- *Cov F:* $\text{Cov } F = \|\hat{\boldsymbol{\Sigma}}_x - \hat{\boldsymbol{\Sigma}}_y\|_F$ measures the Frobenius distance between the global sample covariance matrices. Here, $\hat{\boldsymbol{\Sigma}}_x$ is the weighted sample covariance with degrees-of-freedom correction $1 - \sum_{i=1}^N w_i^2$, and $\hat{\boldsymbol{\Sigma}}_y$ is the reference sample covariance with denominator $M - 1$.

Particle system evaluation. Following DriftLite (Ren et al., 2026), we evaluate particle systems using the NLL difference, MMD and SWD defined above, together with two additional metrics for interparticle structure and the target-energy distribution.

- *Radial Distribution Function Wasserstein-1 Distance:* $W_1^{\text{RDF}} = W_1(g_P, g_Q)$ compares pooled pairwise-distance distributions. Here, the ensemble-averaged interatomic distance distribution g describes the distribution of interparticle distances and characterizes the spatial arrangement of particles.

- *Energy Wasserstein-1 distance:* $W_1^E = W_1(E_{\#}P, E_{\#}Q)$ compares the full target-energy distributions, where $E_{\#}P$ denotes the distribution of $E(\mathbf{x})$ for $\mathbf{x} \sim P$.

ALDP annealing evaluation. We evaluate ALDP annealing using four metrics. For each run reported in Table 2, we compare 25,000 generated samples with all RNE reference configurations at the corresponding temperature. Before evaluation, we center each conformation by subtracting its mean atomic position and apply chirality folding to match the reference enantiomer. To assess pretrained model error at $800\,\mathrm{K}$, we evaluate reverse-diffusion samples generated using the second-order Heun sampler with 512 steps, as described in App. C.6.

- *Ramachandran TVD:* We compare the joint distributions of the backbone dihedral angles (ϕ, ψ) using normalized 64×64 histograms over $[-\pi, \pi)^2$. The angles ϕ and ψ are computed from the atom quadruplets $(4, 6, 8, 14)$ and $(6, 8, 14, 16)$, respectively. Let P_{ij} and Q_{ij} denote the generated and reference probabilities in bin (i, j), respectively. The Ramachandran total variation distance (TVD) is

$$D_{\mathrm{Ram}} = \frac{1}{2} \sum_{i=1}^{64} \sum_{j=1}^{64} |P_{ij} - Q_{ij}| \,.$$

- *Basin-population error:* We compare the relative populations of four regions in the Ramachandran plane, defined by the signs of ϕ and ψ: α_R for $\phi < 0, \psi < 0$; extended/β for $\phi < 0, \psi \geq 0$; $\mathrm{C7_{ax}}$ for $\phi \geq 0, \psi < 0$; and α_L for $\phi \geq 0, \psi \geq 0$. Let p_b and q_b denote the generated and reference probabilities in basin b, respectively. The basin-population error is

$$E_{\mathrm{basin}} = \frac{1}{2} \sum_{b=1}^{4} |p_b - q_b| \,.$$

- *Energy TVD:* We compare the reduced-energy distributions using 100 equal-width bins spanning the MD reference energy range. We evaluate the potential energy $U(\mathbf{x})$ of each conformation using the unconstrained OpenMM ALDP system and convert it to the reduced energy $u(\mathbf{x}) = \beta U(\mathbf{x})$, where $\beta = 1/(k_{\mathrm{B}}T)$. Following the evaluation convention of FKC (Skreta et al., 2025) and RNE (He et al., 2026), we exclude samples outside the reference energy range before normalizing the histograms. Let $P_i^{(E)}$ and $Q_i^{(E)}$ denote the normalized generated and reference bin probabilities, respectively. The energy TVD is

$$D_{\mathrm{energy}} = \frac{1}{2} \sum_{i=1}^{100} \left| P_i^{(E)} - Q_i^{(E)} \right| .$$

- *Distance TVD:* We compare pooled distributions of all $\binom{22}{2} = 231$ pairwise atomic distances using 200 equal-width bins spanning the MD reference distance range. We compute these distances for each conformation and pool them across conformations. Let $P_i^{(d)}$ and $Q_i^{(d)}$ denote the normalized generated and reference bin probabilities, respectively. The distance TVD is

$$D_{\mathrm{distance}} = \frac{1}{2} \sum_{i=1}^{200} \left| P_i^{(d)} - Q_i^{(d)} \right| .$$

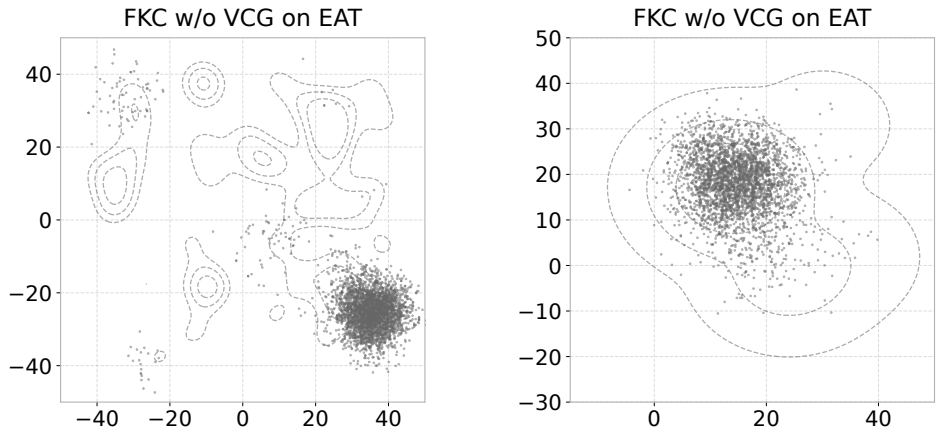
D Additional Experimental Results

D.1 GMM results

Without VCG, FKC on EAT concentrates samples in a limited subset of the target modes under both annealing and reward tilting, leaving other modes poorly represented (Fig. 6).

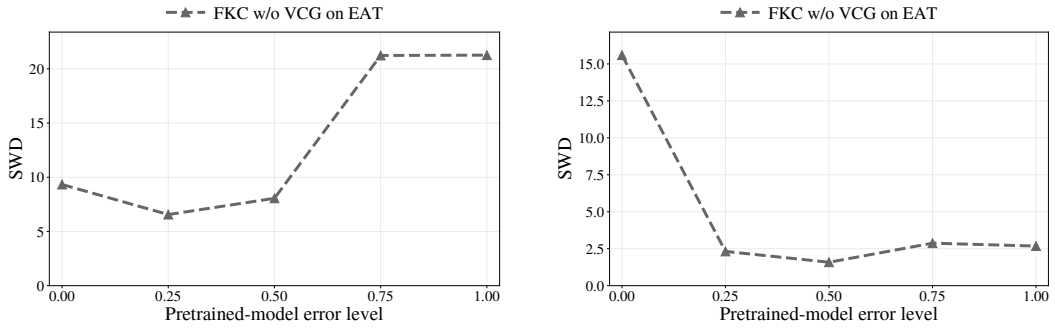
The SWD sweep in Fig. 7 further shows substantial sampling error even at zero pretrained-model error, with no monotonic dependence on the perturbation level in either task. Together with the FKC on EAT results using VCG in Fig. 2, these observations support the essential role of VCG in stablizing Feynman-Kac dynamics.

Table 5 reports the runtime of each method. FKC on EAT and EBFKC on EAT take longer than either MIT variant. Both EAT methods require additional evaluations for the target-energy correction, which add to the sampling cost.



(a) Annealing **(b) Reward tilting**

Figure 6: Sample distribution of FKC w/o VCG on EAT in the GMM task under (a) annealing and (b) reward tilting.



(a) Annealing **(b) Reward tilting**

Figure 7: SWD of FKC w/o VCG on EAT across pretrained-model error levels on the GMM under (a) annealing and (b) reward tilting.

Table 5: Runtime comparison on the 30-dimensional GMM under $\sigma_\mu = 10$ and $\rho = 0.1$. Results are mean $\pm$ std over 5 runs.

Method	Annealing ($\gamma = 2.0$)	Reward tilting ($\sigma = 100$)
FKC on EAT	5.357 ± 0.406 (s)	5.996 ± 0.736 (s)
EBFKC w/o VCG on MIT	3.098 ± 0.537 (s)	3.283 ± 0.445 (s)
EBFKC on MIT	3.626 ± 0.376 (s)	4.061 ± 0.963 (s)
EBFKC on EAT (Ours)	5.267 ± 0.082 (s)	5.503 ± 0.149 (s)

D.2 DOUBLE-WELL RESULTS

Fig. 8 demonstrates that the pretrained energy-based diffusion model is imperfect: its reverse diffusion samples fail to reproduce the reference RDF peaks and exhibit an energy distribution shifted toward higher energies. Fig. 9 and Fig. 10 show the results of annealing and reward tilting, respectively. In both tasks, EBFKC on EAT closely matches the target RDF and energy distribution, whereas FKC on EAT and EBFKC on MIT retain substantial deviations from the reference.

Table 6 reports the runtime of each method on the DW-4 tasks. For DW-4, we evaluate the required second-order derivatives of the target energy directly, without finite differences, so EBFKC on EAT takes reasonable 11% longer than EBFKC on MIT in both tasks.

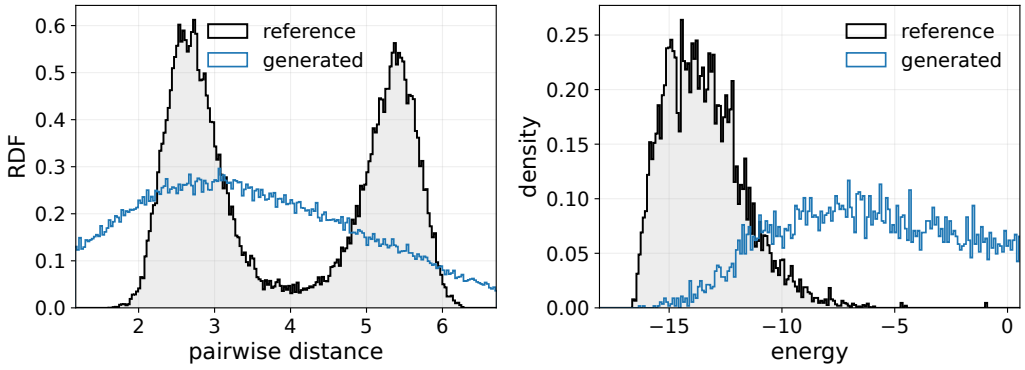
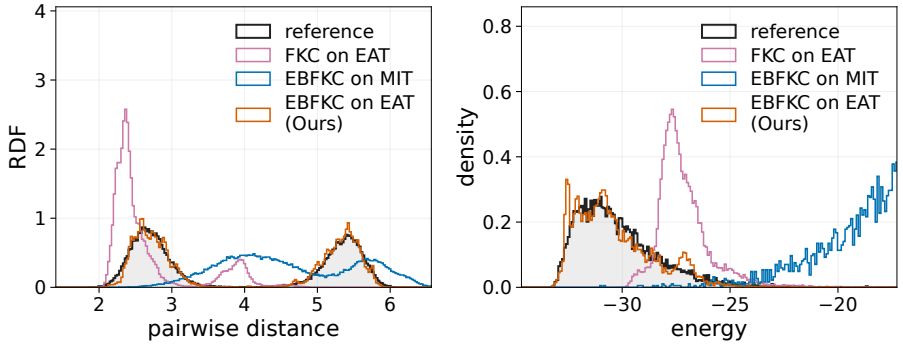
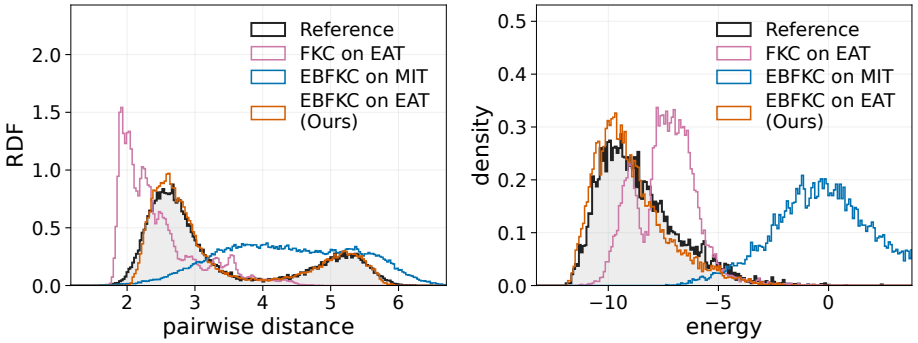


Figure 8: Comparison of the reference distribution and reverse diffusion samples from the pre-trained energy-based diffusion model on the double-well system.



(a) Radial Distribution Function **(b) Energy Distribution**

Figure 9: Comparison of the reference distribution, FKC on EAT, EBFKC on MIT, and EBFKC on EAT for the double-well system under annealing.



(a) Radial Distribution Function **(b) Energy Distribution**

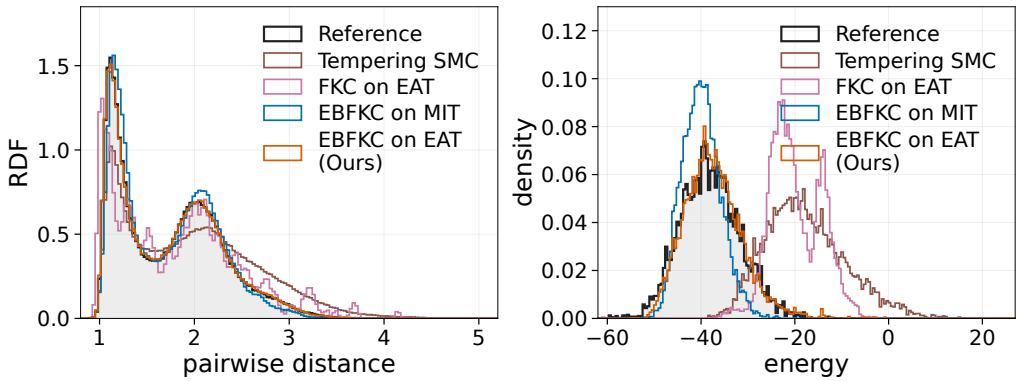
Figure 10: Comparison of the reference distribution, FKC on EAT, EBFKC on MIT, and EBFKC on EAT for the double-well system under reward tilting.

Table 6: Runtime comparison on the DW-4 tasks. Results are mean ± std over 5 runs.

Method	Annealing ($\gamma = 2.0$)	Reward tilting ($\lambda' = 0.5$)
FKC on EAT	412.594 ± 16.254 (s)	416.153 ± 17.053 (s)
EBFKC w/o VCG on MIT	342.789 ± 16.330 (s)	339.989 ± 13.836 (s)
EBFKC on MIT	367.378 ± 14.568 (s)	365.100 ± 18.214 (s)
EBFKC on EAT (Ours)	406.058 ± 16.066 (s)	403.458 ± 14.271 (s)

D.3 Lennard–Jones results

Annealing results. Fig. 11 compares the radial distribution functions and energy distributions of tempering SMC, FKC on EAT, EBFKC on MIT and EBFKC on EAT with the reference under LJ-13 annealing. EBFKC on EAT closely reproduces the reference RDF peaks and energy distribution, whereas EBFKC on MIT deviates from the reference RDF and shifts the energy distribution severely. Tempering SMC shifts the energy distribution toward higher energies, which demonstrates tempering SMC alone is insufficient to accurately sample from the Lennard–Jones system. Table 7 summarizes their performance across five metrics on DW-4 and LJ-13 under annealing and reward tilting.



(a) Radial Distribution Function **(b) Energy Distribution**

Figure 11: Comparison of the reference distribution, Tempering SMC, FKC on EAT, EBFKC on MIT and EBFKC on EAT for the Lennard-Jones system under annealing.

Table 7: Performance comparison on particle systems (DW-4 and LJ-13). Results are reported as mean ± std over 5 runs.

System	Task	Method	$\lvert\Delta\text{NLL}\rvert$	MMD	SWD	W_1^{RDF}	W_1^E
DW	Annealing	FKC on EAT	23.9565 ± 0.9045	0.2659 ± 0.0059	0.6263 ± 0.0511	1.6010 ± 0.0090	23.9565 ± 0.9045
		EBFKC w/o VCG on MIT	18.1923 ± 1.0910	0.4315 ± 0.0645	0.9894 ± 0.1340	0.7005 ± 0.0307	18.1923 ± 1.0910
		EBFKC on MIT	19.7381 ± 0.2039	0.0996 ± 0.0029	0.2341 ± 0.0217	0.7582 ± 0.0069	19.7381 ± 0.2039
		EBFKC on EAT (Ours)	$\mathbf{0.3389 \pm 0.2107}$	$\mathbf{0.0816 \pm 0.0193}$	$\mathbf{0.2032 \pm 0.0441}$	$\mathbf{0.3943 \pm 0.1437}$	$\mathbf{0.3754 \pm 0.1556}$
	Reward tilting	FKC on EAT	7.2620 ± 0.3554	0.1506 ± 0.0037	0.3644 ± 0.0313	0.9265 ± 0.0152	7.2620 ± 0.3554
		EBFKC w/o VCG on MIT	8.1395 ± 1.0000	0.5666 ± 0.2091	1.1923 ± 0.4556	0.7240 ± 0.1778	8.1395 ± 1.0000
		EBFKC on MIT	9.3752 ± 0.0563	0.1500 ± 0.0029	0.3403 ± 0.0113	1.0112 ± 0.0023	9.3752 ± 0.0563
		EBFKC on EAT (Ours)	$\mathbf{0.3769 \pm 0.0568}$	$\mathbf{0.0178 \pm 0.0028}$	$\mathbf{0.0461 \pm 0.0052}$	$\mathbf{0.0690 \pm 0.0065}$	$\mathbf{0.3769 \pm 0.0568}$
LJ	Annealing	Tempering SMC	22.1007 ± 3.3704	0.3914 ± 0.1333	0.4064 ± 0.1320	0.2928 ± 0.0501	22.1007 ± 3.3704
		FKC on EAT	30.7764 ± 3.4780	0.3153 ± 0.1421	0.3384 ± 0.1562	0.2553 ± 0.0397	30.7764 ± 3.4780
		EBFKC w/o VCG on MIT	3.9507 ± 1.1581	0.2068 ± 0.0420	0.2222 ± 0.0311	0.0467 ± 0.0099	4.0748 ± 1.0444
		EBFKC on MIT	3.4750 ± 0.8181	0.1891 ± 0.0747	0.2176 ± 0.0831	0.0411 ± 0.0108	3.5841 ± 0.7376
		EBFKC on EAT (Ours)	$\mathbf{0.9211 \pm 0.3994}$	$\mathbf{0.0600 \pm 0.0059}$	$\mathbf{0.0647 \pm 0.0098}$	$\mathbf{0.0121 \pm 0.0075}$	$\mathbf{0.9411 \pm 0.3991}$
	Reward tilting	Tempering SMC	14.4944 ± 3.6800	0.3141 ± 0.1519	0.3439 ± 0.1917	0.4414 ± 0.2101	14.4944 ± 3.6800
		FKC on EAT	17.5700 ± 0.4250	0.1246 ± 0.0159	0.1536 ± 0.0328	0.3275 ± 0.0051	17.5700 ± 0.4250
		EBFKC w/o VCG on MIT	7.3426 ± 1.3823	0.2026 ± 0.0798	0.2316 ± 0.0901	0.1528 ± 0.0262	7.3493 ± 1.3764
		EBFKC on MIT	6.4575 ± 0.5650	0.1441 ± 0.0356	0.1658 ± 0.0526	0.1358 ± 0.0137	6.4618 ± 0.5605
		EBFKC on EAT (Ours)	$\mathbf{0.1366 \pm 0.0832}$	$\mathbf{0.0345 \pm 0.0041}$	$\mathbf{0.0425 \pm 0.0023}$	$\mathbf{0.0140 \pm 0.0024}$	$\mathbf{0.2454 \pm 0.0837}$

Runtime. Table 8 reports the runtime of each method on the Lennard–Jones benchmark. We use the look-ahead map on LJ-13 and compute the required second-order derivatives using finite differences, whose evaluations can be parallelized across particles. These additional evaluations make the relative runtime of EBFKC on EAT longer than EBFKC on MIT with/without VCG.

Table 8: Runtime comparison on the LJ-13 tasks. Results are mean ± std over 5 runs.

Method	Annealing ($\gamma = 2.0$)	Reward tilting ($\lambda' = 0.5$)
FKC on EAT	1484.691 ± 1.749 (s)	2074.467 ± 2.337 (s)
EBFKC w/o VCG on MIT	670.970 ± 0.693 (s)	1199.705 ± 1.273 (s)
EBFKC on MIT	671.882 ± 0.804 (s)	1201.382 ± 1.070 (s)
EBFKC on EAT (Ours)	1339.710 ± 0.454 (s)	1820.829 ± 1.568 (s)

D.4 ALDP ANNEALING RESULTS

At 800 K, the pretrained model captures the broad shapes of the dihedral marginals but retains visible discrepancies from the reference (Fig. 12). After annealing to 300 K, EBFKC on EAT closely reproduces the reference marginals of both backbone dihedrals (Fig. 13). Both MIT variants underrepresent the leftmost major peak in ϕ and overconcentrate probability near the dominant positive-ψ peak.

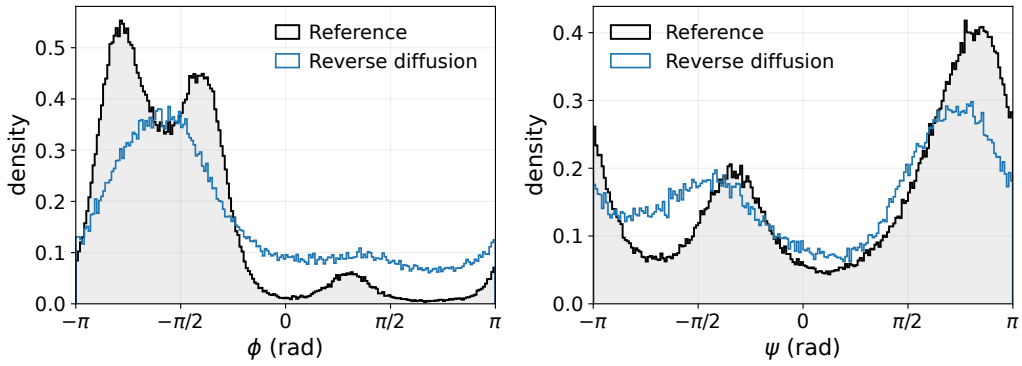


Figure 12: ALDP backbone dihedral marginals at 800 K from the reference distribution and pretrained reverse diffusion: ϕ (left) and ψ (right). The reverse diffusion marginals are estimated from 100,000 generated samples.

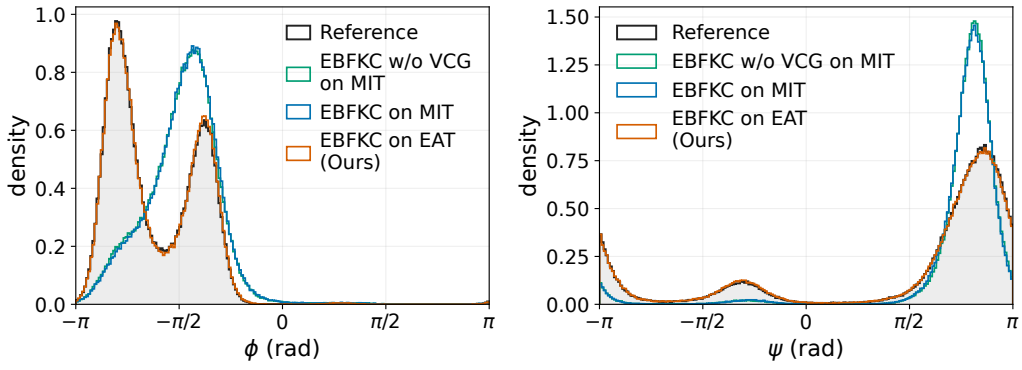


Figure 13: ALDP annealing marginals comparison for ϕ (left) and ψ (right). EBFKC on EAT closely matches the 300 K reference, whereas both MIT variants deviate substantially. The marginals are estimated from all 250,000 samples generated by each method and all reference samples at 300 K.

Table 9 reports the runtime of each method for ALDP annealing using 1000 particles. Unlike the particle system experiments, OpenMM provides only first-order derivatives of the target energy, so the second-order derivatives required by EBFKC on EAT are computed using finite-difference ap-

proximation. Because these additional OpenMM evaluations cannot be parallelized across particles, EBFKC on EAT has a longer runtime than EBFKC on MIT with/without VCG.

Table 9: Runtime comparison for ALDP annealing from 800 K to 300 K. Results are reported as mean ± std over 5 runs.

Method	Sampling time (s)
EBFKC w/o VCG on MIT	142.202 ± 0.636
EBFKC on MIT	142.932 ± 0.119
EBFKC on EAT (Ours)	508.433 ± 1.109

Look-ahead map ablation study. We ablate the denoised look-ahead map in the ALDP annealing experiment to demonstrate its necessity in high-dimensional real-world applications. The ablation reveals two concrete failure modes. *First,* with alive filtering enabled, the sampler completes all 1,000 scheduled annealing transitions but is terminated by the endpoint alive check because none of the 1,000 terminal particles satisfies the joint validity criterion: energy below the cutoff and finite OpenMM energy and force. *Second,* without alive filtering, the 1,000-particle run becomes numerically unstable at the 816th transition when the target-energy correction activates, and all 1,000 particle-wise VCG potentials become non-finite because the target energy is evaluated directly at noisy particle states. Together, these failures show that the look-ahead map is necessary for producing physically meaningful terminal-space predictions before the target energy and its gradient are evaluated by OpenMM. Alive filtering can shield the VCG calculation from occasional invalid particles, but it cannot compensate for the systematic failure caused by removing the look-ahead map.

D.5 Molecular free-energy estimation results

Figures 14 and 15 compare the dihedral marginals obtained by pretrained reverse diffusion with the reference distributions at 300 K for ALDP and Ala-4, respectively. The pretrained ALDP model captures the broad shapes of the reference marginal distributions but fails to reproduce some details, while the pretrained Ala-4 model closely reproduces the marginal distributions with some remaining discrepancies. These results indicate that both pretrained models capture the main features of the reference distributions but remain imperfect.

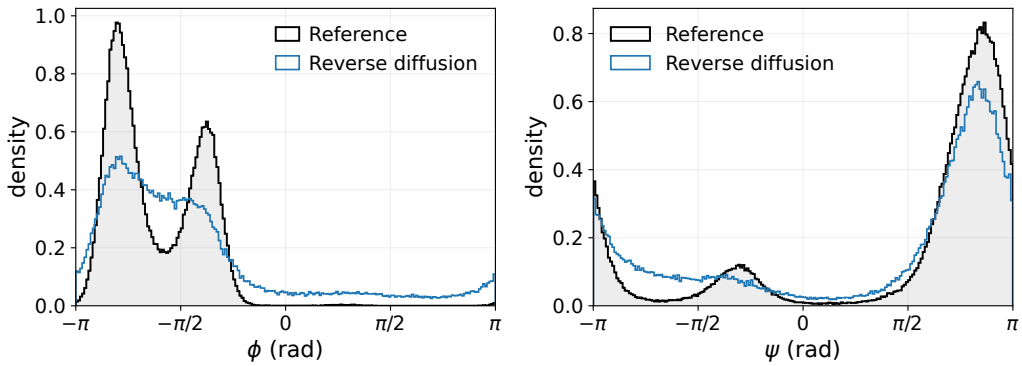


Figure 14: ALDP backbone dihedral marginals at 300 K from the reference distribution and pretrained reverse diffusion: ϕ (left) and ψ (right). The reverse diffusion marginals are estimated from 100,000 generated samples.

Fig. 16 presents the free-energy profiles estimated by EBFKC w/o VCG on MIT for ALDP and Ala-4. For ALDP, the method exhibits large estimation errors and substantial variability across runs, overestimating the free energy in the positive-ϕ region and failing to recover the reference basin there. Comparison with EBFKC on MIT in Fig. 5(a) highlights the benefit of VCG for ALDP free-energy estimation. For Ala-4, the method substantially underestimates the high free-energy barrier in the positive-ϕ_3 region.

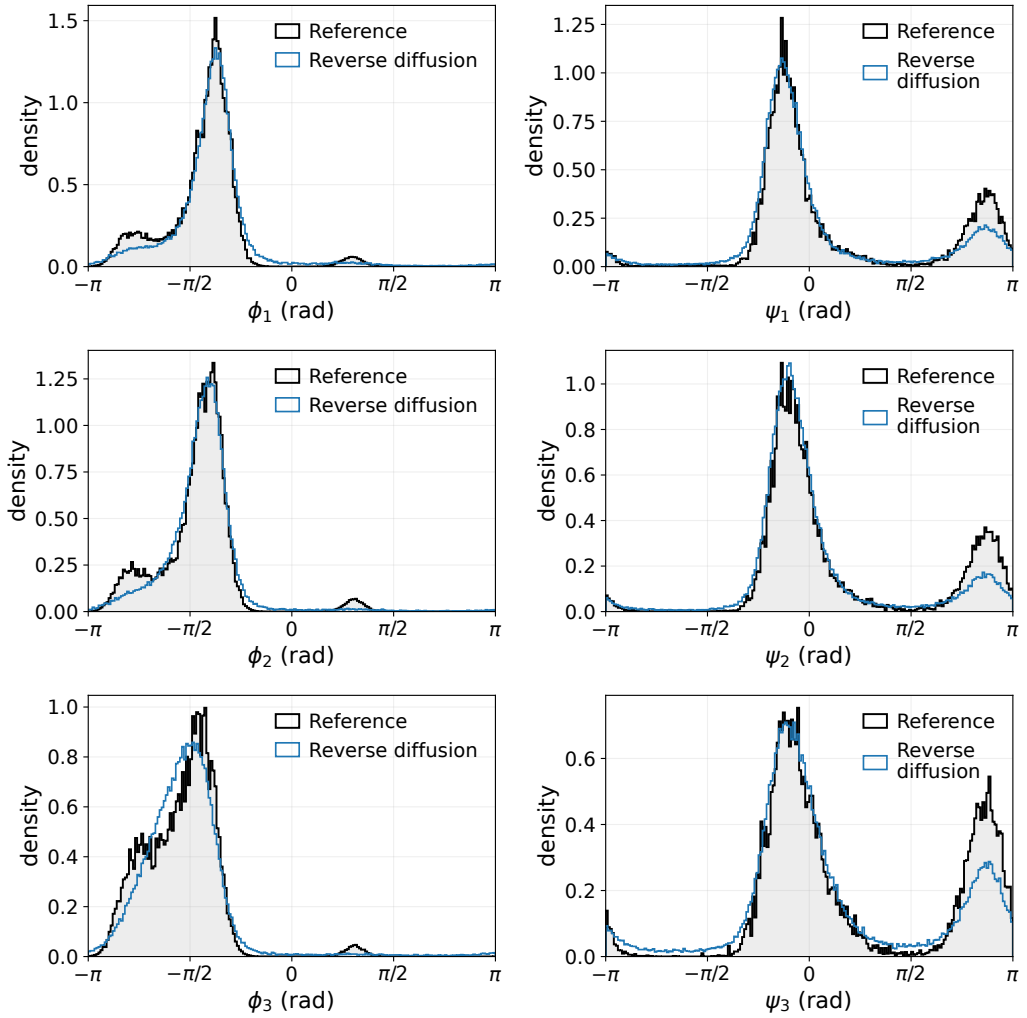
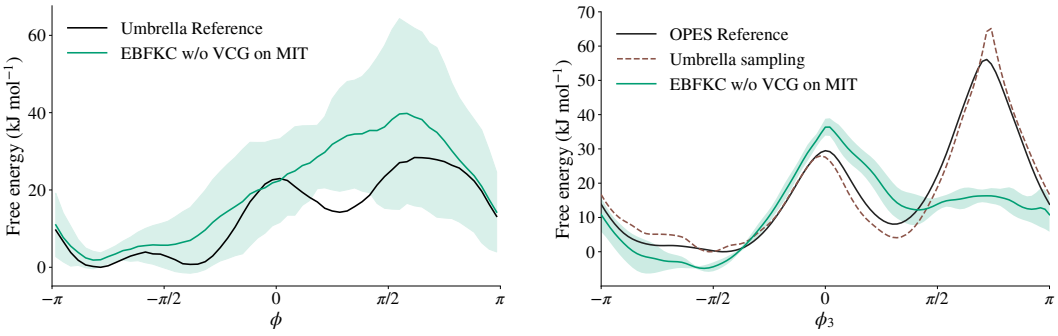


Figure 15: Ala-4 backbone dihedral marginals at 300 K from the reference distribution and pre-trained reverse diffusion. The reverse diffusion marginals are estimated from 100,000 generated samples.



(a) ALDP: compared with umbrella reference **(b) Ala-4: compared with OPES reference**

Figure 16: EBFKC w/o VCG on MIT free-energy estimates for the molecular systems. (a) ALDP along ϕ, compared with traditional umbrella sampling reference. (b) Ala-4 along ϕ_3, compared with the OPES reference and umbrella sampling results. Results are based on 5 independent runs.

Table 10 reports the runtime of each method for ALDP free-energy estimation in window 13 with 200 particles. As in the ALDP annealing experiment, the finite-difference evaluation of the second-order target-energy derivatives introduces additional OpenMM evaluations which cannot be parallelized across particles, so EBFKC on EAT has a longer runtime than EBFKC on MIT with/without VCG. Also we note that the runtime of EBFKC on MIT is slightly shorter than EBFKC w/o VCG on MIT, which is reasonable because the VCG method can reduce the variance of particle weights and thus reduce the number of resampling steps, which can save some MALA runtime.

Table 10: Runtime comparison for ALDP free-energy estimation in window 13 with 200 particles. Results are reported as mean $\pm$ std over 5 runs.

Method	Sampling time (s)
EBFKC w/o VCG on MIT	119.30 ± 10.43
EBFKC on MIT	106.17 ± 3.24
EBFKC on EAT (Ours)	269.19 ± 9.11

Table 11 reports the corresponding runtime for Ala-4 free-energy estimation in window 13 with 12 particles. Because the larger pretrained model makes each baseline sampling step more expensive, target-energy correction contributes a smaller relative runtime overhead in Ala-4 than in ALDP.

Table 11: Runtime comparison for Ala-4 free-energy estimation in window 13 with 12 particles. Results are reported as mean $\pm$ std over 5 runs.

Method	Sampling time (s)
EBFKC w/o VCG on MIT	389.05 ± 14.20
EBFKC on MIT	376.89 ± 17.25
EBFKC on EAT (Ours)	504.87 ± 12.79