

---

# Langevin Diffusion Variational Inference

---

**Tomas Geffner**

University of Massachusetts, Amherst  
tgeffner@cs.umass.edu

**Justin Domke**

University of Massachusetts, Amherst  
domke@cs.umass.edu

## Abstract

Many methods that build powerful variational distributions based on unadjusted Langevin transitions exist. Most of these were developed using a wide range of different approaches and techniques. Unfortunately, the lack of a unified analysis and derivation makes developing new methods and reasoning about existing ones a challenging task. We address this giving a single analysis that unifies and generalizes these existing techniques. The main idea is to augment the target and variational by numerically simulating the underdamped Langevin diffusion process and its time reversal. The benefits of this approach are twofold: it provides a unified formulation for many existing methods, and it simplifies the development of new ones. In fact, using our formulation we propose a new method that combines the strengths of previously existing algorithms; it uses underdamped Langevin transitions and powerful augmentations parameterized by a score network. Our empirical evaluation shows that our proposed method consistently outperforms relevant baselines in a wide range of tasks.

## 1 INTRODUCTION

Several recent work attempts to build powerful variational distributions using unadjusted Hamiltonian Monte Carlo (HMC) transition kernels (Salimans et al., 2015; Wolf et al., 2016; Caterini et al., 2018; Wu et al., 2020; Thin et al., 2021; Zhang et al., 2021; Geffner & Domke, 2021; Chen et al., 2022). In principle, one would like to use the last sample marginal of the HMC chain as variational distribution. Since this marginalization is typically intractable, these methods use auxiliary variables (Agakov & Barber,

2004); they build an augmented variational distribution that includes all samples generated in the chain, an augmented target, and perform variational inference (VI) on these augmented distributions. Training proceeds by maximizing the ELBO using unbiased reparameterization gradients, made possible by using *uncorrected* transitions.

One such method is Unadjusted Langevin Annealing (ULA) (Wu et al., 2020; Thin et al., 2021), which can be seen as an approximation of Annealed Importance Sampling (Neal, 2001; Jarzynski, 1997). The method builds a sequence of densities that gradually bridge an initial approximation to the target, and augments the variational distribution and target using uncorrected overdamped Langevin kernels targeting each of these bridging densities.

While ULA has shown good performance, it has two limitations: It is based on overdamped Langevin dynamics, which are known to suffer from random walk behavior (Neal et al., 2011, §5.2), and it augments the target using an approximation of the Annealed Importance Sampling augmentation, which is known to be suboptimal (Del Moral et al., 2006). These two limitations were addressed independently. Uncorrected Hamiltonian Annealing (UHA) (Geffner & Domke, 2021; Zhang et al., 2021) extends ULA to use *underdamped* Langevin transitions, known to improve convergence over the overdamped variant (Cheng et al., 2018). Meanwhile, Monte Carlo Diffusion (MCD) (Doucet et al., 2022) extends ULA to use better augmentations for the target. Both of these lead to significant performance improvements over ULA, albeit through orthogonal enhancements.

These methods were developed using different approaches: While UHA was developed as a differentiable approximation to Annealed Importance Sampling with underdamped Langevin transitions, MCD was developed by numerically simulating the overdamped Langevin diffusion and its time reversal, approximating intractable terms with a score network (Song & Ermon, 2019). The fact that these methods have different derivations and are based on different techniques makes it difficult to reason about their benefits, drawbacks, and the connections between them. It also means it is not obvious how to combine both of their benefits.

This paper introduces a formulation for Langevin-based

VI that encompasses previously proposed methods. This formulation can be seen as a generalization of MCD (Doucet et al., 2022) that uses underdamped Langevin dynamics, instead of the overdamped variant. Like MCD (Doucet et al., 2022), our approach is based on the analysis of continuous time processes. Its main components are the underdamped Langevin diffusion process and its time reversal, which are numerically simulated to derive the augmentations for the variational approximation and target. We introduce our approach for Langevin-based VI in section 3.

Our method is compatible with multiple numerical simulation schemes, with different choices leading to different algorithms. Section 4 introduces a simulation scheme based on splitting methods (Bou-Rabee & Owhadi, 2010; Melchionna, 2007). We show that this specific scheme can be used to recover ULA, MCD and UHA, providing a unified view for all of them, and shedding light on the connections between them, their benefits and limitations.

Additionally, our formulation facilitates the development of new methods. We use it to propose Langevin Diffusion VI (LDVI), a novel method that combines the best of UHA and MCD: it uses powerful and improved augmentations for the target, like MCD, while enjoying the benefits of underdamped Langevin transitions, like UHA. We evaluate LDVI empirically in section 5, showing that it outperforms ULA, UHA and MCD in a range of inference tasks.

Finally, we explore the importance of the numerical simulation scheme. In section 5.2 we observe that one can also develop methods using a Euler-Maruyama type discretization scheme. Our experimental results therein show that the simulation method used plays a crucial role in the algorithms' performance, suggesting a possible direction to explore to further improve these methods.

## 2 PRELIMINARIES

**Variational Inference.** VI approximates a target distribution  $p(z) = \bar{p}(z)/Z$  (known up to the normalizing constant  $Z$ ) with a simpler distribution  $q$ . It works by finding the parameters of  $q$  that maximize the evidence lower bound

$$\text{ELBO}(q(z) \parallel \bar{p}(z)) = \mathbb{E}_{q(z)} \log \frac{\bar{p}(z)}{q(z)}. \quad (1)$$

Noting that  $\log Z = \text{ELBO}(q(z) \parallel \bar{p}(z)) + \text{KL}(q(z) \parallel p(z))$ , it can be seen that maximizing the ELBO is equivalent to minimizing the KL-divergence from the approximation  $q(z)$  to the target  $p(z)$ .

**MCMC-VI.** Many methods have been developed to use MCMC to build powerful variational approximations. Ideally, one would use the last sample marginal of an MCMC chain as the approximating distribution. However, since computing this marginal is typically intractable, most methods are based on augmentations (Agakov & Barber, 2004)

and variants of Annealed Importance Sampling (Wu et al., 2020; Thin et al., 2021; Geffner & Domke, 2021; Zhang et al., 2021; Doucet et al., 2022). They define a sequence of unnormalized densities  $\bar{\pi}_k(z) = q(z)^{1-\beta_k} \bar{p}(z)^{\beta_k}$ , for  $k = 1, \dots, K-1$  and  $0 < \beta_1 < \dots < \beta_{K-1} < 1$ , forward transitions  $F_k(z_{k+1}|z_k)$  that (approximately) leave  $\pi_k$  invariant, backward transitions  $B_k(z_k|z_{k+1})$ , and build the augmented target and variational distribution as

$$\begin{aligned} q(z_{1:K}) &= q(z_1) \prod_{k=1}^{K-1} F_k(z_{k+1}|z_k) \\ \bar{p}(z_{1:K}) &= \bar{p}(z_K) \prod_{k=1}^{K-1} B_k(z_k|z_{k+1}). \end{aligned} \quad (2)$$

Then, one attempts to tune the forward and backward transitions to maximize the ELBO between these augmented distributions, equivalent to minimizing the KL divergence between them. The chain rule for the KL-divergence (Cover, 1999) then guarantees  $\text{KL}(q(z_K) \parallel p(z_K)) \leq \text{KL}(q(z_{1:K}) \parallel p(z_{1:K}))$ , justifying the use of the marginal of  $q(z_{1:K})$  over  $z_K$  to approximate the original target distribution.

While augmentations bypass intractable marginalizations, they introduce additional looseness in that  $\text{ELBO}(q(z_{1:K}) \parallel p(z_{1:K})) \leq \text{ELBO}(q(z_K) \parallel p(z_K))$ . For a given set of forward transitions  $F_k$ , this inequality can in principle be made tight by using the optimal backward transitions (Del Moral et al., 2006)

$$B_k(z_k|z_{k+1}) = F_k(z_{k+1}|z_k) \frac{q(z_k)}{q(z_{k+1})}. \quad (3)$$

In practice, however, the marginal densities  $q(z_k)$  are not exactly known, so algorithms must use other choices for  $B_k$ . There are two desiderata: the ratio  $B_k/F_k$  must be tractable (required to get a tractable expression for the ELBO between the augmented distributions), and the transitions should be differentiable (not strictly needed, but desirable, as it allows the use of reparameterization gradients to tune all parameters). Most recent methods were developed with these two properties in mind (Salimans et al., 2015; Wolf et al., 2016; Wu et al., 2020; Thin et al., 2021; Geffner & Domke, 2021; Zhang et al., 2021; Doucet et al., 2022; Jankowiak & Phan, 2021). For instance, ULA uses unadjusted overdamped Langevin kernels for both  $F_k$  and  $B_k$ , and UHA extends it to use underdamped Langevin kernels. For the latter, the distributions from eq. (2) are further augmented to include momentum variables  $\rho$ , leading to

$$\begin{aligned} q(z_{1:K}, \rho_{1:K}) &= q(z_1, \rho_1) \prod_{k=1}^{K-1} F_k(z_{k+1}, \rho_{k+1}|z_k, \rho_k) \\ \bar{p}(z_{1:K}, \rho_{1:K}) &= \bar{p}(z_K, \rho_K) \prod_{k=1}^{K-1} B_k(z_k, \rho_k|z_{k+1}, \rho_{k+1}), \end{aligned} \quad (4)$$

and to the augmented ELBO

$$\begin{aligned} & \text{ELBO}(q(z_{1:K}, \rho_{1:K}) \| \bar{p}(z_{1:K}, \rho_{1:K})) = \\ & \mathbb{E}_q \left[ \log \frac{\bar{p}(z_K, \rho_K)}{q(z_1, \rho_1)} + \sum_{k=1}^{K-1} \log \frac{F_k(z_{k+1}, \rho_{k+1} | z_k, \rho_k)}{B_k(z_k, \rho_k | z_{k+1}, \rho_{k+1})} \right]. \end{aligned} \quad (5)$$

### 3 LANGEVIN DIFFUSION VARIATIONAL INFERENCE

This section introduces our approach for Langevin-based VI. It provides a way to build the augmented distributions from eq. (4). Its main components are (1) the underdamped Langevin diffusion process and its time reversal, (2) a numerical simulation scheme to approximately simulate these processes, and (3) a score network (Doucet et al., 2022; Song & Ermon, 2019) used to approximate intractable terms in the time-reversed process. Together, these produce the forward and backward transitions  $F_k$  and  $B_k$  with a tractable ratio. Since our approach is compatible with many simulation schemes, we first introduce it in a general way, and present a specific simulation scheme in section 4.

#### 3.1 Langevin Diffusion

This sub-section introduces the Langevin diffusion process and its time reversal, which will be used to derive the forward and backward transitions in the following sections. Let  $\pi^t(z)$  be a sequence of densities bridging from the starting distribution for  $t = 0$  to the target for  $t = T$ . That is,  $\pi^0(z) = q(z)$  and  $\pi^T(z) = p(z)$ . The Langevin diffusion process is characterized by the following stochastic differential equation (SDE):

$$\begin{aligned} dz^t &= \rho^t dt \\ d\rho^t &= [\nabla \log \pi^t(z^t) - \gamma \rho^t] dt + \sqrt{2\gamma} dw^t, \end{aligned} \quad (6)$$

where  $t \in [0, T]$ ,  $w^t$  is a standard Wiener process,  $\gamma > 0$  is a friction coefficient, and  $(z^0, \rho^0) \sim q(z^0, \rho^0)$ . The forward transitions  $F_k$  will be derived by simulating this process. The motivation behind the use of this process comes from its good convergence properties. Intuitively, evolving eq. (6) yields values for  $(z^t, \rho^t)$  that tend to be close to  $\pi^t(z)\mathcal{N}(\rho|0, I)$ . Thus, one may hope that the marginal density of the process at time  $T$  is close to  $p(z)\mathcal{N}(\rho|0, I)$ , meaning the distribution of the final value  $z^T$  may be close to the target of interest.

The backward transitions  $B_k$ , on the other hand, will be derived by simulating the time-reversed SDE corresponding to eq. (6). Defining  $y^t = z^{T-t}$  and  $\lambda^t = \rho^{T-t}$ , this time-reversed process is characterized by (obtained using results for time-reversed diffusions (Anderson, 1982; Haussmann

& Pardoux, 1986), see appendix B)

$$\begin{aligned} dy^t &= -\lambda^t dt \\ d\lambda^t &= [\gamma \lambda^t - \nabla \log \pi^{T-t}(y^t) + \\ & 2\gamma \nabla_\lambda \log q^{T-t}(y^t, \lambda^t)] dt + \sqrt{2\gamma} dw^t, \end{aligned} \quad (7)$$

where  $q^t$  is the marginal of the forward process at time  $t$ . Formally, this process is initialized with  $(y^0, \lambda^0) \sim q^T(y^0, \lambda^0)$ . However, in what follows, where we use it to define the backward transitions to augment the target, it will be initialized as  $(y^0, \lambda^0) \sim p(y^0, \lambda^0)$ . The motivation for using the reverse time SDE from eq. (7) is that, under exact simulation, it yields the optimal backward transitions from eq. (3) (i.e. no additional looseness in the augmented ELBO).

#### 3.2 Transitions via SDE Simulation

The forward and backward transitions will be obtained by simulating the forward and time-reversed processes for a fixed period of time  $\delta = T/K$ . If we could simulate the above SDEs *exactly*, then

- The forward transition  $F_k(z_{k+1}, \rho_{k+1} | z_k, \rho_k)$  would be obtained by simulating the forward process from time  $t = k\delta$  up to time  $t = (k+1)\delta$ , starting from the initial values  $(z, \rho) = (z_k, \rho_k)$ ,
- The backward transition  $B_k(z_k, \rho_k | z_{k+1}, \rho_{k+1})$  would be obtained by simulating the reverse-time SDE from time  $t = (K-k)\delta$  up to time  $t = (K-k-1)\delta$ , starting from the initial values  $(y, \lambda) = (z_{k+1}, \rho_{k+1})$ .

It can be shown that these backward transitions are optimal. That is, if one could simulate eqs. (6) and (7) exactly to get the forward and backward transitions defined above, the resulting augmentations would be tight in the sense that the augmented ELBO from eq. (5) would have no additional looseness compared to an ELBO defined between the last sample marginals  $q_K(z_K, \rho_K)$  and  $\bar{p}(z_K, \rho_K)$ .

Unfortunately, these transitions are intractable for two reasons. First, the forward marginal density  $q^t$  that appears in the reverse SDE is unknown. Second, it is intractable to exactly simulate or evaluate either of the above SDEs.

**Approximating  $\nabla_\lambda \log q^{T-t}(y^t, \lambda^t)$ .** The first source of intractability of the optimal transitions is the score term  $\nabla_\lambda \log q^{T-t}(y^t, \lambda^t)$ , which is typically unavailable. Inspired by the fact that  $q^{T-t}(y, \lambda)$  is expected to be close to  $\pi^{T-t}(y)\mathcal{N}(\lambda|0, I)$ , we propose to approximate this term as

$$\nabla_\lambda \log q^{T-t}(y^t, \lambda^t) \approx -\lambda^t + s(T-t, y^t, \lambda^t), \quad (8)$$

where  $s: \mathbb{R} \times \mathbb{R}^D \times \mathbb{R}^D \rightarrow \mathbb{R}^D$  is some learnable function approximator. Following recent work (Song et al., 2020; Song & Ermon, 2019; Sohl-Dickstein et al., 2015; Ho et al., 2020; Doucet et al., 2022), we use a neural network, typically referred to as score network, which is trained with

the other parameters to maximize the ELBO. The intuition behind our approximation in eq. (8) comes from considering scenarios where the forward transitions mix fast. In such cases  $q^{T-t}(y^t, \lambda^t)$  will be close to  $\pi^{T-t}(y^t) \mathcal{N}(\lambda^t | 0, I)$ , and thus the approximation  $\nabla_\lambda \log q^{T-t}(y^t, \lambda^t) \approx -\lambda^t$  should work well. (In fact, as we show in section 4.2, several well-known methods are recovered by removing the score network; that is, fixing  $s(t, y, \lambda) = 0$ .)

**Transitions via numerical simulation.** The second source of intractability is that it is rarely possible to simulate the forward and reverse SDEs *exactly*. Thus, we use a numerical simulation scheme to *approximately* simulate them. The requirements for the simulation scheme are (1) it must yield transitions with a tractable ratio, and (2) it must be differentiable, in order to allow unbiased reparameterization gradients (Titsias & Lázaro-Gredilla, 2014; Kingma & Welling, 2013; Rezende et al., 2014). Section 4 presents a scheme that satisfies these.

### 3.3 Framework for Langevin-based VI

Our formulation for Langevin-based VI is based on the transitions described above. To get a specific instance, several choices are required:

- A momentum augmented target  $\bar{p}(z_K, \rho_K) = \bar{p}(z_K) p(\rho_K | z_K)$  that retains original target  $\bar{p}(z)$  as marginal, often defined as  $\bar{p}(z_K) \mathcal{N}(\rho_K | 0, I)$ ,
- A momentum augmented initial approximation  $q(z_1, \rho_1)$ , often defined as  $q(z_1) \mathcal{N}(\rho_1 | 0, I)$ ,
- A score network  $s(t, z, \rho)$  to approximate intractable term involving  $q^t(z, \rho)$ ,
- Forward and backward transitions  $F_k$  and  $B_k$  with a tractable ratio, obtained by numerically simulating the forward and reverse SDEs from eqs. (6) and (7).

For specific choices for these components, we can compute

$$\frac{\bar{p}(z_{1:K}, \rho_{1:K})}{q(z_{1:K}, \rho_{1:K})} = \frac{\bar{p}(z_K, \rho_K)}{q(z_1, \rho_1)} \prod_{k=1}^{K-1} \frac{B_k(z_k, \rho_k | z_{k+1}, \rho_{k+1})}{F_k(z_{k+1}, \rho_{k+1} | z_k, \rho_k)}, \quad (9)$$

required to estimate and optimize the ELBO from eq. (5).

## 4 NUMERICAL SIMULATION SCHEME

This section introduces two numerical simulation schemes, one for the forward SDE and one for the time-reversed SDE, which yield transitions with a tractable ratio. We begin by giving explicit algorithmic representations for these transitions and an expression for their ratio (section 4.1). We then explain how our formulation for Langevin-based VI with these transitions can be used to recover several existing methods, including ULA, MCD and UHA (section 4.2), and also how it can be used to derive new methods (section 4.3).

---

### Algorithm 1 Forward transition $F_k(z_{k+1}, \rho_{k+1} | z_k, \rho_k)$

---

**Require:**  $z_k, \rho_k$ , step-size  $\delta$

Re-sample momentum $\rho'_k \sim m_F(\rho'_k   \rho_k, \gamma, \delta)$	}	Leapfrog step $\tau_{\text{LP}}(z_k, \rho'_k)$
Update $\rho''_k = \rho'_k + \frac{\delta}{2} \nabla \log \pi^{k\delta}(z_k)$		
Update $z_{k+1} = z_k + \delta \rho''_k$		
Update $\rho_{k+1} = \rho''_k + \frac{\delta}{2} \nabla \log \pi^{k\delta}(z_{k+1})$		

**return**  $(z_{k+1}, \rho_{k+1})$

---



---

### Algorithm 2 Backward transition $B_k(z_k, \rho_k | z_{k+1}, \rho_{k+1})$

---

**Require:**  $z_{k+1}, \rho_{k+1}$ , step-size  $\delta$

Update $\rho''_k = \rho_{k+1} - \frac{\delta}{2} \nabla \log \pi_k(z_k)$	}	Inverse leapfrog $\tau_{\text{LP}}^{-1}(z_{k+1}, \rho_{k+1})$
Update $z_k = z_{k+1} - \delta \rho''_k$		
Update $\rho'_k = \rho''_k - \frac{\delta}{2} \nabla \log \pi_k(z_{k+1})$		

Re-sample momentum  $\rho_k \sim m_B(\rho_k | \rho'_k, z_k, \gamma, \delta)$

**return**  $(z_k, \rho_k)$

---

## 4.1 Forward and Backward Transitions

The forward transitions used to approximately simulate to forward SDE are shown in algorithm 1. They consist of two steps: (partial) momentum resampling from some distribution  $m_F$  (see section 4.1.1), followed by a single leapfrog integrator step typically used to simulate Hamiltonian dynamics (Neal et al., 2011; Betancourt, 2017) (denoted by  $\tau_{\text{LP}}$  in algorithm 1, which consists on sequential deterministic updates to the variables  $\rho, z$ , and  $\rho$ ). As explained in section 4.1.1, these transitions are derived by simulating the forward SDE from eq. (6) using splitting methods (Bou-Rabee & Owhadi, 2010; Melchionna, 2007).

The backward transitions used to approximately simulate the time-reversed SDE are shown in algorithm 2. They also consist of two steps: the *inverse* of a single leapfrog integrator step used to simulate Hamiltonian dynamics, followed by a (partial) momentum resampling from some distribution  $m_B$ . We include their derivation and details for the momentum resampling distribution  $m_B$  in section 4.1.1.

In order to use these transitions for Langevin-based VI, we need an expression for their ratio. This is given in lemma 1, proved in appendix D.

**Lemma 1.** *Let  $F_k(z_{k+1}, \rho_{k+1} | z_k, \rho_k)$  and  $B_k(z_k, \rho_k | z_{k+1}, \rho_{k+1})$  be the transitions defined in algorithms 1 and 2,  $m_F$  and  $m_B$  the momentum resampling distributions used in these transitions,  $\delta$  the discretization step-size, and  $\gamma > 0$  the damping coefficient. Then,*

$$\frac{B_k(z_k, \rho_k | z_{k+1}, \rho_{k+1})}{F_k(z_{k+1}, \rho_{k+1} | z_k, \rho_k)} = \frac{m_B(\rho_k | \rho'_k, z_k, \gamma, \delta)}{m_F(\rho'_k | \rho_k, \gamma, \delta)}, \quad (10)$$

where  $\rho'_k$  is as defined in algorithms 1 and 2, given by  $(z_k, \rho'_k) = \tau_{\text{LP}}^{-1}(z_{k+1}, \rho_{k+1})$ .

Using the transitions from algorithms 1 and 2 and their ratio given in lemma 1 we can get an exact expression for the

augmented ELBO from eq. (5). While computing this augmented ELBO exactly is typically intractable, an unbiased estimate can be obtained using a sample from  $q(z_{1:K}, \rho_{1:K})$ , as shown in algorithm 3.

---

**Algorithm 3** Generating the augmented ELBO (eq. (5)).

---

Sample  $(z_1, \rho_1) \sim q(z_1, \rho_1)$ .  
 Initialize estimator as  $\mathcal{L} \leftarrow -\log q(z_1, \rho_1)$ .  
**for**  $k = 1, 2, \dots, K - 1$  **do**  
     Run  $F_k$  (alg. 1) on  $(z_k, \rho_k)$ , store  $\rho'_k, z_{k+1}, \rho_{k+1}$ .  
     Update  $\mathcal{L} \leftarrow \mathcal{L} + \log \frac{m_B(\rho_k | \rho'_k, z_k, \gamma, \delta)}{m_F(\rho'_k | \rho_k, \gamma, \delta)}$ .  
 Update  $\mathcal{L} \leftarrow \mathcal{L} + \log \bar{p}(z_K, \rho_K)$ .  
**return**  $\mathcal{L}$

---

#### 4.1.1 Derivation of Forward and Backward Transitions

We now show the derivation for the forward and backward transitions using splitting methods (Bou-Rabee & Owhadi, 2010; Melchionna, 2007), which have been observed to work well for Langevin processes (Leimkuhler & Matthews, 2013; Monmarché, 2021). Simply put, splitting methods split an SDE into multiple simpler components, simulate each component for a time-step of size  $\delta$ , and then combine the solutions sequentially to build the  $\delta$ -sized step for the original SDE.

**Forward transitions.** These are obtained by approximately simulating the forward SDE using a splitting method. Following Monmarché (Monmarché, 2021), we split the SDE in three components,  $\mathcal{A}_F$ ,  $\mathcal{B}_F$  and  $\mathcal{O}_F$ ,<sup>1</sup>

$$\underbrace{\begin{bmatrix} dz^t \\ d\rho^t \end{bmatrix}}_{\mathcal{A}_F} = \begin{bmatrix} \rho^t dt \\ 0 \end{bmatrix}, \quad \underbrace{\begin{bmatrix} dz^t \\ d\rho^t \end{bmatrix}}_{\mathcal{B}_F} = \begin{bmatrix} 0 \\ \nabla \log \pi^t(z^t) dt \end{bmatrix},$$

$$\underbrace{\begin{bmatrix} dz^t \\ d\rho^t \end{bmatrix}}_{\mathcal{O}_F} = \begin{bmatrix} 0 \\ -\gamma \rho^t dt + \sqrt{2\gamma} dw^t \end{bmatrix},$$

each one simpler than the original SDE, and then build the forward transition by sequentially composing the simulations for components  $\mathcal{O}_F \mathcal{B}_F \mathcal{A}_F \mathcal{B}_F$ . The final forward transition shown in algorithm 1 can be obtained by noting that each of the individual components can be simulated with the following strategies:

**Simulating  $\mathcal{A}_F$ .** This can be done exactly. Given initial values  $(z^{t_0}, \rho^{t_0})$  at time  $t_0$ , simulating  $\mathcal{A}_F$  for a time  $\delta$  results in  $(z^{t_0+\delta}, \rho^{t_0+\delta}) = (z^{t_0} + \delta \rho^{t_0}, \rho^{t_0})$ .

<sup>1</sup>A similar split was used in the context of generative modeling by Dockhorn et al. (Dockhorn et al., 2021), albeit for a different (simpler) diffusion which targets a Gaussian using a different definition for the bridging densities, as typically done with diffusion models (Sohl-Dickstein et al., 2015).

**Simulating  $\mathcal{B}_F$ .** Given initial values  $(z^{t_0}, \rho^{t_0})$  at time  $t_0$ , and using that  $\pi^{t_0} \approx \pi^{t_0+\delta}$  for small  $\delta$ , simulating  $\mathcal{B}_F$  for a time  $\delta$  results in  $(z^{t_0+\delta}, \rho^{t_0+\delta}) = (z^{t_0}, \rho^{t_0} + \delta \nabla \log \pi^{t_0}(z^{t_0}))$ .

**Simulating  $\mathcal{O}_F$ .** This can be done exactly, as  $\mathcal{O}_F$  corresponds to an Ornstein–Uhlenbeck process. Given an initial value of  $\rho^{t_0}$  at time  $t_0$ , simulating  $\mathcal{O}_F$  for a time  $\delta$  gives  $\rho^{t_0+\delta} \sim \mathcal{N}(\rho^{t_0+\delta} | \eta \rho^{t_0}, (1 - \eta^2)I)$ , where  $\eta = \exp(-\gamma\delta)$ . However, as we will see next, exact simulation for the corresponding component of the reverse SDE is not possible. Thus, it may be useful to simulate  $\mathcal{O}_F$  approximately as well, using the Euler-Maruyama scheme (Maruyama, 1955; Bayram et al., 2018), which gives  $\rho^{t_0+\delta} \sim \mathcal{N}(\rho^{t_0+\delta} | \rho^{t_0}(1 - \gamma\delta), 2\gamma\delta I)$ . We use  $m_F$  to denote generically the momentum resampling distribution used, which could be any of the ones just described.

In summary, simulating  $\mathcal{O}_F$  yields the momentum resampling step, while composing the simulations for  $\mathcal{B}_F \mathcal{A}_F \mathcal{B}_F$  yields the leapfrog integration step (note that since  $\mathcal{B}_F$  is simulated twice, it is done with a step-size of  $\delta/2$ ).

**Backward transitions.** Like for the forward transitions, these are derived by splitting the reverse SDE in three components,  $\mathcal{A}_B$ ,  $\mathcal{B}_B$  and  $\mathcal{O}_B$  (using our approximation for the score term),

$$\underbrace{\begin{bmatrix} dy^t \\ d\lambda^t \end{bmatrix}}_{\mathcal{A}_B} = \begin{bmatrix} -\lambda^t dt \\ 0 \end{bmatrix}, \quad \underbrace{\begin{bmatrix} dy^t \\ d\lambda^t \end{bmatrix}}_{\mathcal{B}_B} = \begin{bmatrix} 0 \\ -\nabla \log \pi^{T-t}(y^t) dt \end{bmatrix},$$

$$\underbrace{\begin{bmatrix} dy^t \\ d\lambda^t \end{bmatrix}}_{\mathcal{O}_B} = \begin{bmatrix} 0 \\ -\gamma \lambda^t dt + 2\gamma s(T - t, y^t, \lambda^t) dt + \sqrt{2\gamma} dw^t \end{bmatrix},$$

Then, we construct the backward transition by sequentially composing the simulations for components  $\mathcal{B}_B \mathcal{A}_B \mathcal{B}_B \mathcal{O}_B$ , where the sequence  $\mathcal{B}_B \mathcal{A}_B \mathcal{B}_B$  yields the inverse of the leapfrog integrator step, and  $\mathcal{O}_B$  yields the momentum resampling step. The derivation follows the one for the forward transitions closely, with one main difference: simulating component  $\mathcal{O}_B$  has an additional difficulty, due to the presence of the term involving the score network  $s$ . While in general  $\mathcal{O}_B$  cannot be simulated exactly (unless we fix  $s = 0$ ), it can be done approximately using the Euler-Maruyama method, which results in the momentum resampling distribution  $m_B(\lambda_{t+\delta} | \lambda_t, y_t, \gamma, \delta) = \mathcal{N}(\lambda_{t+\delta} | \lambda_t(1 - \gamma\delta) + 2\gamma\delta s(T - t, y_t, \lambda_t), 2\gamma\delta I)$ . We give further details in appendix C.

## 4.2 Recovering Known Methods

As mentioned previously, ULA, MCD and UHA were originally derived using different techniques and approaches. Some of these methods use overdamped Langevin dynamics, while others use the underdamped variant; some were

derived as approximations of Annealed Importance Sampling, while others emerged from an analysis of continuous time diffusion processes. This section’s main purpose is to show that all of these methods can be derived in a unified way using the formulation for Langevin-based VI from section 3.3 with the numerical simulation schemes introduced above. We begin by briefly giving details about ULA, MCD and UHA (including their different derivations), followed by an explanation of how these methods can be recovered with our approach.

**ULA** (Wu et al., 2020; Thin et al., 2021) works directly with unadjusted overdamped Langevin kernels (i.e. no momentum variables), defining the forward transitions as

$$F_k(z_{k+1}|z_k) = \mathcal{N}(z_{k+1}|z_k + \delta \nabla \log \pi_k(z_k), 2\delta I). \quad (11)$$

Then, using the fact that  $F_k(z_{k+1}|z_k)$  is approximately reversible with respect to  $\pi_k$  when the step-size  $\delta$  is small, it defines the backward transitions as  $B_k(z_k|z_{k+1}) = F_k(z_k|z_{k+1})$ . The ratio between these transitions, and thus the augmented ELBO, are straightforward to compute (see appendix D). Broadly speaking, the method can be seen as a differentiable approximation to Annealed Importance Sampling with overdamped Langevin transitions.

**Theorem 4.1.** *ULA is recovered by the formulation from section 3.3 with  $\bar{p}(z_K, \rho_k) = \bar{p}(z_K)\mathcal{N}(\rho_K|0, I)$ ,  $q(z_1, \rho_1) = q(z_1)\mathcal{N}(\rho_1|0, I)$ ,  $s(t, z, \rho) = 0$ , and the transitions from algorithms 1 and 2 with exact momentum resampling (possible due to removing the score network) with  $\eta = 0$  (high friction limit).*

**MCD** (Doucet et al., 2022) was developed by studying the overdamped Langevin diffusion process, given by

$$dz^t = \nabla \log \pi^t(z^t)dt + \sqrt{2}dw^t. \quad (12)$$

It uses unadjusted overdamped Langevin kernels for the forward transitions (i.e. simulating eq. (12) with the Euler-Maruyama scheme), and uses backward transitions derived by simulating the reverse-time diffusion corresponding to eq. (12), also with the Euler-Maruyama scheme, using a score network to approximate intractable terms.

**Theorem 4.2.** *MCD is recovered by the formulation from section 3.3 with  $s(t, z, \rho) = \tilde{s}(t, z)$ ,  $\bar{p}(z_K, \rho_k) = \bar{p}(z_K)\mathcal{N}(\rho_K|\tilde{s}(T, z_K), I)$ ,  $q(z_1, \rho_1) = q(z_1)\mathcal{N}(\rho_1|\tilde{s}(\delta, z_1))$ , the forward transitions from algorithm 1 with exact momentum resampling for  $\eta = 0$  (high friction limit), and the backward transition from algorithm 2 using the momentum resampling distribution described in appendix D.*

**UHA** (Geffner & Domke, 2021; Zhang et al., 2021) was developed as an approximation to Annealed Importance Sampling using underdamped Langevin dynamics. It uses unadjusted underdamped Langevin kernels for the forward

transitions, and the unadjusted reversal of a Metropolis adjusted underdamped Langevin kernel for the backward transitions (simply put, Geffner and Domke (Geffner & Domke, 2021) and Zhang et al. (Zhang et al., 2021) derived an exact expression for the reversal of a Metropolis adjusted underdamped Langevin kernel, and proposed to remove the correction step to define the backward transition).

**Theorem 4.3.** *UHA is recovered by the formulation from section 3.3 with  $\bar{p}(z_K, \rho_k) = \bar{p}(z_K)\mathcal{N}(\rho_K|0, I)$ ,  $q(z_1, \rho_1) = q(z_1)\mathcal{N}(\rho_1|0, I)$ ,  $s(t, z, \rho) = 0$ , and the transitions from algorithms 1 and 2 with exact momentum resampling (possible due to removing the score network) with a learnable  $\eta$ .*

We include proofs in appendix D. All follow similar steps, we get the exact transitions and expression for the ELBO given the specific choices made in each case, and compare to that of the original method, verifying their equivalence.

### 4.3 LDVI: A New Method

Apart from recovering many existing methods, new algorithms can be derived using the proposed simulation scheme. As an example, we propose Langevin Diffusion VI (LDVI), a novel method that combines the benefits of MCD (backward transitions aided by a score network) with the benefits of UHA (underdamped dynamics). It is obtained by using the formulation from section 3.3 with  $\bar{p}(z_K, \rho_K) = \bar{p}(z_K)\mathcal{N}(\rho_K|0, I)$ ,  $q(z_1, \rho_1) = q(z_1)\mathcal{N}(\rho_1|0, I)$ , a full score network  $s(t, z, \rho)$ , and the transitions from algorithms 1 and 2 with the momentum resampling distributions

$$\begin{aligned} m_F(\rho'_k|\rho_k, \gamma, \delta) &= \mathcal{N}(\rho'_k|\rho_k(1 - \gamma\delta), 2\gamma\delta I) \\ m_B(\rho_k|\rho'_k, z_k, \gamma, \delta) &= \mathcal{N}(\rho_k|\rho'_k(1 - \gamma\delta) + \\ &\quad 2\gamma\delta s(k\delta, z_k, \rho'_k), 2\gamma\delta I), \end{aligned}$$

which are obtained by simulating components  $\mathcal{O}_F$  and  $\mathcal{O}_B$  using the Euler-Maruyama scheme (Maruyama, 1955; Bayram et al., 2018).

## 5 EXPERIMENTS

This section presents an empirical evaluation of different methods that follow our framework. We are interested in the effect that different choices have on the final performance. Specifically, we are interested in studying the benefits of using underdamped dynamics (UHA and LDVI) instead of the overdamped variant (ULA and MCD), the benefits of using powerful backward transitions aided by learnable score networks (MCD and LDVI), and the benefits of combining both improvements (LDVI) against each individual one (MCD and UHA). We explore this empirically in section 5.1.

We are also interested in how the numerical simulation scheme used affects the methods’ performance. We explore

	Ionosphere				Sonar			
	ULA	MCD	UHA	LDVI	ULA	MCD	UHA	LDVI
$K = 8$	-116.4 <sub>.05</sub>	-114.6 <sub>.01</sub>	-115.6 <sub>.05</sub>	<b>-114.4</b> <sub>.02</sub>	-122.4 <sub>.1</sub>	-117.2 <sub>.1</sub>	-120.1 <sub>.02</sub>	<b>-116.3</b> <sub>.03</sub>
$K = 16$	-115.4 <sub>.01</sub>	-113.6 <sub>.03</sub>	-114.4 <sub>.03</sub>	<b>-113.1</b> <sub>.01</sub>	-119.9 <sub>.02</sub>	-114.4 <sub>.02</sub>	-116.8 <sub>.08</sub>	<b>-112.6</b> <sub>.04</sub>
$K = 32$	-114.5 <sub>.05</sub>	-112.9 <sub>.05</sub>	-113.4 <sub>.03</sub>	<b>-112.4</b> <sub>.03</sub>	-117.4 <sub>.1</sub>	-112.4 <sub>.05</sub>	-113.9 <sub>.1</sub>	<b>-110.6</b> <sub>.08</sub>
$K = 64$	-113.8 <sub>.02</sub>	-112.5 <sub>.04</sub>	-112.8 <sub>.04</sub>	<b>-112.1</b> <sub>.01</sub>	-115.3 <sub>.05</sub>	-111.1 <sub>.7</sub>	-111.9 <sub>.02</sub>	<b>-109.7</b> <sub>.02</sub>
$K = 128$	-113.1 <sub>.04</sub>	-112.2 <sub>.02</sub>	-112.3 <sub>.02</sub>	<b>-111.9</b> <sub>.01</sub>	-113.5 <sub>.03</sub>	-110.2 <sub>.04</sub>	-110.6 <sub>.07</sub>	<b>-109.1</b> <sub>.03</sub>
$K = 256$	-112.7 <sub>.01</sub>	-112.1 <sub>.02</sub>	-112.1 <sub>.02</sub>	<b>-111.7</b> <sub>.01</sub>	-112.1 <sub>.08</sub>	-109.7 <sub>.03</sub>	-109.7 <sub>.05</sub>	<b>-108.9</b> <sub>.02</sub>

Table 1: **Combining underdamped dynamics with score networks, as done by LDVI, yields better results than all other methods for both datasets.** ELBO (higher is better, standard deviations in gray) achieved after training by different methods for different values of  $K$  for a logistic regression model with two datasets, *ionosphere* ( $d = 35$ ) and *sonar* ( $d = 61$ ). Plain VI achieves an ELBO of  $-124.1_{.15}$  nats with the *ionosphere* dataset, and  $-138.6_{.2}$  nats with the *sonar* dataset. Best result for each dataset and value of  $K$  highlighted.

	Brownian motion				Lorenz system			
	ULA	MCD	UHA	LDVI	ULA	MCD	UHA	LDVI
$K = 8$	-1.9 <sub>.05</sub>	-1.4 <sub>.06</sub>	-1.6 <sub>.03</sub>	<b>-1.1</b> <sub>.03</sub>	-1168.2 <sub>.1</sub>	-1168.1 <sub>.1</sub>	-1166.3 <sub>.1</sub>	<b>-1166.1</b> <sub>.06</sub>
$K = 16$	-1.5 <sub>.06</sub>	-0.8 <sub>.04</sub>	-1.1 <sub>.03</sub>	<b>-0.5</b> <sub>.03</sub>	-1165.7 <sub>.1</sub>	-1165.6 <sub>.1</sub>	-1163.1 <sub>.3</sub>	<b>-1162.2</b> <sub>.07</sub>
$K = 32$	-1.1 <sub>.05</sub>	-0.4 <sub>.05</sub>	-0.5 <sub>.04</sub>	<b>0.1</b> <sub>.04</sub>	-1163.2 <sub>.04</sub>	-1163.3 <sub>.04</sub>	-1160.3 <sub>.05</sub>	<b>-1157.6</b> <sub>.1</sub>
$K = 64$	-0.7 <sub>.03</sub>	-0.1 <sub>.1</sub>	0.1 <sub>.02</sub>	<b>0.5</b> <sub>.03</sub>	-1160.9 <sub>.04</sub>	-1161.1 <sub>.04</sub>	-1157.7 <sub>.05</sub>	<b>-1153.7</b> <sub>.1</sub>
$K = 128$	-0.3 <sub>.03</sub>	0.2 <sub>.04</sub>	0.4 <sub>.01</sub>	<b>0.7</b> <sub>.01</sub>	-1158.9 <sub>.05</sub>	-1158.9 <sub>.05</sub>	-1155.4 <sub>.07</sub>	<b>-1153.1</b> <sub>.1</sub>
$K = 256$	-0.1 <sub>.02</sub>	0.5 <sub>.01</sub>	0.6 <sub>.01</sub>	<b>0.9</b> <sub>.02</sub>	-1157.2 <sub>.06</sub>	-1157.1 <sub>.06</sub>	-1153.3 <sub>.1</sub>	<b>-1151.1</b> <sub>.2</sub>

Table 2: **LDVI yields better results than all other methods for both models.** ELBO (higher is better, standard deviations in gray) achieved after training by different methods for different values of  $K$  for two time series models ( $d = 32$  for the Brownian motion model and  $d = 90$  for Lorenz system). Plain VI achieves an ELBO of  $-4.4_{.02}$  nats on the Brownian motion model and  $-1187.8_{.4}$  nats on the Lorenz system model. Best result for each model and value of  $K$  highlighted.

	Random effect regression (seeds)			
	ULA	MCD	UHA	LDVI
$K = 8$	-75.5 <sub>.02</sub>	-75.1 <sub>.05</sub>	<b>-74.9</b> <sub>.01</sub>	<b>-74.9</b> <sub>.01</sub>
$K = 16$	-75.2 <sub>.04</sub>	-74.6 <sub>.04</sub>	-74.6 <sub>.01</sub>	<b>-74.5</b> <sub>.03</sub>
$K = 32$	-74.9 <sub>.03</sub>	-74.3 <sub>.03</sub>	<b>-74.2</b> <sub>.02</sub>	<b>-74.2</b> <sub>.02</sub>
$K = 64$	-74.6 <sub>.01</sub>	-74.1 <sub>.02</sub>	-74.1 <sub>.01</sub>	<b>-73.9</b> <sub>.05</sub>
$K = 128$	-74.3 <sub>.03</sub>	-73.9 <sub>.01</sub>	-73.8 <sub>.02</sub>	<b>-73.7</b> <sub>.01</sub>
$K = 256$	-74.1 <sub>.01</sub>	-73.7 <sub>.01</sub>	-73.7 <sub>.02</sub>	<b>-73.6</b> <sub>.01</sub>

Table 3: **LDVI yields better results than all other methods.** ELBO (higher is better, standard deviations in gray) achieved after training by different methods for different values of  $K$  for a random effect regression model with the *seeds* dataset ( $d = 26$ ). Plain VI achieves an ELBO of  $-77.1_{.02}$  nats. Best result for each value of  $K$  highlighted.

this in section 5.2, where we propose and evaluate empirically an alternative simulation scheme based on a simpler splitting than the one introduced in section 4.

In all cases, we use the different methods to perform inference on a wide range of tasks for values of  $K \in$

$\{8, 16, 32, 64, 128, 256\}$ , and report mean ELBO achieved after training, with standard deviations computed over three different random seeds. For all methods we set  $q(z)$  to a mean-field Gaussian, initialized to a maximizer of the ELBO, and train all parameters using Adam for 150000 steps. We repeat all simulations for the learning rates  $10^{-3}$ ,  $10^{-4}$  and  $10^{-5}$ , and keep the best one for each method and model. For all methods we tune the initial distribution  $q(z)$ , discretization step-size  $\delta$ , and the bridging densities’ parameters  $\beta$ . For LDVI and UHA we also tune the damping coefficient  $\gamma > 0$ , and for LDVI and MCD we tune the score network  $s$ , which has two hidden layers with residual connections (He et al., 2016). We implement all methods using Jax (Bradbury et al., 2018).

## 5.1 Underdamped dynamics and score networks

**Logistic Regression.** Table 1 shows results achieved by ULA, MCD, UHA and LDVI on a logistic regression model with two datasets, *ionosphere* (Sigillito et al., 1989) ( $d = 35$ ) and *sonar* (Gorman & Sejnowski, 1988) ( $d = 61$ ). It can be observed that going from overdamped to underdamped dynamics yields significant performance improvements: For the *sonar* dataset, UHA with  $K = 64$  bridging densities

	Brownian motion				Logistic regression (sonar)			
	UHA <sub>EM</sub>	UHA	LDVI <sub>EM</sub>	LDVI	UHA <sub>EM</sub>	UHA	LDVI <sub>EM</sub>	LDVI
$K = 8$	-2.8 <sub>.4</sub>	-1.6 <sub>.03</sub>	-2.8 <sub>.4</sub>	-1.1 <sub>.03</sub>	-124.1 <sub>.1</sub>	-120.1 <sub>.02</sub>	-118.5 <sub>.1</sub>	-116.3 <sub>.03</sub>
$K = 16$	-2.2 <sub>.04</sub>	-1.1 <sub>.03</sub>	-1.4 <sub>.03</sub>	-0.5 <sub>.03</sub>	-119.9 <sub>.08</sub>	-116.8 <sub>.08</sub>	-114.4 <sub>.05</sub>	-112.6 <sub>.04</sub>
$K = 32$	-1.6 <sub>.02</sub>	-0.5 <sub>.04</sub>	-0.5 <sub>.02</sub>	0.1 <sub>.04</sub>	-116.4 <sub>.1</sub>	-113.9 <sub>.1</sub>	-111.7 <sub>.04</sub>	-110.6 <sub>.08</sub>
$K = 64$	-0.9 <sub>.04</sub>	0.1 <sub>.02</sub>	0.1 <sub>.05</sub>	0.5 <sub>.03</sub>	-113.8 <sub>.1</sub>	-111.9 <sub>.02</sub>	-110.3 <sub>.04</sub>	-109.7 <sub>.02</sub>
$K = 128$	-0.4 <sub>.03</sub>	0.4 <sub>.01</sub>	0.4 <sub>.04</sub>	0.7 <sub>.01</sub>	-111.9 <sub>.1</sub>	-110.6 <sub>.07</sub>	-109.6 <sub>.04</sub>	-109.1 <sub>.03</sub>
$K = 256$	0.1 <sub>.04</sub>	0.6 <sub>.01</sub>	0.6 <sub>.05</sub>	0.9 <sub>.02</sub>	-110.7 <sub>.1</sub>	-109.7 <sub>.05</sub>	-109.1 <sub>.06</sub>	-108.9 <sub>.02</sub>

Table 4: ELBO (higher is better, standard deviations in gray) achieved by UHA<sub>EM</sub> and LDVI<sub>EM</sub>, the variants of UHA and LDVI that use the transitions from algorithms 4 and 5.

performs better than ULA with  $K = 256$ , and LDVI with  $K = 64$  performs better than MCD with  $K = 256$ . Similarly, it can be observed that the use of score networks for the backward transitions also yields significant gains: For the *sonar* dataset, MCD and LDVI with  $K = 64$  outperform ULA and UHA, respectively, with  $K = 256$ . Finally, results show that combining both improvements is beneficial as well, as it can be seen that LDVI outperforms all other methods for all datasets and values of  $K$ .

**Time series.** We consider two time series models obtained by discretizing two different SDEs, one modeling a Brownian motion with a Gaussian observation model ( $d = 32$ ), and other modeling a Lorenz system, a three-dimensional nonlinear dynamical system used to model atmospheric convection ( $d = 90$ ). We give details for these models in appendix A. Both were obtained from the ‘‘Inference Gym’’ (Sountsov et al., 2020). Results are shown in table 2. The conclusions are similar to the ones for the logistic regression models: underdamped dynamics and score networks yield gains in performance, and LDVI, which combines both improvements, performs better than all other methods.

**Random effect regression.** Table 3 shows results on a random effect regression model with the *seeds* dataset (Crowder, 1978) ( $d = 26$ ) (model details in appendix A). The same conclusions hold, both UHA and MCD perform better than ULA, and LDVI performs better than all other methods.

## 5.2 Effect of Numerical Simulation Scheme

All the methods studied so far are based on the simulation scheme introduced in section 4. We note that other simulations methods, which yield different transitions, could be used. We are interested in studying how the simulation scheme used affects methods’ performance.

We consider the forward and backward transitions shown in algorithms 4 and 5, obtained by simulating the SDEs using the Euler-Maruyama scheme as explained in appendix E (where we also give an expression for the ratio of the tran-

sitions). We propose two methods using these transitions: One that uses a full score network  $s$ , and other one that uses no score network (i.e. fixes  $s = 0$ ). Intuitively, these can be seen as variants of LDVI and UHA that use this new simulation scheme, so we term them LDVI<sub>EM</sub> and UHA<sub>EM</sub>.<sup>2</sup>

**Algorithm 4** Forward transition  $F_k(z_{k+1}, \rho_{k+1}|z_k, \rho_k)$  obtained with modified Euler-Maruyama

---

Resample momentum  $\rho_{k+1} \sim \mathcal{N}(\rho_k(1 - \gamma\delta) + \delta\nabla \log \pi_{k\delta}(z_k), 2\gamma\delta I)$ .  
 Update position  $z_{k+1} = z_k + \delta\rho_{k+1}$ .  
**return**  $(z_{k+1}, \rho_{k+1})$

---

**Algorithm 5** Backward transition  $B_k(z_k, \rho_k|z_{k+1}, \rho_{k+1})$  obtained with modified Euler-Maruyama

---

Update position  $z_k = z_{k+1} - \delta\rho_{k+1}$ .  
 Resample momentum  $\rho_k \sim \mathcal{N}(\rho_{k+1}(1 - \delta\gamma) - \delta\nabla \log \pi_{k\delta}(z_k) + 2\delta\gamma s(k\delta, z_k, \rho_{k+1}), 2\delta\gamma I)$ .  
**return**  $(z_k, \rho_k)$

---

Table 4 shows results for the Brownian motion model and the logistic regression model with the *sonar* dataset (full results in appendix E). It can be observed that each of LDVI<sub>EM</sub> and UHA<sub>EM</sub> performs worse than its counterpart using the simulation scheme from section 4, LDVI and UHA. Interestingly, for the Brownian motion model, UHA<sub>EM</sub> is also outperformed by ULA. This sheds light on the importance of the simulation scheme used: for some models, the benefits obtained by using underdamped dynamics and a score network may be lost by using a poorly chosen simulation scheme.

## Acknowledgements

This material is based upon work supported in part by the National Science Foundation under Grant No. 1908577.

<sup>2</sup>Using this simulation it is not directly clear how to take the high-friction limit. Thus, it is unclear whether they can be used to recover methods analogous to ULA and MCD.



## References

- Agakov, F. V. and Barber, D. An auxiliary variational method. In *International Conference on Neural Information Processing*, pp. 561–566. Springer, 2004.
- Anderson, B. D. Reverse-time diffusion equation models. *Stochastic Processes and their Applications*, 12(3):313–326, 1982.
- Bayram, M., Partal, T., and Buyukoz, G. O. Numerical methods for simulation of stochastic differential equations. *Advances in Difference Equations*, 2018(1):1–10, 2018.
- Betancourt, M. A conceptual introduction to hamiltonian monte carlo. *arXiv preprint arXiv:1701.02434*, 2017.
- Bou-Rabee, N. and Owhadi, H. Long-run accuracy of variational integrators in the stochastic context. *SIAM Journal on Numerical Analysis*, 48(1):278–297, 2010.
- Bradbury, J., Frostig, R., Hawkins, P., Johnson, M. J., Leary, C., Maclaurin, D., Necula, G., Paszke, A., VanderPlas, J., Wanderman-Milne, S., and Zhang, Q. JAX: composable transformations of Python+NumPy programs. 2018. URL <http://github.com/google/jax>.
- Caterini, A. L., Doucet, A., and Sejdinovic, D. Hamiltonian variational auto-encoder. In *Advances in Neural Information Processing Systems*, 2018.
- Chen, N., Xu, Z., and Campbell, T. Bayesian inference via sparse hamiltonian flows. *arXiv preprint arXiv:2203.05723*, 2022.
- Cheng, X., Chatterji, N. S., Bartlett, P. L., and Jordan, M. I. Underdamped langevin mcmc: A non-asymptotic analysis. In *Conference on Learning Theory*, pp. 300–323. PMLR, 2018.
- Cover, T. M. *Elements of information theory*. John Wiley & Sons, 1999.
- Crowder, M. J. Beta-binomial anova for proportions. *Applied statistics*, pp. 34–37, 1978.
- Del Moral, P., Doucet, A., and Jasra, A. Sequential monte carlo samplers. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(3):411–436, 2006.
- Dockhorn, T., Vahdat, A., and Kreis, K. Score-based generative modeling with critically-damped langevin diffusion. *arXiv preprint arXiv:2112.07068*, 2021.
- Doucet, A., Grathwohl, W. S., Matthews, A. G. d. G., and Strathmann, H. Annealed importance sampling meets score matching. In *ICLR Workshop on Deep Generative Models for Highly Structured Data*, 2022.
- Geffner, T. and Domke, J. Mcmc variational inference via uncorrected hamiltonian annealing. *Advances in Neural Information Processing Systems*, 34, 2021.
- Gorman, R. P. and Sejnowski, T. J. Analysis of hidden units in a layered network trained to classify sonar targets. *Neural networks*, 1(1):75–89, 1988.
- Goudie, R. J., Turner, R. M., De Angelis, D., and Thomas, A. Multibugs: a parallel implementation of the bugs modelling framework for faster bayesian inference. *Journal of statistical software*, 95, 2020.
- Haussmann, U. G. and Pardoux, E. Time reversal of diffusions. *The Annals of Probability*, pp. 1188–1205, 1986.
- He, K., Zhang, X., Ren, S., and Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.
- Ho, J., Jain, A., and Abbeel, P. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33:6840–6851, 2020.
- Jankowiak, M. and Phan, D. Surrogate likelihoods for variational annealed importance sampling. *arXiv preprint arXiv:2112.12194*, 2021.
- Jarzynski, C. Nonequilibrium equality for free energy differences. *Physical Review Letters*, 78(14):2690, 1997.
- Kingma, D. P. and Welling, M. Auto-encoding variational bayes. In *Proceedings of the International Conference on Learning Representations*, 2013.
- Leimkuhler, B. and Matthews, C. Rational construction of stochastic numerical methods for molecular sampling. *Applied Mathematics Research eXpress*, 2013(1):34–56, 2013.
- Maruyama, G. Continuous markov processes and stochastic equations. *Rendiconti del Circolo Matematico di Palermo*, 4(1):48–90, 1955.
- Melchionna, S. Design of quasisymplectic propagators for langevin dynamics. *The Journal of chemical physics*, 127(4):044108, 2007.
- Monmarché, P. High-dimensional mcmc with a standard splitting scheme for the underdamped langevin diffusion. *Electronic Journal of Statistics*, 15(2):4117–4166, 2021.
- Neal, R. M. Annealed importance sampling. *Statistics and computing*, 11(2):125–139, 2001.
- Neal, R. M. et al. Mcmc using hamiltonian dynamics. *Handbook of markov chain monte carlo*, 2(11):2, 2011.
- Rezende, D. J., Mohamed, S., and Wierstra, D. Stochastic backpropagation and approximate inference in deep generative models. In *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*, pp. 1278–1286, 2014.
- Salimans, T., Kingma, D., and Welling, M. Markov chain monte carlo and variational inference: Bridging the gap. In *International Conference on Machine Learning*, pp. 1218–1226, 2015.
- Sigillito, V. G., Wing, S. P., Hutton, L. V., and Baker, K. B. Classification of radar returns from the ionosphere using neural networks. *Johns Hopkins APL Technical Digest*, 10(3):262–266, 1989.

- Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N., and Ganguli, S. Deep unsupervised learning using nonequilibrium thermodynamics. In *International Conference on Machine Learning*, pp. 2256–2265. PMLR, 2015.
- Song, Y. and Ermon, S. Generative modeling by estimating gradients of the data distribution. *Advances in Neural Information Processing Systems*, 32, 2019.
- Song, Y., Sohl-Dickstein, J., Kingma, D. P., Kumar, A., Ermon, S., and Poole, B. Score-based generative modeling through stochastic differential equations. *arXiv preprint arXiv:2011.13456*, 2020.
- Sountsov, P., Radul, A., and contributors. Inference gym. 2020. URL [https://pypi.org/project/inference\\_gym](https://pypi.org/project/inference_gym).
- Thin, A., Kotelevskii, N., Doucet, A., Durmus, A., Moulines, E., and Panov, M. Monte carlo variational auto-encoders. In *International Conference on Machine Learning*, pp. 10247–10257. PMLR, 2021.
- Titsias, M. and Lázaro-Gredilla, M. Doubly stochastic variational bayes for non-conjugate inference. In *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*, pp. 1971–1979, 2014.
- Wolf, C., Karl, M., and van der Smagt, P. Variational inference with hamiltonian monte carlo. *arXiv preprint arXiv:1609.08203*, 2016.
- Wu, H., Köhler, J., and Noé, F. Stochastic normalizing flows. *Advances in Neural Information Processing Systems*, 33: 5933–5944, 2020.
- Zhang, G., Hsu, K., Li, J., Finn, C., and Grosse, R. B. Differentiable annealed importance sampling and the perils of gradient noise. *Advances in Neural Information Processing Systems*, 34, 2021.

## A MODEL DETAILS

### A.1 Time Series Models

This models are presented following their descriptions in the Inference Gym documentation (Sountsov et al., 2020). Both are obtained by discretizing an SDE and using a Gaussian observation model.

**Brownian motion with Gaussian observation noise** The model is given by

$$\begin{aligned} \alpha_{\text{inn}} &\sim \text{LogNormal}(\text{loc} = 0, \text{scale} = 2) \\ \alpha_{\text{obs}} &\sim \text{LogNormal}(\text{loc} = 0, \text{scale} = 2) \\ x_1 &\sim \mathcal{N}(\text{loc} = 0, \text{scale} = \alpha_{\text{inn}}) \\ x_i &\sim \mathcal{N}(\text{loc} = x_{i-1}, \text{scale} = \alpha_{\text{inn}}) && i = 2, \dots, 30 \\ y_i &\sim \mathcal{N}(\text{loc} = x_i, \text{scale} = \alpha_{\text{obs}}) && i = 1, \dots, 30. \end{aligned}$$

The goal is to do inference over variables  $\alpha_{\text{inn}}$ ,  $\alpha_{\text{obs}}$  and  $x_i$  ( $i = 1, \dots, 30$ ), given the observations  $y_i$ , for  $i \in \{1, \dots, 10\} \cup \{20, \dots, 30\}$  (i.e. the ten middle observations are missing).

**Lorenz system** The model is given by

$$\begin{aligned} x_1 &\sim \mathcal{N}(\text{loc} = 0, \text{scale} = 1) \\ y_1 &\sim \mathcal{N}(\text{loc} = 0, \text{scale} = 1) \\ z_1 &\sim \mathcal{N}(\text{loc} = 0, \text{scale} = 1) \\ x_i &\sim \mathcal{N}(\text{loc} = 10(y_{i-1} - x_{i-1}), \text{scale} = \alpha_{\text{inn}}) && i = 2, \dots, 30 \\ y_i &\sim \mathcal{N}(\text{loc} = x_{i-1}(28 - z_{i-1}) - y_{i-1}, \text{scale} = \alpha_{\text{inn}}) && i = 2, \dots, 30 \\ z_i &\sim \mathcal{N}(\text{loc} = x_{i-1}y_{i-1} - \frac{8}{3}z_{i-1}, \text{scale} = \alpha_{\text{inn}}) && i = 2, \dots, 30, \\ o_i &\sim \mathcal{N}(\text{loc} = x_i, \text{scale} = 1) && i = 2, \dots, 30, \end{aligned}$$

where  $\alpha_{\text{inn}} = 0.1$  (determined by the discretization step-size used for the original SDE). The goal is to do inference over  $x_i, y_i, z_i$  for  $i = 1, \dots, 30$ , given observed values  $o_i$  for  $i \in \{1, \dots, 10\} \cup \{20, \dots, 30\}$ .

### A.2 Random effect regression

This model can be found in the MultiBUGS (Goudie et al., 2020) documentation. It is essentially a random effects regression model, given by

$$\begin{aligned} \tau &\sim \text{Gamma}(0.01, 0.01) \\ a_0 &\sim \mathcal{N}(0, 10) \\ a_1 &\sim \mathcal{N}(0, 10) \\ a_2 &\sim \mathcal{N}(0, 10) \\ a_{12} &\sim \mathcal{N}(0, 10) \\ b_i &\sim \mathcal{N}\left(0, \frac{1}{\sqrt{\tau}}\right) && i = 1, \dots, 21 \\ \text{logits}_i &= a_0 + a_1 x_i + a_2 y_i + a_{12} x_i y_i + b_i && i = 1, \dots, 21 \\ r_i &\sim \text{Binomial}(\text{logits}_i, N_i) && i = 1, \dots, 21. \end{aligned}$$

The goal is to do inference over the variables  $\tau, a_0, a_1, a_2, a_{12}$  and  $b_i$  for  $i = 1, \dots, 21$ , given observed values for  $x_i, y_i$  and  $N_i$ . The data used was obtained from (Crowder, 1978), which models the germination proportion of seeds arranged in a factorial layout by seed and type of root.

## B TIME-REVERSED SDE FOR UNDERDAMPED LANGEVIN PROCESS

Consider a diffusion process of the form

$$dx^t = f(z^t, t)dt + G(z^t, t)dw^t, \quad (13)$$

where  $t \in 0, T$ . Defining  $u^t = x^{T-t}$ , the process that inverts the one above is given by (Haussmann & Pardoux, 1986; Anderson, 1982; Dockhorn et al., 2021)

$$du^t = [-f(u^t, T-t) + G(u^t, T-t)G(u^t, T-t)^\top \nabla \log q^{T-t}(u^t)] dt + G(u^t, T-t)dw^t. \quad (14)$$

This result can be used to derive the time-reversed diffusion for the underdamped Langevin process, by expressing the forward process as (using  $x = [z, \rho]$ )

$$\begin{bmatrix} dz^t \\ d\rho^t \end{bmatrix} = \underbrace{\begin{bmatrix} \rho^t \\ \nabla \log \pi^t(z^t) - \gamma \rho^t \end{bmatrix}}_{f(x^t, t)} dt + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & \sqrt{2\gamma} \end{bmatrix}}_{G(x^t, t)} \begin{bmatrix} dw_1^t \\ dw_2^t \end{bmatrix}, \quad (15)$$

and applying eq. (14).

## C DERIVATION OF BACKWARD TRANSITIONS FROM ALGORITHM 2

The time-reversed SDE is split into three components as

$$\begin{bmatrix} dy^t \\ d\lambda^t \end{bmatrix} = \underbrace{\begin{bmatrix} -\lambda^t dt \\ 0 \end{bmatrix}}_{\mathcal{A}_B} + \underbrace{\begin{bmatrix} 0 \\ -\nabla \log \pi^{T-t}(y^t) dt \end{bmatrix}}_{\mathcal{B}_B} + \underbrace{\begin{bmatrix} 0 \\ -\gamma \rho^t dt + 2\gamma s(T-t, y^t, \lambda^t) dt + \sqrt{2\gamma} dw^t \end{bmatrix}}_{\mathcal{O}_B}.$$

Then, the final transitions are given by sequentially composing the simulations for components  $\mathcal{B}_B \mathcal{A}_B \mathcal{B}_B \mathcal{O}_B$ .

**Simulating  $\mathcal{A}_B$ :** This can be done exactly. Given initial values  $(y^{t_0}, \lambda^{t_0})$  at time  $t_0$ , simulating  $\mathcal{A}_B$  for a time  $\delta$  results in  $(y^{t_0+\delta}, \lambda^{t_0+\delta}) = (y^{t_0} + \delta \lambda^{t_0}, \lambda^{t_0})$ .

**Simulating  $\mathcal{B}_B$ :** Given initial values  $(y^{t_0}, \lambda^{t_0})$  at time  $t_0$ , and using that  $\pi^{T-t_0} \approx \pi^{T-t_0+\delta}$  for small  $\delta$ , simulating  $\mathcal{B}_B$  for a time  $\delta$  results in  $(y^{t_0+\delta}, \lambda^{t_0+\delta}) = (y^{t_0}, \lambda^{t_0} + \delta \nabla \log \pi^{T-t_0}(y^{t_0}))$ .

**Simulating  $\mathcal{O}_B$ :** In contrast to the forward transitions, this component cannot be simulated exactly due to the presence of the term involving  $s$  (it can be done exactly if we fix  $s = 0$ ). This component can be simulated approximately, for instance, using the Euler-Maruyama scheme. Given an initial values  $y^{t_0}$  and  $\lambda^{t_0}$  at time  $t_0$ , simulating  $\mathcal{O}_B$  for a time  $\delta$  gives  $\lambda^{t_0+\delta} \sim \mathcal{N}(\lambda^{t_0+\delta} | \lambda^{t_0} (1 - \gamma \delta) + 2\gamma \delta s(T-t, y^{t_0}), 2\gamma \delta I)$ . We will use  $m_B$  to denote generically the momentum resampling distribution used.

Finally, the backward transitions from algorithm 2 are obtained by sequentially combining the results above, and transforming back to the variables  $z$  and  $\rho$  using the fact that  $z^t = y^{T-t}$  and  $\rho^t = \lambda^{T-t}$ .

## D PROOFS

### D.1 Proof of lemma 1

This proof follows closely the one from Geffner and Domke (Geffner & Domke, 2021). We will derive expressions for the forward and backward transitions separately, and take the ratio at the end. In the derivation, we replace all delta functions by Gaussians with variance  $\Delta$ , and take the limit  $\Delta \rightarrow 0$  with the final expressions. We use  $(z', \rho') = \tau_{\text{LP},k}(z, \rho)$  to denote the leapfrog integration step typically used by HMC targeting the bridging density  $\pi_k$  (see algorithm 1 for a definition), which is invertible and volume preserving (Jacobian determinant equals one).

**Forward transitions** We will divide the forward transition from algorithm 1 in two steps:

1. Resampling step:  $z'_k, \rho'_k \sim m_F(\rho'_k | \rho_k, \gamma, \delta) \mathcal{N}(z'_k | z_k, \Delta)$ ,
2. Deterministic step:  $(z_{k+1}, \rho_{k+1}) = \tau_{\text{LP},k}(z'_k, \rho'_k)$ .

Using that  $\tau_{\text{LP},k}$  is invertible and volume preserving, we get

$$F_k(z_{k+1}, \rho_{k+1} | z_k, \rho_k) = m_F \left( (\tau_{\text{LP},k}^{-1})^\rho(z_{k+1}, \rho_{k+1}) \middle| \rho_k, \gamma, \delta \right) \mathcal{N} \left( (\tau_{\text{LP},k}^{-1})^z(z_{k+1}, \rho_{k+1}) \middle| z_k, \Delta \right), \quad (16)$$

where  $(\tau_{\text{LP},k}^{-1})^\rho(z_{k+1}, \rho_{k+1})$  is defined as the second component of  $\tau_{\text{LP},k}^{-1}(z_{k+1}, \rho_{k+1})$ , and similarly  $(\tau_{\text{LP},k}^{-1})^z(z_{k+1}, \rho_{k+1})$  as its first component.

**Backward transitions** Similarly, we divide the backward transitions from algorithm 2 in two steps:

1. Deterministic step:  $(z'_k, \rho'_k) = \tau_{\text{LP},k}^{-1}(z_{k+1}, \rho_{k+1})$ ,
2. Resampling step:  $z_k, \rho_k \sim \mathcal{N}(z_k | z'_k, \Delta) m_B(\rho_k | \rho'_k, z_k, \gamma, \delta)$ .

Then, we get

$$B_k(z_k, \rho_k | z_{k+1}, \rho_{k+1}) = \mathcal{N} \left( z_k \middle| (\tau_{\text{LP},k}^{-1})^z(z_{k+1}, \rho_{k+1}), \Delta \right) m_B \left( \rho_k \middle| (\tau_{\text{LP},k}^{-1})^\rho(z_{k+1}, \rho_{k+1}), z_k, \gamma, \delta \right). \quad (17)$$

Finally, using  $\rho'_k = (\tau_{\text{LP},k}^{-1})^\rho(z_{k+1}, \rho_{k+1})$  and  $z'_k = (\tau_{\text{LP},k}^{-1})^z(z_{k+1}, \rho_{k+1})$  to simplify notation, taking the ratio between  $F_k$  and  $B_k$  yields

$$\frac{B_k(z_k, \rho_k | z_{k+1}, \rho_{k+1})}{F_k(z_{k+1}, \rho_{k+1} | z_k, \rho_k)} = \frac{m_B(\rho_k | \rho'_k, z_k, \gamma, \delta)}{m_F(\rho'_k | \rho_k, \gamma, \delta)}, \quad (18)$$

since the ratio between the Gaussian densities cancel.

## D.2 Proof of theorem 4.1

**ULA summary** ULA uses the following transitions

$$\begin{aligned} F_k(z_{k+1}|z_k) &= \mathcal{N}(z_{k+1}|z_k + \epsilon \nabla \log \pi_k(z_k), 2\epsilon) \\ B_k(z_k|z_{k+1}) &= \mathcal{N}(z_k|z_{k+1} + \epsilon \nabla \log \pi_k(z_{k+1}), 2\epsilon) \end{aligned} \quad (19)$$

The forward transition yields

$$z_{k+1} = z_k + \epsilon \nabla \log \pi_k(z_k) + \sqrt{2\epsilon} \xi_k, \quad (20)$$

where  $\xi_k \sim \mathcal{N}(0, I)$ , for  $k = 1, \dots, K - 1$ .

Then, ULA's lower bound can be written as (using  $\mathcal{N}(a|b, c)$  with  $c \geq 0$  to denote the pdf of a Gaussian with mean  $b$  and variance  $c$  evaluated at  $a$ )

$$\frac{p(z_K)}{q(z_1)} \prod_{k=1}^{K-1} \frac{B_k(z_k|z_{k+1})}{F_k(z_{k+1}|z_k)} = \frac{p(z_K)}{q(z_1)} \prod_{k=1}^{K-1} \frac{\mathcal{N}(z_k|z_{k+1} + \epsilon \nabla \log \pi_k(z_{k+1}), 2\epsilon)}{\mathcal{N}(z_{k+1}|z_k + \epsilon \nabla \log \pi_k(z_k), 2\epsilon)} \quad (21)$$

$$= \frac{p(z_K)}{q(z_1)} \prod_{k=1}^{K-1} \frac{\mathcal{N}(-\epsilon \nabla \log \pi_k(z_k) - \epsilon \nabla \log \pi_k(z_{k+1}) - \sqrt{2\epsilon} \xi_k | 0, 2\epsilon)}{\mathcal{N}(\sqrt{2\epsilon} \xi_k | 0, 2\epsilon)} \quad (22)$$

$$= \frac{p(z_K)}{q(z_1)} \prod_{k=1}^{K-1} \frac{\mathcal{N}(\sqrt{\frac{\epsilon}{2}} \log \pi_k(z_k) + \sqrt{\frac{\epsilon}{2}} \nabla \log \pi_k(z_{k+1}) + \xi_k | 0, 1)}{\mathcal{N}(\xi_k | 0, 1)}, \quad (23)$$

where we obtain eq. (22) using the expression for  $z_{k+1}$  from eq. (20).

**Recovering ULA** Using exact momentum resampling for the forward transitions gives  $\rho'_k \sim \mathcal{N}(\eta \rho_k, (1 - \eta^2)I)$ , where  $\eta = \exp(-\gamma\delta)$ . Using  $\eta = 0$  (high friction limit) gives  $\rho'_k \sim \mathcal{N}(0, I)$ , which is used in place of  $m_F(\rho'_k|\rho_k, \gamma, \delta)$  in algorithm 1. The final forward transitions  $F_k(z_{k+1}, \rho_{k+1}|z_k, \rho_k)$  are thus given by

$$\begin{aligned} \text{Simulate } \mathcal{O}_F: \quad \rho'_k &\sim \mathcal{N}(0, I) && \rightarrow \rho'_k = \xi_k \sim \mathcal{N}(0, I) \\ \text{Simulate } \mathcal{B}_F: \quad \rho''_k &= \rho'_k + \frac{\delta}{2} \nabla \log \pi_k(z_k) && \rightarrow \rho''_k = \xi_k + \frac{\delta}{2} \nabla \log \pi_k(z_k) \\ \text{Simulate } \mathcal{A}_F: \quad z_{k+1} &= z_k + \delta \rho''_k && \rightarrow z_{k+1} = z_k + \frac{\delta^2}{2} \nabla \log \pi_k(z_k) + \delta \xi_k \\ \text{Simulate } \mathcal{B}_F: \quad \rho_{k+1} &= \rho''_k + \frac{\delta}{2} \nabla \log \pi_k(z_{k+1}) && \rightarrow \rho_{k+1} = \xi_k + \frac{\delta}{2} \nabla \log \pi_k(z_k) + \frac{\delta}{2} \nabla \log \pi_k(z_{k+1}). \end{aligned} \quad (24)$$

It can be seen that the forward dynamics for  $z$  are exactly the same as those used by ULA (eq. (20)), using  $\epsilon = \delta^2/2$ .

Thanks to removing the score network (i.e.  $s(t, z, \rho) = 0$ ), exact momentum resampling is possible for the backward transitions  $B_k(z_k, \rho_k|z_{k+1}, \rho_{k+1})$  as well. Similarly to the forward transitions, this gives  $\rho_k \sim \mathcal{N}(0, I)$ , which should be used in place of  $m_B(\rho_k|\rho'_k, z_k, \gamma, \delta)$  in algorithm 2.

Finally, using the transitions with these resampling distributions, the momentum augmented distributions given by  $\bar{p}(z_K, \rho_k) = \bar{p}(z_K) \mathcal{N}(\rho_k|0, I)$  and  $q(z_1, \rho_1) = q(z_1) \mathcal{N}(\rho_1|0, I)$ , and the result from lemma 1, we get

$$\frac{p(z_{1:K}, \rho_{1:K})}{q(z_{1:K}, \rho_{1:K})} = \frac{p(z_K)}{q(z_1)} \prod_{k=1}^{K-1} \frac{\mathcal{N}(\rho_{k+1}|0, I)}{\mathcal{N}(\rho'_k|0, I)} \quad (25)$$

$$= \frac{p(z_K)}{q(z_1)} \prod_{k=1}^{K-1} \frac{\mathcal{N}(\xi_k + \frac{\delta}{2} \nabla \log \pi_k(z_k) + \frac{\delta}{2} \nabla \log \pi_k(z_{k+1}) | 0, I)}{\mathcal{N}(\xi_k | 0, I)}, \quad (26)$$

where the second line is obtained by replacing  $\rho_{k+1}$  and  $\rho'_k$  by their respective expressions from eq. (24). Taking  $\epsilon = \delta^2/2$  gives the ratio used by ULA (eq. (23)), showing that our framework with the choices from theorem 4.1 recovers ULA.

### D.3 Proof of theorem 4.2

**MCD summary** MCD uses the same forward transitions as ULA

$$F_k(z_{k+1}|z_k) = \mathcal{N}(z_k + \epsilon \nabla \log \pi_k(z_k), 2\epsilon), \quad (27)$$

and backward transitions given by

$$B_k(z_k|z_{k+1}) = \mathcal{N}(z_{k+1} + \epsilon \nabla \log \pi_k(z_{k+1}) + 2\epsilon \tilde{s}(k\delta, z_{k+1}), 2\epsilon). \quad (28)$$

Using these transitions, and writing  $z_{k+1} = z_k + \epsilon \nabla \log \pi_k(z_k) + \sqrt{2\epsilon} \xi_k$  (where  $\xi_k \sim \mathcal{N}(0, I)$ ), MCD yields the ratio

$$\begin{aligned} \frac{\bar{p}(z_{1:K})}{q(z_{1:K})} &= \frac{p(z_K)}{q(z_1)} \prod_{k=1}^{K-1} \frac{B_k(z_k|z_{k+1})}{F_k(z_{k+1}|z_k)} \\ &= \frac{\bar{p}(z_K)}{q(z_1)} \prod_{k=1}^{K-1} \frac{\mathcal{N}(\sqrt{\frac{\epsilon}{2}} \nabla \log \pi_k(z_k) + \sqrt{\frac{\epsilon}{2}} \nabla \log \pi_k(z_{k+1}) + \sqrt{2\epsilon} \tilde{s}(k\delta, z_{k+1}) + \xi_k | 0, 1)}{\mathcal{N}(\xi_k | 0, 1)}. \end{aligned} \quad (29)$$

**Recovering MCD** The forward transitions used by our framework to recover MCD are exactly the same as the ones used in the proof of theorem 4.1, shown in eq. (24). Therefore, the forward dynamics for  $z$  are given by  $z_{k+1} = z_k + \frac{\delta^2}{2} \nabla \log \pi_k(z_k) + \delta \xi_k$ , which is exactly the same as the forward dynamics used by MCD, taking  $\epsilon = \delta^2/2$ .

Deriving the backward transitions requires simulating component  $\mathcal{O}_B$ . Using  $s(T-t, y_t, \lambda_t) = s(T-t, y_t)$  (as stated in theorem 4.2), this component is given by

$$\begin{bmatrix} dy^t \\ d\lambda^t \end{bmatrix} = \begin{bmatrix} 0 \\ -\gamma \rho^t dt + 2\gamma s(T-t, y^t) dt + \sqrt{2\gamma} dw^t \end{bmatrix}. \quad (30)$$

Since  $dy^t = 0$ , we get that  $y^t$  is constant as a function of  $t$ . However, the term  $s(T-t, y^t)$  is not a constant with respect to time, due to its first argument. Thus, in general, exact simulation for this component is intractable. However, approximating  $s(T-t, y^t) \approx s(T-t_0, y^{t_0})$  for  $t \in [t_0, t_0 + \delta]$ , we can simulate it as  $\lambda^{t_0+\delta} \sim \mathcal{N}(\eta \lambda^{t_0} + 2s(T-t_0, y^t)(1-\eta), (1-\eta^2)I)$ , where  $\eta = \exp(-\gamma\delta)$ .<sup>3</sup> Taking  $\eta = 0$  and expressing the transitions in terms of  $z$  and  $\rho$  yields the backward transitions  $B_k(z_k, \rho_k | z_{k+1}, \rho_{k+1})$  from algorithm 2 with momentum resampling distribution given by  $m_B(\rho_k | \rho'_k, z_k, \gamma, \delta) = \mathcal{N}(2s(k\delta, z_k), I)$ .

Using the above transitions, the momentum augmented distributions given by

$$\bar{p}(z_K, \rho_K) = \bar{p}(z_K) \mathcal{N}(\rho_K | 2s(K\delta, z_K), I) \quad \text{and} \quad q(z_1, \rho_1) = q(z_1) \mathcal{N}(\rho_1 | 2s(\delta, z_1), I),$$

and the result from lemma 1, our framework yields

$$\frac{\bar{p}(z_{1:K}, \rho_{1:K})}{q(z_{1:K}, \rho_{1:K})} = \frac{\bar{p}(z_K) \mathcal{N}(\rho_K | 2s(K\delta, z_K), I)}{q(z_1) \mathcal{N}(\rho_1 | 2s(\delta, z_1), I)} \prod_{k=1}^{K-1} \frac{\mathcal{N}(\rho_k | 2s(k\delta, z_k), I)}{\mathcal{N}(\rho'_k | 0, I)}. \quad (31)$$

Finally, replacing  $\rho_{k+1}$  and  $\rho'_k$  by their expressions from eq. (24), this ratio can be written as

$$\frac{\bar{p}(z_{1:K}, \rho_{1:K})}{q(z_{1:K}, \rho_{1:K})} = \frac{p(z_K)}{q(z_1)} \prod_{k=1}^{K-1} \frac{\mathcal{N}(\xi_k + \frac{\delta}{2} \nabla \log \pi_k(z_k) + \frac{\delta}{2} \nabla \log \pi_k(z_{k+1}) - 2s((k+1)\delta, z_{k+1}) | 0, I)}{\mathcal{N}(\xi_k | 0, I)}, \quad (32)$$

which recovers the ratio used by MCD taking  $\epsilon = \frac{\delta^2}{2}$  and  $s((k+1)\delta, z_{k+1}) = -\frac{1}{2} \sqrt{2\epsilon} \tilde{s}(k\delta, z_{k+1})$ . This shows that our framework with the choices from theorem 4.2 recovers MCD.

<sup>3</sup>This is obtained by noting that the process  $d\lambda^t = a dt - \gamma \lambda^t dt + \sqrt{2\gamma} dw_t$ , where  $a$  is a constant, admits exact simulation as  $\lambda^{t_0+\delta} \sim \mathcal{N}(\lambda^{t_0} \eta + \frac{a}{\gamma}(1-\eta), (1-\eta^2)I)$ , where  $\eta = \exp(-\gamma\delta)$ . Our result follows from setting  $a = \gamma s(T-t_0, y^{t_0})$ .

#### D.4 Proof of theorem 4.3

For this proof we will use  $\tau_{\text{LP}}(z, \rho)$  to denote a single step of the leapfrog integrator typically used by HMC (see algorithm 1 for the definition), and  $\gamma(z, \rho)$  to denote the operator that negates the momentum variables, that is,  $(z, -\rho) = \gamma(z, \rho)$ .

UHA uses forward transitions  $F_k(z_{k+1}, \rho_{k+1} | z_k, \rho_k)$  that consist of three steps

1. Resample momentum as  $\rho'_k \sim \mathcal{N}(\eta\rho_k, (1 - \eta^2)I)$ , where  $\eta = \exp(-\gamma\delta)$ ,
2. Apply a leapfrog step  $\tau_{\text{LP}}$  followed by a negation of the momentum, which gives  $(z''_k, \rho''_k) = (\gamma \circ \tau_{\text{LP}})(z_k, \rho'_k)$ ,
3. Negate the momentum, which gives  $(z_{k+1}, \rho_{k+1}) = \gamma(z''_k, \rho''_k)$ .

Combining steps (2) and (3) gives  $(z_{k+1}, \rho_{k+1}) = (\gamma \circ \gamma \circ \tau_{\text{LP}})(z_k, \rho'_k)$ , which can be simplified to  $(z_{k+1}, \rho_{k+1}) = \tau_{\text{LP}}(z_k, \rho'_k)$ , since  $\gamma$  is an involution (self-inverting). Thus, UHA transitions can be expressed as a sequence of two steps: momentum resampling, followed by an application of a leapfrog step used by HMC. This is exactly the same as the forward transition from algorithm 1 with exact momentum resampling.

Similarly, the backward transitions  $B_k(z_k, \rho_k | z_{k+1}, \rho_{k+1})$  used by UHA also consist of three steps, given by

1. Negate the momentum, which gives  $(z''_k, \rho''_k) = \gamma(z_{k+1}, \rho_{k+1})$ .
2. Apply a leapfrog step  $\tau_{\text{LP}}$  followed by a negation of the momentum, which gives  $(z_k, \rho'_k) = (\gamma \circ \tau_{\text{LP}})(z''_k, \rho''_k)$ ,
3. Resample momentum as  $\rho_k \sim \mathcal{N}(\eta\rho'_k, (1 - \eta^2)I)$ , where  $\eta = \exp(-\gamma\delta)$ .

Using the fact that  $\gamma \circ \tau_{\text{LP}} = (\gamma \circ \tau_{\text{LP}})^{-1} = \tau_{\text{LP}}^{-1} \circ \gamma^{-1}$  (Neal et al., 2011), steps (1) and (2) above can be combined as  $((\gamma \circ \tau_{\text{LP}}) \circ \gamma)(z_{k+1}, \rho_{k+1}) = (\tau_{\text{LP}}^{-1} \circ \gamma^{-1} \circ \gamma)(z_{k+1}, \rho_{k+1}) = \tau_{\text{LP}}^{-1}(z_{k+1}, \rho_{k+1})$ . This shows that the backward transitions used by UHA can be expressed as a sequence of two steps: the inverse of a leapfrog step used by HMC, followed by exact resampling of the momentum. This is exactly the same as the backward transition from algorithm 2 with exact momentum resampling (possible due to removing the score network).

This shows that the forward and backward transitions used by UHA are recovered by our framework with the simulation scheme from section 4, using the choices stated in theorem 4.3. Finally, using the momentum augmented distributions given by

$$\bar{p}(z_K, \rho_K) = \bar{p}(z_K)\mathcal{N}(\rho_K | 0, I) \quad \text{and} \quad q(z_1, \rho_1) = q(z_1)\mathcal{N}(\rho_1 | 0, I),$$

and the result from lemma 1, our framework yields the ratio

$$\begin{aligned} \frac{\bar{p}(z_{1:K}, \rho_{1:K})}{q(z_{1:K}, \rho_{1:K})} &= \frac{\bar{p}(z_K)\mathcal{N}(\rho_K | 0, I)}{q(z_1)\mathcal{N}(\rho_1 | 0, I)} \prod_{k=1}^{K-1} \frac{\mathcal{N}(\rho_k | 0, I)}{\mathcal{N}(\rho'_k | 0, I)} \\ &= \frac{\bar{p}(z_K)}{q(z_1)} \prod_{k=1}^{K-1} \frac{\mathcal{N}(\rho_{k+1} | 0, I)}{\mathcal{N}(\rho'_k | 0, I)}, \end{aligned}$$

which is exactly the ratio used by UHA. This shows that our framework with the choices from theorem 4.3 recovers UHA.



Table 5: ELBO achieved after training by different methods for different values of  $K$  for a logistic regression model with the *ionosphere* ( $d = 35$ ) dataset. Higher is better. Plain VI achieves an ELBO of  $-124.1$  nats. Best result for each value of  $K$  highlighted.

	Logistic regression ( <i>Ionosphere</i> )					
	ULA	MCD	UHA	LDVI	UHA <sub>EM</sub>	LDVI <sub>EM</sub>
$K = 8$	-116.4	-114.6	-115.6	<b>-114.4</b>	-117.7	-115.5
$K = 16$	-115.4	-113.6	-114.4	<b>-113.1</b>	-115.9	-113.8
$K = 32$	-114.5	-112.9	-113.4	<b>-112.4</b>	-114.6	-112.9
$K = 64$	-113.8	-112.5	-112.8	<b>-112.1</b>	-113.6	-112.4
$K = 128$	-113.1	-112.2	-112.3	<b>-111.9</b>	-113.1	-112.1
$K = 256$	-112.7	-112.1	-112.1	<b>-111.7</b>	-112.5	-111.9

## E SIMPLER DISCRETIZATION SCHEME FROM SECTION 5.2

This section shows the derivation of the forward and backward transitions from algorithms 4 and 5 together and an expression for their ratio, and results on all datasets using the resulting method.

### E.1 Transitions

**Forward transitions** The forward transitions  $F_k^{\text{em}}(z_{k+1}, \rho_{k+1} | z_k, \rho_k)$  from algorithm 4 are obtained by splitting the forward SDE as

$$\begin{bmatrix} dz^t \\ d\rho^t \end{bmatrix} = \underbrace{\begin{bmatrix} \rho^t dt \\ 0 \end{bmatrix}}_{\mathcal{U}_F} + \underbrace{\begin{bmatrix} 0 \\ \nabla \log \pi^t(z^t) dt - \gamma \rho^t dt + \sqrt{2\gamma} dw^t \end{bmatrix}}_{\mathcal{V}_F},$$

and by sequentially composing the simulations for components  $\mathcal{V}_F \mathcal{U}_F$ . Component  $\mathcal{V}_F$  is simulated using the Euler-Maruyama scheme, while  $\mathcal{U}_F$  is simulated exactly. This yields the forward transitions from algorithm 4.

**Backward transitions** The backward transitions  $B_k^{\text{em}}(z_k, \rho_k | z_{k+1}, \rho_{k+1})$  from algorithm 5 are obtained by splitting the time-reversed SDE as

$$\begin{bmatrix} dy^t \\ d\lambda^t \end{bmatrix} = \underbrace{\begin{bmatrix} -\lambda^t dt \\ 0 \end{bmatrix}}_{\mathcal{U}_B} + \underbrace{\begin{bmatrix} 0 \\ -\nabla \log \pi^{T-t}(y^t) dt - \gamma \lambda^t dt + 2\gamma s(T-t, y^t, \lambda^t) dt + \sqrt{2\gamma} dw^t \end{bmatrix}}_{\mathcal{V}_B}$$

and by sequentially composing the simulations for components  $\mathcal{U}_B \mathcal{V}_B$ . Component  $\mathcal{U}_B$  is simulated exactly, while  $\mathcal{V}_B$  is simulated using the Euler-Maruyama scheme. This yields the backward transitions from algorithm 5.

**Ratio between transitions** The ratio between the transitions from algorithms 4 and 5 is given by

$$\frac{B_k^{\text{em}}(z_k, \rho_k | z_{k+1}, \rho_{k+1})}{F_k^{\text{em}}(z_{k+1}, \rho_{k+1} | z_k, \rho_k)} = \frac{\mathcal{N}(\rho_k | \rho_{k+1} (1 - \delta\gamma) - \delta \nabla \log \pi_{k\delta}(z_k) + 2\delta\gamma s(k\delta, z_k, \rho_{k+1}), 2\delta\gamma I)}{\mathcal{N}(\rho_{k+1} | \rho_k (1 - \gamma\delta) + \delta \nabla \log \pi_{k\delta}(z_k), 2\gamma\delta I)}. \quad (33)$$

This can be obtained following a similar reasoning as the one used to prove lemma 1.

### E.2 Results on all models

Results for all models are shown in tables 5, 6, 7, 8 and 9. In addition to UHA<sub>EM</sub> and LDVI<sub>EM</sub>, the tables include results for ULA, MCD, UHA and LDVI as well, to facilitate comparisons. It can be observed that, for all models, using the simpler transitions from algorithms 4 and 5 (i.e. UHA<sub>EM</sub> and LDVI<sub>EM</sub>) lead to worse results than those obtained using the transitions from algorithms 1 and 2 (i.e. UHA and LDVI).

Table 6: ELBO achieved after training by different methods for different values of  $K$  for a logistic regression model with the *sonar* ( $d = 61$ ) dataset. Higher is better. Plain VI achieves an ELBO of  $-138.6$  nats. Best result for each value of  $K$  highlighted.

Logistic regression ( <i>Sonar</i> )						
	ULA	MCD	UHA	LDVI	UHA <sub>EM</sub>	LDVI <sub>EM</sub>
$K = 8$	-122.4	-117.2	-120.1	<b>-116.3</b>	-124.1	-118.5
$K = 16$	-119.9	-114.4	-116.8	<b>-112.6</b>	-119.9	-114.4
$K = 32$	-117.4	-112.4	-113.9	<b>-110.6</b>	-116.4	-111.7
$K = 64$	-115.3	-111.1	-111.9	<b>-109.7</b>	-113.8	-110.3
$K = 128$	-113.5	-110.2	-110.6	<b>-109.1</b>	-111.9	-109.6
$K = 256$	-112.1	-109.7	-109.7	<b>-108.9</b>	-110.7	-109.1

Table 7: ELBO achieved after training by different methods for different values of  $K$  for the Brownian motion model ( $d = 32$ ). Higher is better. Plain VI achieves an ELBO of  $-4.4$  nats. Best result for each value of  $K$  highlighted.

Brownian motion						
	ULA	MCD	UHA	LDVI	UHA <sub>EM</sub>	LDVI <sub>EM</sub>
$K = 8$	-1.9	-1.4	-1.6	<b>-1.1</b>	-2.8	-2.8
$K = 16$	-1.5	-0.8	-1.1	<b>-0.5</b>	-2.2	-1.4
$K = 32$	-1.1	-0.4	-0.5	<b>0.1</b>	-1.6	-0.5
$K = 64$	-0.7	-0.1	0.1	<b>0.5</b>	-0.9	0.1
$K = 128$	-0.3	0.2	0.4	<b>0.7</b>	-0.4	0.4
$K = 256$	-0.1	0.5	0.6	<b>0.9</b>	0.1	0.6

Table 8: ELBO achieved after training by different methods for different values of  $K$  for the Lorenz system model ( $d = 90$ ). Higher is better. Plain VI achieves an ELBO of  $-1187.8$  nats. Best result for each value of  $K$  highlighted.

Lorenz system						
	ULA	MCD	UHA	LDVI	UHA <sub>EM</sub>	LDVI <sub>EM</sub>
$K = 8$	-1168.2	-1168.1	-1166.3	<b>-1166.1</b>	-1170.5	-1170.5
$K = 16$	-1165.7	-1165.6	-1163.1	<b>-1162.2</b>	-1169.8	-1166.8
$K = 32$	-1163.2	-1163.3	-1160.3	<b>-1157.6</b>	-1167.9	-1162.9
$K = 64$	-1160.9	-1161.1	-1157.7	<b>-1153.7</b>	-1161.3	-1161.4
$K = 128$	-1158.9	-1158.9	-1155.4	<b>-1153.1</b>	-1158.1	-1163.4
$K = 256$	-1157.2	-1157.1	-1153.3	<b>-1151.1</b>	-1163.1	-1154.6

Table 9: ELBO achieved after training by different methods for different values of  $K$  for a random effect regression model with the *seeds* dataset ( $d = 26$ ). Higher is better. Plain VI achieves an ELBO of  $-77.1$  nats. Best result for each value of  $K$  highlighted.

Random effect regression (seeds)						
	ULA	MCD	UHA	LDVI	UHA <sub>EM</sub>	LDVI <sub>EM</sub>
$K = 8$	-75.5	-75.1	<b>-74.9</b>	<b>-74.9</b>	-75.9	-75.5
$K = 16$	-75.2	-74.6	-74.6	<b>-74.5</b>	-75.1	-75.1
$K = 32$	-74.9	-74.3	<b>-74.2</b>	<b>-74.2</b>	-74.8	-74.8
$K = 64$	-74.6	-74.1	-74.1	<b>-73.9</b>	-74.4	-74.4
$K = 128$	-74.3	-73.9	-73.8	<b>-73.7</b>	-74.1	-74.1
$K = 256$	-74.1	-73.7	-73.7	<b>-73.6</b>	-73.9	-73.7