DENOISING DIFFUSION SAMPLERS

Anonymous authors
Paper under double-blind review

ABSTRACT

Denoising diffusion models are a very popular class of generative models which provide state-of-the-art results in a variety of domains such as image and speech synthesis. One adds gradually noise to data using a diffusion to transform the data distribution into a Gaussian distribution. Samples from the generative model are then obtained by simulating an approximation of the time-reversal of this diffusion initialized by Gaussian samples. Practically, the intractable score terms appearing in the time-reversed process are approximated using score matching techniques. We explore here a similar idea to sample approximately from unnormalized probability density functions and estimate their normalizing constants. We consider a process where the target density diffuses towards a Gaussian. Denoising Diffusion Samplers (DDS) are obtained by approximating the corresponding time-reversal. Score matching is not applicable in this context and an alternative variational inference approach is used. However, we can leverage some of the ideas introduced in generative modeling to this Monte Carlo sampling task. Similarly, we can adapt existing theoretical results from denoising diffusion models to provide theoretical guarantees for DDS. We discuss the connections between DDS, optimal control and Schrödinger bridge and we finally demonstrate them experimentally on a variety of sampling tasks.

1 INTRODUCTION

Let π be a probability density on $\mathbb{R}^d$ of the form

$$\pi(x) = \frac{\gamma(x)}{Z}, \qquad Z = \int_{\mathbb{R}^d} \gamma(x)\mathrm{d}x, \tag{1}$$

where $\gamma : \mathbb{R}^d \to \mathbb{R}^+$ can be evaluated pointwise however the normalizing constant Z is intractable. We are here interested in both estimating Z and obtaining approximate samples from π.

A large variety of Monte Carlo techniques has been developed in the literature to address this problem. In particular Annealed Importance Sampling (AIS) (Neal, 2001) and its Sequential Monte Carlo (SMC) extensions (Del Moral et al., 2006) are often regarded as the "gold standard" approach to compute normalizing constants. These methods rely on the introduction of a sequence of annealed versions of the target distribution which are explored using a combination of Markov chain Monte Carlo (MCMC) kernels, importance sampling and resampling steps.

In machine learning, variational inference techniques have also been proposed. One considers a flexible family of easy-to-sample distributions q^θ whose parameters are optimized by minimizing a suitable metric w.r.t. π, typically the reverse Kullback–Leibler discrepancy $\mathrm{KL}(q^\theta||\pi)$. Typical choices for q_θ include mean-field approximation (Jordan et al., 1998) or more sophisticated techniques such as normalizing flows (Papamakarios et al., 2019). However, these methods can fail to capture complex targets; e.g. it was shown that normalising flows become pathological when approximating targets with supports having complicated topologies (Cornish et al., 2020).

To bypass these problems and improve the flexibility of the variational family $q^\theta(x)$, it is useful to model it as the marginal distribution of an auxiliary extended target distribution; i.e. $q^\theta(x) = \int q^\theta(x,u)\mathrm{d}u$. As this marginal is typically not tractable, θ is then learned by minimizing a discrepancy measure between $q^\theta(x,u)$ and an extended target distribution $p^\theta(x,u) = \pi(x)p^\theta(u|x)$ for $\int p^\theta(u|x)\mathrm{d}u = 1$ for all x (Agakov & Barber, 2004). AIS can be thought actually as a specific example of such a procedure where $q^\theta(x,u)$ is the joint distribution of the Markov chain defined by a

sequence of MCMC kernels whose final state is x while $p^\theta(x,u)$ is the corresponding AIS extended target distribution (Neal, 2001). To be able to minimize easily $\mathrm{KL}(q^\theta||p^\theta)$ using stochastic gradient and the reparameterization trick, various differentiable version of AIS using differentiable Langevin and Hamiltonian MCMC proposals have been proposed (Wu et al., 2020; Geffner & Domke, 2021; Thin et al., 2021; Zhang et al., 2021; Doucet et al., 2022; Geffner & Domke, 2022).

In this work, we follow a different approach inspired by Denoising Diffusion Probabilistic Models (DDPM), a powerful class of generative models (Sohl-Dickstein et al., 2015; Ho et al., 2020; Song et al., 2021b). In this context, one adds noise progressively to data using a diffusion to transform the complex data distribution into a Gaussian distribution. The time-reversal of this diffusion can then be used to transform a Gaussian sample into a sample from the target. While superficially similar to Langevin dynamics, this process can mix fast even in high dimension as it inherits the mixing properties of the forward diffusion (De Bortoli et al., 2021, Theorem 1), typically an Orstein–Ulhenbeck process with dimension-free convergence rates. Practically, the time-reversal of this diffusion needs to be approximated as its drift involves the logarithmic derivatives of the marginal distributions of the forward diffusion known as the scores.

These results motivate us to investigate a novel class of samplers termed Denoising Diffusion Samplers (DDS). As for generative modeling, we consider a forward diffusion which transforms progressively the initial target distribution into a Gaussian distribution, this defines an extended target distribution $p(x,u)$. DDS are obtained by approximating the time-reversal of such processes $q^\theta(x,u)$. What distinguishes DDS from DDPM is that we cannot simulate sample paths from the diffusion we want to time-reverse in the sampling context as the initial state is distributed from the intractable target. Hence score matching ideas to approximate the score terms are not applicable. This requires the development of alternative techniques and we focused here again on $\mathrm{KL}(q^\theta||p^\theta)$. However, we show how one can exploit the similarities between DDS and DDPM to leverage and adapt some of the ideas developed in generative modeling to Monte Carlo sampling.

We also detail the connections between DDS and KL and path integral control (Kappen et al., 2012; Heng et al., 2020) and Schrödinger bridges (Beghi, 1996; Bernton et al., 2019; De Bortoli et al., 2021; Vargas et al., 2021; Zhang & Chen, 2022b; Zhang et al., 2021). Finally we compare experimentally DDS to state-of-the-art Monte Carlo methods.

2 Denoising Diffusion Samplers: Continuous Time

We start here by formulating DDS in continuous-time. This continuous-time formulation is not useful practically per se as one cannot simulate exactly the resulting stochastic and ordinary differential equations. However this provides insight on the structure of time-reversal we want to approximate. Numerical integrators are discussed in the following section.

2.1 Forward and backward diffusions

Consider the forward noising diffusion given by an Ornstein–Uhlenbeck process[1]

$$\mathrm{d}x_t = -\beta_t x_t \mathrm{d}t + \sigma\sqrt{2\beta_t}\mathrm{d}B_t, \qquad x_0 \sim \pi, \tag{2}$$

where $(B_t)_{t\in[0,T]}$ is a d-dimension Brownian motion and $t \to \beta_t$ is a non-decreasing function. This diffusion induces the path measure $\mathcal{P}$ on the time interval $[0,T]$. Its transition density satisfies

$$p_{t|0}(x_t|x_0) = \mathcal{N}\left(x_t; x_0 \exp\left(-\int_0^t \beta_s \mathrm{d}s\right), \sigma^2\left(1-\exp\left(-2\int_0^t \beta_s \mathrm{d}s\right)\right) I\right). \tag{3}$$

From now on, we will always consider a scenario where $\int_0^T \beta_s \mathrm{d}s \gg 1$. If we denote by p_t the marginal density of x_t under the forward diffusion (2), this ensures that $p_T(x_T) \approx \mathcal{N}(x_T; 0, \sigma^2 I)$. We can thus think of this forward diffusion as transporting (approximately) the target density π to this Gaussian density.

Following (Haussmann & Pardoux, 1986), the time-reversal of this diffusion $(y_t)_{t\in[0,T]} = (x_{T-t})_{t\in[0,T]}$ is given by

$$\mathrm{d}y_t = \beta_{T-t}\Big\{y_t + 2\sigma^2 \nabla \ln p_{T-t}(y_t)\Big\}\mathrm{d}t + \sigma\sqrt{2\beta_{T-t}}\mathrm{d}W_t, \qquad y_0 \sim p_T, \tag{4}$$

[1]This is referred to as a Variance Preserving diffusion in Song et al. (2021b).

where $(W_t)_{t\in[0,T]}$ is another d-dimensional Brownian motion. By definition this time-reversal starts from $p_T(y_0) \approx \mathcal{N}(y_0; 0, \sigma^2 I)$ and at time T provides $y_T \sim \pi$. This suggests that if we could simulate approximately the diffusion (4), then we would obtain approximate samples from π. This idea is at the core of DDPM.

Practically, this requires however being able to approximate the logarithmic derivatives $(\nabla \ln p_t(x))_{t\in[0,T]}$ known as the scores. To achieve this, DDPM rely on (denoising) score matching (Hyvärinen, 2005; Vincent, 2011) as it is easy to sample approximately from (2). This is not possible in our scenario as sampling from (2) requires sampling $x_0 \sim \pi$ which is impossible by assumption.

2.2 Reference diffusion and value function

In our context, it will useful to reformulate the path measure of interest $\mathcal{P}$ induced by (2) as the solution to a variational problem defined through a simple reference process. This reference process is defined by the diffusion following the stochastic differential equation (2) but initialized at $p_0^{\text{ref}}(x_0) = \mathcal{N}(x_0; 0, \sigma^2 I)$. We denote by $\mathcal{P}^{\text{ref}}$ the resulting path measure whose marginals satisfies $p_t^{\text{ref}}(x_t) = \mathcal{N}(x_t; 0, \sigma^2 I)$ for all $t \in [0, T]$. From the chain rule for KL, it is easy to check that $\mathcal{P}$ can be identified as the path measure minimizing the reverse (or forward) KL discrepancy w.r.t. $\mathcal{P}^{\text{ref}}$ over the set of path measures $\mathcal{Q}$ with marginal $q_0(x_0) = \pi(x_0)$ at time $t = 0$, that is

$$\mathcal{P} = \arg\max_{\mathcal{Q}} \Big\{ \mathrm{KL}(\mathcal{Q}||\mathcal{P}^{\text{ref}}) : q_0 = \pi \Big\}.$$

A time-reversal representation of $\mathcal{P}^{\text{ref}}$ is given by $(y_t)_{t\in[0,T]} = (x_{T-t})_{t\in[0,T]}$ satisfying

$$\mathrm{d}y_t = \underbrace{\beta_{T-t}\Big\{ y_t + 2\sigma^2 \nabla \ln p_{T-t}^{\text{ref}}(y_t) \Big\}}_{:=-\beta_{T-t} y_t} \mathrm{d}t + \sigma\sqrt{2\beta_{T-t}}\mathrm{d}W_t, \qquad y_0 \sim p_0^{\text{ref}}, \tag{5}$$

Hence we can also rewrite the time-reversal (4) of $\mathcal{P}$ as

$$\mathrm{d}y_t = -\beta_{T-t}\Big\{ y_t - 2\sigma^2 \nabla \ln \phi_{T-t}(y_t) \Big\}\mathrm{d}t + \sigma\sqrt{2\beta_{T-t}}\mathrm{d}W_t, \qquad y_0 \sim p_T, \tag{6}$$

where $\phi_t(x_t) = p_t(x_t)/p_t^{\text{ref}}(x_t)$ is a so-called value function which satisfies a Kolmogorov equation such that $\phi_t(x_t) = \mathbb{E}_{\mathcal{P}^{\text{ref}}}[\phi_0(x_0)|x_t]$, the expectation being w.r.t. the law of the reference process.

2.3 Learning the time-reversal

To approximate the time-reversal of $\mathcal{P}$, (4) suggests to consider a path measure $\mathcal{Q}_\theta$ whose time-reversal is induced by the diffusion

$$\mathrm{d}y_t = \beta_{T-t}\Big\{ y_t + 2\sigma^2 s_\theta(T-t, y_t) \Big\}\mathrm{d}t + \sigma\sqrt{2\beta_{T-t}}\mathrm{d}W_t, \qquad y_0 \sim \mathcal{N}(0, \sigma^2 I), \tag{7}$$

so that $s_\theta(t, x) \approx \nabla \ln p_t(x)$. In practice, we would parameterize $s_\theta(t, x)$ by a neural network and minimize

$$\mathrm{KL}(\mathcal{Q}_\theta||\mathcal{P}) = \mathrm{KL}(p_T||\mathcal{N}(0, \sigma^2 I)) + \mathbb{E}_{y_0\sim\mathcal{N}(0,\sigma^2 I)}[\mathrm{KL}(\mathcal{Q}_\theta(\cdot|y_0)||\mathcal{P}(\cdot|y_0))], \tag{8}$$

where by Girsanov's formula (Klebaner, 2012) one obtains

$$\mathbb{E}_{y_0\sim\mathcal{N}(0,\sigma^2 I)}[\mathrm{KL}(\mathcal{Q}_\theta(\cdot|y_0)||\mathcal{P}(\cdot|y_0))] = \sigma^2 \mathbb{E}_{\mathcal{Q}_\theta}\Big[\int_0^T \beta_{T-t}||s_\theta(T-t, y_t) - \nabla \ln p_{T-t}(y_t)||^2 \mathrm{d}t \Big]. \tag{9}$$

This expression of the KL is reminiscent of the expression obtained by (Song et al., 2021a, Theorem 1) in the context of DDPM. However, there is a key difference as the expectation appearing in (9) is here w.r.t. $\mathcal{Q}_\theta$ and not w.r.t. $\mathcal{P}$. Hence, we cannot use integration by parts to get rid of the intractable score functions $(\nabla \ln p_t(x)))_{t\in[0,T]}$.

Instead, we take inspiration from (6) and consider a path measure $\mathcal{Q}_\theta$ whose time-reversal is induced by the diffusion

$$\mathrm{d}y_t = -\beta_{T-t}\Big\{ y_t - 2\sigma^2 f_\theta(T-t, y_t) \Big\}\mathrm{d}t + \sigma\sqrt{2\beta_{T-t}}\mathrm{d}W_t, \qquad y_0 \sim \mathcal{N}(0, \sigma^2 I). \tag{10}$$

so that $y_t \sim q_{T-t}^\theta$. Obviously this equivalent to consider (7) for $f_\theta(t, x) = s_\theta(t, x) - \nabla \ln p_t^{\text{ref}}(x) = s_\theta(t, x) + x/\sigma^2$ but will simplify notation.

Proposition 1. *From Girsanov theorem, we have*

$$\ln\left(\frac{\mathrm{d}\mathcal{Q}^\theta}{\mathrm{d}\mathcal{P}^{\text{ref}}}(y_{[0,T]})\right) = \sigma^2 \int_0^T \beta_{T-t}||f_\theta(T-t,y_t)||^2\mathrm{d}t + \sigma \int_0^T \sqrt{2\beta_{T-t}} f_\theta(T-t,y_t)^\top \mathrm{d}W_t. \quad (11)$$

It follows that

$$\begin{aligned}\mathrm{KL}(\mathcal{Q}_\theta||\mathcal{P})) &= \mathrm{KL}(\mathcal{Q}_\theta||\mathcal{P}^{\text{ref}}) + \mathbb{E}_{y_T\sim q_0^\theta}\Big[\ln\left(\frac{\mathcal{N}(y_T;0,\sigma^2 I)}{\pi(y_T)}\right)\Big] \\ &= \mathbb{E}_{\mathcal{Q}_\theta}\Big[\sigma^2\int_0^T \beta_{T-t}||f_\theta(T-t,y_t)||^2\mathrm{d}t + \ln\left(\frac{\mathcal{N}(y_T;0,\sigma^2 I)}{\pi(y_T)}\right)\Big]. \quad (12)\end{aligned}$$

For θ^* minimizing this KL, approximate samples from π can be obtained by sampling the approximate time reversal (10) and returning $y_T \sim q_0^\theta$. We can also obtain an unbiased estimate of Z by using the following importance sampling identity

$$\hat{Z} = \frac{\gamma(y_T)}{\mathcal{N}(y_T;0,\sigma^2 I)}\frac{\mathrm{d}\mathcal{P}^{\text{ref}}}{\mathrm{d}\mathcal{Q}^\theta}(y_{[0,T]}), \quad \text{for } y_{[0,T]} \sim \mathcal{Q}_\theta, \quad (13)$$

where the second term on the r.h.s. of (13) can be computed directly from (11). Obviously practically, the stochastic differential equation (10) as well the integrals in (11) need to be time-discretized. We will discuss a numerical integrator in the next section.

2.4 Ordinary differential equation formulation

One can define an ordinary differential equation (ODE) admitting the same marginal distributions as the noising diffusion (2) (see e.g. (Song et al., 2021b))

$$\mathrm{d}x_t = -\beta_t\Big\{x_t + \sigma^2\nabla\ln p_t(x_t)\Big\}\mathrm{d}t \quad (14)$$

Now from $f_{\theta^*}(t,x) \approx \nabla \ln \phi_t(x)$, we obtain $x_t + \sigma^2\nabla\ln p_t(x) \approx \sigma^2 f_{\theta^*}(t,x)$. So it is also possible to sample approximately from π by integrating the ODE over $[0,T]$

$$\mathrm{d}y_t = \sigma^2\beta_{T-t}f_{\theta^*}(T-t,y_t)\mathrm{d}t, \qquad y_0 \sim \mathcal{N}(0,\sigma^2 I). \quad (15)$$

We denote by $\bar{q}_{T-t}^\theta$ the distribution of y_t so that $y_T \sim \bar{q}_0^\theta$ is an approximate sample from π. We can then obtain an unbiased estimate of Z by using the following importance sampling identity

$$\hat{Z} = \frac{\gamma(y_T)}{\bar{q}_0^\theta(y_T)}, \quad \text{for } y_T \sim \bar{q}_0^\theta. \quad (16)$$

It is indeed possible to use compute this estimate as, contrary to the proposal q_0^θ induced by the diffusion (10), we can compute pointwise $\bar{q}_0^\theta$ using the instantaneous change of variables formula (Chen et al., 2018)

$$\ln \bar{q}_0^\theta(y_T) = \ln\mathcal{N}(y_0;0,\sigma^2 I) - \sigma^2\int_0^T \beta_{T-t}\nabla\cdot f_{\theta^*}(T-t,y_t)\mathrm{d}t \quad (17)$$

where $(y_t)_{t\in[0,T]}$ arises from the ODE (15). The expensive divergence term can be estimated using the Hutchinson estimator (Grathwohl et al., 2018; Song et al., 2021b).

3 Denoising Diffusion Samplers: Discrete-Time

In the previous section, we have discussed DDS in continuous-time. Obviously, it is impossible in practice to sample exactly the diffusion (10), estimate an unbiased gradient of the KL between path measures $\mathrm{KL}(\mathcal{Q}_\theta||\mathcal{P}))$ in (12) or integrate exactly the ODE (15). We propose principled discrete-time approximating addressing these challenges.

3.1 FORWARD AND BACKWARD PROCESSES

Consider a discretization step $\delta > 0$ such that $K = T/\delta$ is an integer. The time-discretized forward noising diffusion (2) corresponds to a discrete-time "forward" Markov process of initial density $x_0 \sim \pi(x_0)$ and Markov transition densities $x_k \sim p_{k|k-1}(x_k|x_{k-1})$ for $k = 1, ..., K$. We denote the resulting marginal densities by $x_k \sim p_k$. We would like to sample the time-reversal of this process, that is performing ancestral sampling, i.e. sample $x_K \sim p_K(x_k)$ then sample $x_{k-1} \sim p_{k-1|k}(x_{k-1}|x_k)$ for $k = K, ..., 1$ where $p_{k-1|k}(x_{k-1}|x_k) = p_{k|k-1}(x_k|x_{k-1})p_{k-1}(x_{k-1})/p_k(x_k)$, which corresponds to a discretization of (4).

3.2 REFERENCE PROCESS AND VALUE FUNCTIONS

As in continuous-time, it is useful to introduce a simple reference process of joint density $p^{\text{ref}}(x_{0:K})$ which is such that $x_0 \sim \mathcal{N}(x_0; 0, \sigma^2 I)$ and $x_k \sim p_{k|k-1}(x_k|x_{k-1})$. This time-discretization of $\mathcal{P}^{\text{ref}}$ will be designed to ensure that the marginals do satisfy $p^{\text{ref}}_k(x_k) = \mathcal{N}(x_k; 0, \sigma^2 I)$ as for the continuous-time process.

It will prove convenient to obtain $p^{\text{ref}}(x_{0:K})$ by time discretizing its time-reversal (5). As it is an Ornstein–Ulhenbeck process, we can integrate it exactly so that the marginals all follow $\mathcal{N}(x_0; 0, \sigma^2 I)$ and, for a regular time discretization $t_k = k\delta$ so that x_k in discrete-time corresponds to x_{t_k}. We obtain $p^{\text{ref}}_{k-1|k}(x_{k-1}|x_k) = \mathcal{N}(x_{k-1}; \sqrt{1-\alpha_k}x_k, \sigma^2\alpha_k I)$ for $\alpha_k = 1 - \exp\left(-2\int_{(k-1)\delta}^{k\delta} \beta_s \mathrm{d}s\right)$; that is the discrete-time time-reversal $(y_k)_{k=0}^K = (x_{K-k})_{k=0}^K$ is sampled using

$$y_{k+1} = \sqrt{1-\alpha_{K-k}}y_k + \sigma\sqrt{\alpha_{K-k}}\varepsilon_k, \qquad \varepsilon_k \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, I). \tag{18}$$

The intractable backward Markov densities $p_{k-1|k}$ one wants to approximate can also be rewritten as twisted versions of $p^{\text{ref}}_{k-1|k}$ by some value functions $(\phi_k)_{k=0}^K$. For completeness, this simple result is proven in appendix and can be thought as the discrete-time counterpart of (6).

Proposition 2. *We have for $k = 1, ..., K$*

$$p_{k-1|k}(x_{k-1}|x_k) = \frac{\phi_{k-1}(x_{k-1})p^{\text{ref}}_{k-1|k}(x_{k-1}|x_k)}{\phi_k(x_k)}, \quad \textit{for} \quad \phi_k(x_k) = \frac{p_k(x_k)}{p^{\text{ref}}_k(x_k)}. \tag{19}$$

The value functions $(\phi_k)_{k=1}^K$ satisfy $\phi_k(x_t) = \mathbb{E}_{p^{\text{ref}}}[\phi_0(x_0)|x_k]$.

This is consistent with earlier results (2) as $\phi_{k-1}(x) \approx \phi_k(x)$ for $\delta \ll 1$ so a Taylor expansion of $\nabla \ln \phi_k$ around x_k shows that $p_{k-1|k}(x_{k-1}|x_k) = p^{\text{ref}}_{k-1|k}(x_{k-1}|x_k)\exp(\ln \phi_{k-1}(x_{k-1}) - \ln \phi_k(x_k)) \approx \mathcal{N}(x_{k-1}; \sqrt{1-\alpha_k}x_k + \alpha_k\sigma^2\nabla \ln \phi_k(x_k), \alpha_k\sigma^2 I)$.

3.3 LEARNING THE BACKWARD KERNELS

In practice, we want to approximate the backward Markov transition densities $p_{k-1|k}$. By numerically integrating (5) using an exponential integrator time type scheme (see e.g. (De Bortoli, 2022; Zhang & Chen, 2022a)), we obtain

$$y_{k+1} = \sqrt{1-\alpha_{K-k}}y_k + 2\sigma^2(1-\sqrt{1-\alpha_{K-k}})\nabla \ln \phi_{T-t_k}(y_k) + \sigma\sqrt{\alpha_{K-k}}\varepsilon_k. \tag{20}$$

We use the same integrator for (10) initialized using $y_0 \sim \mathcal{N}(0, \sigma^2 I)$

$$y_{k+1} = \sqrt{1-\alpha_{K-k}}y_k + 2\sigma^2(1-\sqrt{1-\alpha_{K-k}})f_\theta(K-k, y_k) + \sigma\sqrt{\alpha_{K-k}}\varepsilon_k, \tag{21}$$

where $f_\theta(k, x) \approx \nabla \ln \phi_t(x)$. This defines the time reversal of the proposal $q^\theta(x_{0:K}) = \mathcal{N}(x_K; 0, \sigma^2 I)\prod_{k=1}^K q^\theta_{k-1|k}(x_{k-1}|x_k)$ with a slight abuse of notation we will write $q^\theta(y_{0:K}) = \mathcal{N}(y_0; 0, \sigma^2 I)\prod_{k=1}^K q^\theta_{k-1|k}(y_{K-k+1}|y_{K-k})$ for the time-reversal. Similarly $p^{\text{ref}}(y_{0:K}) = \mathcal{N}(y_0; 0, \sigma^2 I)\prod_{k=1}^K p^{\text{ref}}_{k-1|k}(y_{K-k+1}|y_{K-k})$.

While we could adapt the KL decomposition developed in (Sohl-Dickstein et al., 2015; Ho et al., 2020) to the reverse KL (see Appendix A.3), we found empirically that it is very unstable in our context. Instead, we will be relying on the discrete-time counterpart of (11).

Proposition 3. *The log density ratio between $q^\theta(y_{0:K})$ and $p^{\text{ref}}(y_{0:K})$ satisfies*

$$\ln\left(\frac{q_\theta(y_{0:K})}{p^{\text{ref}}(y_{0:K})}\right) = 2\sigma^2 \sum_{k=1}^{K} \frac{(1-\sqrt{1-\alpha_k})^2}{\alpha_k} ||f_\theta(k, y_{K-k})||^2 + 2\sigma \sum_{k=1}^{K} \frac{(1-\sqrt{1-\alpha_k})}{\sqrt{\alpha_k}} f_\theta(k, y_{K-k})^\top \epsilon_k \tag{22}$$

where $\epsilon_k := \frac{1}{\sigma\sqrt{\alpha_{K-k}}}(y_{k+1} - \sqrt{1-\alpha_{K-k}}y_k - 2\sigma^2(1-\sqrt{1-\alpha_{K-k}})\beta_k f_\theta(K-k, y_k)) \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, I)$ *for* $y_{0:K} \sim q_\theta(y_{0:K})$. *It follows that*

$$\text{KL}(q^\theta||p) = \text{KL}(q^\theta||p^{\text{ref}}) + \mathbb{E}_{q_0^\theta}\left[\ln\left(\frac{\pi(y_K)}{\mathcal{N}(y_K; 0, \sigma^2 I)}\right)\right] \tag{23}$$

$$= \mathbb{E}_{q^\theta}\left[2\sigma^2 \sum_{k=1}^{K} \frac{(1-\sqrt{1-\alpha_k})^2}{\alpha_k} ||f_\theta(k, y_{K-k})||^2 + \ln\left(\frac{\mathcal{N}(y_K; 0, \sigma^2 I)}{\pi(y_K)}\right)\right]. \tag{24}$$

We compute an unbiased gradient of this objective using the reparameterization trick and the JAX software package (Bradbury et al., 2018) for automatic differentiation. Unfortunately, contrary to DDPM, $q^\theta_{k|K}$ is not available in closed form for $k < K-1$ so we can neither mini-batches over the time index k without having to simulate the process until the minimum sampled time nor reparameterize x_k as in Ho et al. (2020).

Once we obtain the parameter θ^* minimizing $\text{KL}(q^\theta||p)$, DDS samples from q^{θ^*} using (21) for $\theta = \theta^*$. The final sample y_K has a distribution $q_0^{\theta^*}$ approximating π by design. By using importance sampling for the unnormalized target $Zp(y_{0:K})$ and the proposal $q_\theta(y_{0:K})$, we obtain the following unbiased estimate of the normalizing constant

$$\hat{Z} = \frac{\gamma(y_K)}{\mathcal{N}(y_K; 0, \sigma^2 I)} \frac{p^{\text{ref}}(y_{0:K})}{q^\theta(y_{0:K})}, \quad \text{for } y_{0:K} \sim q_\theta(y_{0:K}) \tag{25}$$

where the second term on the r.h.s. of (25) is given by the exponential of the opposite of (22).

4 Algorithm Settings

In our implementation we follow a parameterisation where f_θ is implemented using the PIS-GRAD network proposed in (Zhang & Chen, 2022b):

$$f_\theta(k, x) = \text{NN}_1(k, x; \theta) + \text{NN}_2(k; \theta) \circ \nabla \ln \pi(x), \tag{26}$$

where the final layers are initialised to 0 in order to make the path regularisation term null. Motivated by IDDPM (Nichol & Dhariwal, 2021) we parametrise the α_k sequence using their proposed schedule:

$$\alpha_k^{1/2} \propto \alpha_{\max}^{1/2} \cos^2\left(\frac{\pi}{2}\frac{1-k/K+s}{1+s}\right), \tag{27}$$

with $s = 0.008$ as in (Nichol & Dhariwal, 2021). Finally, it is known that optimization through unrolled computation graphs can be chaotic and introduce exploding/vanishing gradients (Parmas et al., 2018; Metz et al., 2020) which can numerically introduce bias. We experienced this in practice and found that detaching the target score term (i.e. $\nabla \ln \pi^\perp$) made optimization more stable for both approaches without affecting the final result. We adopt this across experiments, an ablation of this feature can be seen in Appendix B.4.1. This has been successfully applied in related prior work (Greff et al., 2019) where detached scores are provided as features to the recognition networks. Alternative approaches exist based on smoothing the chaotic loss-landscape (Vicol et al., 2021) but we leave this for future work.

5 Intepretation, Related Work and Extensions

5.1 DDS as Variational Inference on Extended Space

We can reinterpret DDS as a variational inference procedure approximating $\pi(x_0)$. A standard VI approach would consider a tractable variational approximation family $q_0^\theta(x_0)$ and minimizing

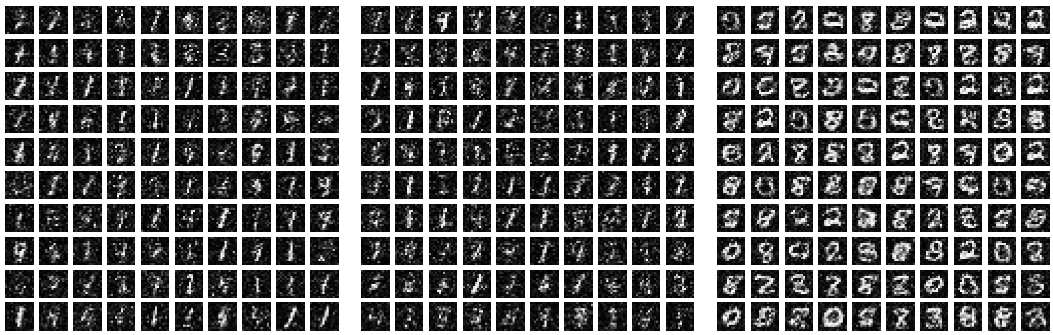

Figure 1: a) DDS , b) PIS , c) SMC. We can see how SMC mode collapses dramatically as it only seems to have 5 image classes. For PIS we identified 7 classes whilst for DDS we found 9 out of 10.

$\mathrm{KL}(q_0^\theta||\pi)$ w.r.t. θ. However, for complex targets, it is difficult to come up with a variational family that is expressive enough. So we consider here the case where $q_0^\theta(x_0)$ is built as the final time marginal of a reverse-time Markov chain, i.e. $q_0^\theta(x_0) = \int q^\theta(x_{0:K})\mathrm{d}x_{1:K}$. The price to pay for this flexibility is that $q_0^\theta(x_0)$ cannot be evaluated pointwise so we need to perform VI for the path distribution $q^\theta(x_{0:K})$ and hence also introduce a corresponding extended target distribution $p(x_{0:K}) = \pi(x_0)p(x_{1:K}|x_0)$ on this path space where $p(x_{1:K}|x_0) = \prod_{k=1}^K p_{k|k-1}(x_k|x_{k-1})$. This instrumental conditional distribution $p(x_{1:K}|x_0)$ acts as a regularizer as evidenced by the chain rule for KL

$$\mathrm{KL}(q^\theta||p) = \mathrm{KL}(q_0^\theta||\pi) + \mathbb{E}_{q_0^\theta}[\mathrm{KL}(q^\theta(\cdot|x_0)||p(\cdot|x_0))]. \tag{28}$$

Moreover, $\mathrm{KL}(q^\theta||p)$ also satisfies (23).

DDS can be thought of as the counterpart of the optimized AIS procedures that have been recently proposed in (Doucet et al., 2022; Geffner & Domke, 2022). In the AIS scenario, we fix the proposal on path space q, typically a inhomogeneous time-discretized overdamped/underdamped Langevin process ensuring that $q(x_0) \approx \pi(x_0)$. We then learn an extended target $p^\theta(x_{0:K}) = \pi(x_0)\prod_{k=1}^K p_{k|k-1}^\theta(x_k|x_{k-1})$ by minimizing $\mathrm{KL}(q||p^\theta)$. It was shown in Doucet et al. (2022) that this corresponds to a score matching loss for appropriate parameterization of $p_{k|k-1}^\theta(x_k|x_{k-1})$. This approach is thus much closer mathematically to DDPM than DDS.

5.2 DDS AS KL AND PATH INTEGRAL CONTROL

The reverse KL we minimize can be expressed as (see proof of Propostion 3)

$$\mathrm{KL}(q^\theta||p) = \int q^\theta(x_{0:K})\left\{\ln\left(\frac{\mathcal{N}(x_0;0,\sigma^2 I)}{\pi(x_0)}\right) + \sum_{k=1}^K \ln\left(\frac{q_{k-1|k}^\theta(x_{k-1}|x_k)}{p_{k-1|k}^{\mathrm{ref}}(x_{k-1}|x_k)}\right)\right\}\mathrm{d}x_{0:K}. \tag{29}$$

This problem is a KL control problem (Kappen et al., 2012). In continuous-time, (12) corresponds to a path integral control problem; see e.g. (Kappen & Ruiz, 2016).

In Heng et al. (2020), it was proposed to use KL control so as to sample from a target π and estimate its normalizing constant Z. However, the proposed algorithm therein is significantly different from the one proposed here. First it relies on p^{ref} being defined by a discretized non-homogeneous dynamics such that $p_K(x_K)$ is typically not approximating a known distribution. Second it approximates the value functions ϕ_k using regression using simple functions so as to be able to sample the twisted kernels (19). Third it relies on SMC to provide an initial estimate of p. This limits the applicability of this methodology to fairly simple models and provides Monte Carlo estimates which are not differentiable.

5.3 CONNECTIONS TO SCHRÖDINGER BRIDGES

The Schrödinger Bridge problem is a standard problem in probability, applied mathematics and optimal control. It takes the following form in discrete time. Given a reference density $p^{\mathrm{ref}}(x_{0:K})$,

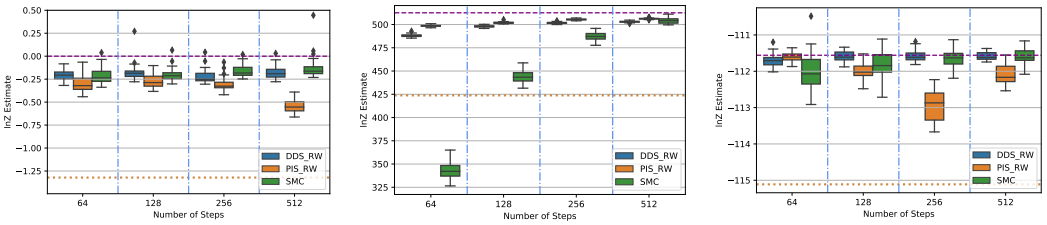

Figure 2: a) Funnel distribution, b) LGCP distribution, c) Ionosphere dataset.

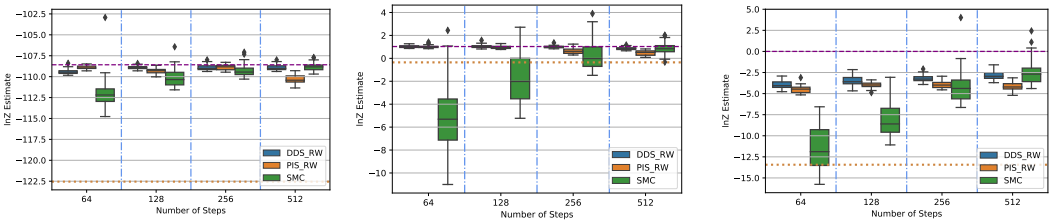

Figure 3: a) Sonar dataset, b) Brownian motion target, c) NICE target.

we aim to find the density $p(x_{0:K})$ such that

$$p = \arg\min_q \Big\{ \mathrm{KL}(q||p^{\mathrm{ref}}) : q_0 = \mu_0, \;\; q_K = \mu_K \Big\}. \tag{30}$$

where μ_0, μ_K are prescribed probability densities. This problem can be solved using iterative proportional fitting (IPF) (Fortet, 1940; Kullback, 1968) which is defined by the following recursion with initialization $p^0 = p^{\mathrm{ref}}$

$$p^{2n+1} = \arg\min_q \Big\{ \mathrm{KL}(q||p^{2n}) : \; q_K = \mu_K \Big\}, \tag{31}$$

$$p^{2n+2} = \arg\min_q \Big\{ \mathrm{KL}(q||p^{2n+1}) : q_0 = \mu_0 \Big\}. \tag{32}$$

For $\mu_0(x_0) = \pi(x_0)$ and $\mu_K(x_K) = \mathcal{N}(x_K; 0, \sigma^2 I)$, IPF would converge in one iteration if we had $p_K^{\mathrm{ref}}(x_K) = \mathcal{N}(x_K; 0, \sigma^2 I)$. As $p_K^{\mathrm{ref}}(x_K) \approx \mathcal{N}(x_K; 0, \sigma^2 I)$, we can thus think of DDS as approximating this SB.

For $q_0 = \pi$ and $q_K = \delta_0$, IPF also converges in one iteration if we have $p_K^{\mathrm{ref}}(x_K) = \delta_0(x_K)$. Designing such a reference process is simple, set e.g. a pinned Brownian motion $p_K^{\mathrm{ref}}(x_K) = \delta_0(x_K) \prod_{k=0}^{K-1} \mathcal{N}(x_k; x_{k+1}, \delta\sigma^2 I)$. This was discussed in discrete-time in (Beghi, 1996) and in continuous-time in (Follmer, 1984; Dai Pra, 1991; Tzen & Raginsky, 2019). Algorithmically, it was proposed by Zhang et al. (2021) to approximate the solution by some projection on some eigenfunctions. In parallel, Zhang & Chen (2022b) and Vargas (2021) have proposed to approximate this SB by using a neural network parameterization of the gradient of the logarithm of value function which is trained by minimizing a reverse KL. DDS can thus be seen as an alternative to this approach which relies on a reference dynamics corresponding to an Ornstein–Uhlenbeck process instead of relying on a reference dynamics corresponding to a pinned Brownian motion. Alternative reference dynamics based on stopping time could also be defined and have proved useful in the generative modeling context (Ye et al., 2022).

6 Experiments

In this section we compare DDS across several sampling baselines and standard benchmarking tasks. Across all tasks we compare our approach to PIS (Zhang & Chen, 2022b), SMC (Del Moral et al., 2006; Zhou et al., 2016) and Mean Field-VI (MF-VI) with a Gaussian variational distribution. In addition to the aforementioned benchmark we compare to AIS (Neal, 2001) based approaches and their

recent optimized variants using score matching (Doucet et al., 2022; Geffner & Domke, 2022) across three standard Bayesian inference models. Finally we explore a recently proposed task (Doucet et al., 2022) that uses a pre-trained normalising flow as a target, within this setting we propose a novel benchmarking criteria that allows us to asses mode collapse in high dimensions and explore the benefits of incorporating inductive biases into f_θ. For each of the approaches we consider we carefully tune their hyper-parameters (e.g. step size, diffusion coefficient, and such), details can be found in Appendix B.1.

6.1 Benchmarking Targets

We first discuss two standard target distributions which are often us to benchmark methods; see e.g. (Neal, 2003; Arbel et al., 2021; Heng et al., 2020; Zhang & Chen, 2022b).

Funnel Distribution: We consider the ten-dimensional 'funnel' distribution (Neal, 2003):

$$x_0 \sim \mathcal{N}(0, \sigma_f^2), \quad x_{1:9}|x_0 \sim \mathcal{N}(0, \exp(x_0)I), \tag{33}$$

where $\sigma_f^2 = 9$. This is a well-known challenging example. We would like to highlight that we found a bug in (Zhang & Chen, 2022b) where they accidentally set $\sigma_f = 1$ when evaluating PIS but not when comparing to SMC and other methods. This makes a significant difference in the results and thus we had to re-run our own faithful implementation of PIS across all experiments [2].

Log Gaussian Cox process: We consider a log Gaussian Cox process arising in spatial statistics to model the positions of pine saplings (Møller et al., 1998). We use a $d = M \times M = 1600$ grid, resulting in the target density

$$\gamma(x) = \mathcal{N}(x; \mu, K) \prod_{i \in [1:M]^2} \exp(x_i y_i - a \exp(x_i)). \tag{34}$$

6.2 Bayesian Models

In this section we explore three standard Bayesian inference models and compare them with standard AIS and MCD benchmarks in addition to the SMC and VI benchmarks presented so far.

Logistic Regression: We evaluate the performance on a Bayesian logistic regression model:

$$w \sim \mathcal{N}(0, \sigma_w^2 I), \quad y_i \sim \text{Bernoulli}(\text{sigmoid}(w^\top x_i)).$$

on two datasets, Ionosphere (Sigillito et al., 1989) ($d = 32$) and Sonar (Gorman & Sejnowski, 1988) ($d = 61$).

Brownian Motion: Here we consider a discretised Brownian motion with a Gaussian observation model as a time series model. Furthermore the diffusion coefficient of the Brownian motion is treated as latent making the total dimension of this time series model 32. This model was obtained from the Sountsov et al. (2020) software package and is specified in more detail in Appendix B.2.

6.3 Mode Collapse In High Dimensions

Normalizing Flow Evaluation: Following Doucet et al. (2022) we train NICE (Dinh et al., 2014) on a down-sampled 14×14 variant of MNIST (LeCun & Cortes, 2010). We use the trained model as our target distribution. As we can generate samples from our target we chose to evaluate the methods samples by measuring the Wasserstein distance between target samples and the respective methods samples. Finally we display the samples in Figure 1. This constitutes of a new evaluation criteria for this benchmark which allows to asses mode collapse for samplers in high dimensional settings.

[2] We reached out to the authors of Zhang & Chen (2022b) who acknowledged the error and instructed their benchmarking results should be re-evaluated.

Method	$\ln Z$ Estimate	$\mathcal{W}_\gamma^2(\pi_{\text{true}}, \hat{\pi})$
DDS	-3.204 ± 0.645	658.079
PIS	-3.933 ± 0.754	658.778
SMC	-4.255 ± 2.043	750.245
MCD	$-6.25 \pm$ -	-

Table 1: Results on NICE target. For MCD we used the results from Doucet et al. (2022)

7 Discussion

In this paper, we have explored the use of denoising diffusion models ideas to sample unnormalized probability distributions and estimate their normalizing constants. We have compared our results to state-of-the-art Monte Carlo methods, including SMC samplers (Del Moral et al., 2006; Dai et al., 2020), optimized AIS samplers (Doucet et al., 2022) and PIS (Zhang & Chen, 2022b; Vargas, 2021). We found overall that properly tuned DDS are competitive with SMC and typically numerically more stable than PIS. However this comes at the cost of a non-negligible training time compared to SMC.

The benefits of DDS over SMC are that they require much lower memory to run as they do not rely on a large collection of interacting particles and allow the parallel simulation of multiple samples once trained. We believe that it is a promising direction that is worth further exploring. There has been much work devoted to improving successfully DDPM over the past two years and some of these techniques should also benefit the development of more powerful DDS samplers.

References

Felix V Agakov and David Barber. An auxiliary variational method. In *Advances in Neural Information Processing Systems*, 2004.

Michael Arbel, Alex Matthews, and Arnaud Doucet. Annealed flow transport Monte Carlo. In *International Conference on Machine Learning*, 2021.

Alessandro Beghi. On the relative entropy of discrete-time Markov processes with given end-point densities. *IEEE Transactions on Information Theory*, 42(5):1529–1535, 1996.

Espen Bernton, Jeremy Heng, Arnaud Doucet, and Pierre E Jacob. Schrödinger bridge samplers. *arXiv preprint arXiv:1912.13170*, 2019.

James Bradbury, Roy Frostig, Peter Hawkins, Matthew James Johnson, Chris Leary, Dougal Maclaurin, George Necula, Adam Paszke, Jake VanderPlas, Skye Wanderman-Milne, and Qiao Zhang. JAX: composable transformations of Python+NumPy programs, 2018. URL `http://github.com/google/jax`.

Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. *Advances in neural information processing systems*, 31, 2018.

Rob Cornish, Anthony Caterini, George Deligiannidis, and Arnaud Doucet. Relaxing bijectivity constraints with continuously indexed normalising flows. In *International Conference on Machine Learning*, 2020.

Chenguang Dai, Jeremy Heng, Pierre E Jacob, and Nick Whiteley. An invitation to sequential Monte Carlo samplers. *arXiv preprint arXiv:2007.11936*, 2020.

Paolo Dai Pra. A stochastic control approach to reciprocal diffusion processes. *Applied Mathematics and Optimization*, 23(1):313–329, 1991.

Valentin De Bortoli. Convergence of denoising diffusion models under the manifold hypothesis. *arXiv preprint arXiv:2208.05314*, 2022.

Valentin De Bortoli, James Thornton, Jeremy Heng, and Arnaud Doucet. Diffusion Schrödinger bridge with applications to score-based generative modeling. In *Advances in Neural Information Processing Systems*, 2021.

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

Laurent Dinh, David Krueger, and Yoshua Bengio. Nice: Non-linear independent components estimation. *arXiv preprint arXiv:1410.8516*, 2014.

Arnaud Doucet, Will Sussman Grathwohl, Alexander G de G Matthews, and Heiko Strathmann. Score-based diffusion meets annealed importance sampling. In *Advances in Neural Information Processing Systems*, 2022.

H. Follmer. An entropy approach to the time reversal of diffusion processes. *Lecture Notes in Control and Information Sciences*, 69:156–163, 1984.

Robert Fortet. Résolution d'un systeme d'équations de M. Schrödinger. *J. Math. Pure Appl. IX*, 1: 83–105, 1940.

Tomas Geffner and Justin Domke. MCMC variational inference via uncorrected Hamiltonian annealing. In *Advances in Neural Information Processing Systems*, 2021.

Tomas Geffner and Justin Domke. Langevin diffusion variational inference. *arXiv preprint arXiv:2208.07743*, 2022.

R Paul Gorman and Terrence J Sejnowski. Analysis of hidden units in a layered network trained to classify sonar targets. *Neural networks*, 1(1):75–89, 1988.

Will Grathwohl, Ricky TQ Chen, Jesse Bettencourt, Ilya Sutskever, and David Duvenaud. Ffjord: Free-form continuous dynamics for scalable reversible generative models. In *International Conference on Learning Representations*, 2018.

Klaus Greff, Raphaël Lopez Kaufman, Rishabh Kabra, Nick Watters, Christopher Burgess, Daniel Zoran, Loic Matthey, Matthew Botvinick, and Alexander Lerchner. Multi-object representation learning with iterative variational inference. In *International Conference on Machine Learning*, pp. 2424–2433. PMLR, 2019.

Ulrich G Haussmann and Etienne Pardoux. Time reversal of diffusions. *The Annals of Probability*, 14(3):1188–1205, 1986.

Jeremy Heng, Adrian N Bishop, George Deligiannidis, and Arnaud Doucet. Controlled sequential Monte Carlo. *The Annals of Statistics*, 48(5):2904–2929, 2020.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems*, 2020.

Aapo Hyvärinen. Estimation of non-normalized statistical models by score matching. *The Journal of Machine Learning Research*, 6:695–709, 2005.

Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the fokker–planck equation. *SIAM journal on mathematical analysis*, 29(1):1–17, 1998.

Hilbert J Kappen, Vicenç Gómez, and Manfred Opper. Optimal control as a graphical model inference problem. *Machine learning*, 87(2):159–182, 2012.

Hilbert Johan Kappen and Hans Christian Ruiz. Adaptive importance sampling for control and inference. *Journal of Statistical Physics*, 162(5):1244–1266, 2016.

Diederik P Kingma and Max Welling. Auto-encoding variational Bayes. *arXiv preprint arXiv:1312.6114*, 2013.

Fima C Klebaner. *Introduction to Stochastic Calculus with Applications*. Imperial College Press, 2012.

Solomon Kullback. Probability densities with given marginals. *The Annals of Mathematical Statistics*, 39(4):1236–1243, 1968.

Yann LeCun and Corinna Cortes. MNIST handwritten digit database. http://yann.lecun.com/exdb/mnist/, 2010. URL http://yann.lecun.com/exdb/mnist/.

Luke Metz, Niru Maheswaranathan, Ruoxi Sun, C Daniel Freeman, Ben Poole, and Jascha Sohl-Dickstein. Using a thousand optimization tasks to learn hyperparameter search strategies. *arXiv preprint arXiv:2002.11887*, 2020.

Jesper Møller, Anne Randi Syversveen, and Rasmus Plenge Waagepetersen. Log Gaussian Cox processes. *Scandinavian Journal of Statistics*, 25(3):451–482, 1998. ISSN 03036898, 14679469.

Radford M Neal. Annealed importance sampling. *Statistics and Computing*, 11(2):125–139, 2001.

Radford M. Neal. Slice sampling. *The Annals of Statistics*, 31:705–767, 06 2003.

Alexander Quinn Nichol and Prafulla Dhariwal. Improved denoising diffusion probabilistic models. In *International Conference on Machine Learning*, pp. 8162–8171. PMLR, 2021.

George Papamakarios, Eric Nalisnick, Danilo Jimenez Rezende, Shakir Mohamed, and Balaji Lakshminarayanan. Normalizing flows for probabilistic modeling and inference. *arXiv preprint arXiv:1912.02762*, 2019.

Paavo Parmas, Carl Edward Rasmussen, Jan Peters, and Kenji Doya. Pipps: Flexible model-based policy search robust to the curse of chaos. In *International Conference on Machine Learning*, pp. 4065–4074. PMLR, 2018.

Vincent G Sigillito, Simon P Wing, Larrie V Hutton, and Kile B Baker. Classification of radar returns from the ionosphere using neural networks. *Johns Hopkins APL Technical Digest*, 10(3): 262–266, 1989.

Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *International Conference on Machine Learning*, 2015.

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

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

Pavel Sountsov, Alexey Radul, and contributors. Inference gym, 2020. URL https://pypi.org/project/inference_gym.

Achille Thin, Nikita Kotelevskii, Alain Durmus, Eric Moulines, Maxim Panov, and Arnaud Doucet. Monte Carlo variational auto-encoders. In *International Conference on Machine Learning*, 2021.

Belinda Tzen and Maxim Raginsky. Theoretical guarantees for sampling and inference in generative models with latent diffusions. In *Conference on Learning Theory*, pp. 3084–3114. PMLR, 2019.

Francisco Vargas. Machine-learning approaches for the empirical schrödinger bridge problem. Technical report, University of Cambridge, Computer Laboratory, 2021.

Francisco Vargas, Andrius Ovsianas, David Fernandes, Mark Girolami, Neil Lawrence, and Nikolas Nüsken. Bayesian learning via neural Schrödinger-Föllmer flows. *arXiv preprint arXiv:2111.10510*, 2021.

Paul Vicol, Luke Metz, and Jascha Sohl-Dickstein. Unbiased gradient estimation in unrolled computation graphs with persistent evolution strategies. In *International Conference on Machine Learning*, pp. 10553–10563. PMLR, 2021.

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

Hao Wu, Jonas Köhler, and Frank Noé. Stochastic normalizing flows. In *Advances in Neural Information Processing Systems*, 2020.

Mao Ye, Lemeng Wu, and Qiang Liu. First hitting diffusion models. *arXiv preprint arXiv:2209.01170*, 2022.

Benjamin Zhang, Tuhin Sahai, and Youssef Marzouk. Sampling via controlled stochastic dynamical systems. In *I (Still) Can't Believe It's Not Better! NeurIPS 2021 Workshop*, 2021.

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

Qinsheng Zhang and Yongxin Chen. Fast sampling of diffusion models with exponential integrator. *arXiv preprint arXiv:2204.13902*, 2022a.

Qinsheng Zhang and Yongxin Chen. Path integral sampler: a stochastic control approach for sampling. In *International Conference on Learning Representations*, 2022b.

Yan Zhou, Adam M Johansen, and John AD Aston. Toward automatic model comparison: an adaptive sequential monte carlo approach. *Journal of Computational and Graphical Statistics*, 25 (3):701–726, 2016.

A Appendix

A.1 Proof of Proposition 2

Proof. We have by Bayes' rule

$$\begin{aligned} p_{k-1|k}(x_{k-1}|x_k) &= \frac{p_{k-1}(x_{k-1})p_{k|k-1}(x_k|x_{k-1})}{p_k(x_k)} \\ &= \frac{\phi_{k-1}(x_{k-1})p^{\text{ref}}_{k-1}(x_{k-1})p_{k|k-1}(x_k|x_{k-1})}{\phi_k(x_k)p^{\text{ref}}_k(x_{k-1})} \\ &= \frac{\phi_{k-1}(x_{k-1})p^{\text{ref}}_{k-1|k}(x_{k-1}|x_k)}{\phi_k(x_k)}, \end{aligned}$$

where we have used the fact that $p_k(x_k) = \phi_k(x_k)p^{\text{ref}}_k(x_k)$, $p^{\text{ref}}_{k|k-1}(x_k|x_{k-1}) = p_{k|k-1}(x_k|x_{k-1})$ and Bayes' rule again. Now we have $\int p_{k-1|k}(x_{k-1}|x_k)\mathrm{d}x_{k-1} = 1$ for any x_k, so the forward recursion (19) for the value function follows directly. By iterating this forward recursion, we obtain

$$\phi_k(x_k) = \int \phi_{k-1}(x_{k-1})p^{\text{ref}}_{k-1|k}(x_{k-1}|x_k)\mathrm{d}x_{k-1}, \qquad \phi_0(x_0) = \frac{\pi(x_0)}{p^{\text{ref}}_0(x_0)}. \tag{35}$$

It follows that

$$\phi_k(x_k) = \int \phi_0(x_0)\ \ p^{\text{ref}}_{0|k}(x_0|x_k)\mathrm{d}x_0. \tag{36}$$

□

A.2 Proof of Proposition 3

Proof. We have

$$\ln\left(\frac{q_\theta(y_{0:K})}{p(y_{0:K})}\right) = \ln\left(\frac{q_\theta(y_{0:K})}{p^{\text{ref}}(y_{0:K})}\right) + \ln\left(\frac{\mathcal{N}(y_0; 0, \sigma^2 I)}{\pi(y_0)}\right). \tag{37}$$

Now by construction, we have from (18) that $p^{\text{ref}}_{k-1|k}(y_{K-k+1}|y_{K-k}) = \mathcal{N}(y_{K-k+1}; \sqrt{1-\alpha_k}y_{K-k}, \sigma^2\alpha_k I)$ and from (21) we have $q^{\theta}_{k-1|k}(y_{K-k+1}|y_{K-k}) = \mathcal{N}(y_{K-k+1}; \sqrt{1-\alpha_k}y_{K-k} + 2\sigma^2(1-\sqrt{1-\alpha_k})f_\theta(k, y_{K-k}), \sigma^2\alpha_k I)$.

It follows that

$$\begin{aligned}
&\ln\left(\frac{q_\theta(y_{0:K})}{p^{\text{ref}}(y_{0:K})}\right) \qquad (38)\\
&= \ln\left(\frac{q_K(y_0)}{p^{\text{ref}}_K(y_0)}\right) + \sum_{k=1}^{K} \ln\left(\frac{q^{\theta}_{k-1|k}(y_{K-k+1}|y_{K-k})}{p^{\text{ref}}_{k-1|k}(y_{K-k+1}|y_{K-k})}\right)\\
&= \sum_{k=1}^{K} \frac{1}{2\alpha_k\sigma^2}\Big[||y_{K-k+1} - \sqrt{1-\alpha_k}y_{K-k}||^2 - ||y_{K-k+1} - \sqrt{1-\alpha_k}y_{K-k} - 2\sigma^2(1-\sqrt{1-\alpha_k})f(k, y_{K-k})||^2\Big]
\end{aligned}$$

where we have exploited the fact that $q_K(y_0) = p^{\text{ref}}_K(y_0) = \mathcal{N}(y_0; 0, \sigma^2 I)$.

Now using $\epsilon_k := \frac{1}{\sigma\sqrt{\alpha_k}}(y_{K-k+1} - \sqrt{1-\alpha_k}y_{K-k} - 2\sigma^2(1-\sqrt{1-\alpha_k})f(k, y_{K-k}))$, we can rewrite (38) as

$$\begin{aligned}
\ln\left(\frac{q_\theta(y_{0:K})}{p^{\text{ref}}(y_{0:K})}\right) &= \frac{1}{2}\sum_{k=1}^{K}\Big[||\epsilon_k + 2\sigma\frac{(1-\sqrt{1-\alpha_k})}{\sqrt{\alpha_k}}f_\theta(k, y_{K-k})||^2 - ||\epsilon_k||^2\Big]\\
&= 2\sigma^2\sum_{k=1}^{K}\frac{(1-\sqrt{1-\alpha_k})^2}{\alpha_k}||f_\theta(k, y_{K-k})||^2 + 2\sigma\sum_{k=1}^{K}\frac{(1-\sqrt{1-\alpha_k})}{\sqrt{\alpha_k}}f_\theta(k, y_{K-k})^\top\epsilon_k
\end{aligned} \qquad (39)$$

Now (22) follows directly from (37) and (39).

□

A.3 Alternative Kullback–Leibler decomposition

A KL decomposition similar in spirit to the one developed for DDPM (Ho et al., 2020) can be be derived. It leverages the fact that

$$p_{k-1|k,0}(x_{k-1}|x_k, x_0) = \mathcal{N}(x_k; \tilde{\mu}_k(x_k, x_0), \sigma^2\tilde{\beta}_k I) \qquad (40)$$

for $\tilde{\mu}_k(x_k, x_0) = \frac{\sqrt{\bar{\alpha}_{k-1}}\beta_k}{1-\bar{\alpha}_k}x_0 + \frac{\sqrt{\alpha_k}(1-\bar{\alpha}_{k-1})}{1-\bar{\alpha}_k}x_k, \tilde{\beta}_k = \frac{1-\bar{\alpha}_{k-1}}{1-\bar{\alpha}_k}\beta_k$.

Proposition 4. *The reverse KL* $\mathrm{KL}(q^\theta||p)$ *satisfies*

$$\begin{aligned}
\mathrm{KL}(q^\theta||p) = \mathbb{E}_{q^\theta}[\mathrm{KL}(q_K(x_K)||p_{K|0}(x_K|x_0))] + \sum_{k=2}^{K}\mathrm{KL}(q^{\theta}_{k-1|k}(x_{k-1}|x_k)||p_{k-1|0,k}(x_{k-1}|x_0, x_k))\\
+ \mathrm{KL}(q^{\theta}_{0|1}(x_0|x_1)||\pi(x_0))\Big].
\end{aligned}$$

So for $q^{\theta}_{k-1|k}(x_{k-1}|x_k) = \mathcal{N}(x_{k-1}; \sqrt{1-\beta_k}x_k + \sigma^2\beta_k f_\theta(k, x_k), \sigma^2\beta_k I)$*, the terms* $\mathrm{KL}(q^{\theta}_{k-1|k}(x_{k-1}|x_k)||p_{k-1|0,k}(x_{k-1}|x_0, x_k))$ *are KL between two Gaussian distributions and can be calculated analytically.*

We found this elegant decomposition to be however numerically unstable and prone to diverging in our reverse KL setting. Empirically we observed the term $\mathrm{KL}(q^{\theta}_{0|1}(x_0|x_1)||\pi(x_0))$ becoming unstable once approaching the true value of $\ln Z$ this happened since the term $x_1 = x_1^\theta$ has attached gradients and is able to become large making the term $\mathrm{KL}(q^{\theta}_{0|1}(x_0|x_1)||\pi(x_0))$ unstable. Furthermore we notice that this formulation of the KL has no term that prevents x_1 from becoming large thus leading to numerical instabilities, note that this is not an issue with DDPM since in their setting x_1 is fixed and has no gradients.

Proof. The reverse KL can be decomposed as follows

$$\begin{aligned}\mathrm{KL}(q^\theta||p) &= \mathbb{E}_{q^\theta}\Big[\ln\Big(\frac{q^\theta(x_{0:K})}{p(x_{0:K})}\Big)\Big] = \mathbb{E}_{q^\theta}\Big[\ln\Big(\frac{q^\theta(x_{0:K})}{\pi(x_0)p(x_{1:K}|x_0)}\Big)\Big] \\ &= \mathbb{E}_{q^\theta}\Big[\ln\Big(\frac{q^\theta(x_{0:K})}{p(x_{1:K}|x_0)}\Big)\Big] - \mathbb{E}_{q^\theta}[\ln\pi(x_0)] \\ &= L(\theta) - \mathbb{E}_{q^\theta}[\ln\pi(x_0)]\end{aligned}$$

where

$$L(\theta) = \mathbb{E}_{q^\theta}\Big[\ln\Big(\frac{q^\theta(x_{0:K})}{p(x_{1:K}|x_0)}\Big)\Big] = \mathbb{E}_{q^\theta}\Big[\ln q_K(x_K) + \sum_{k=1}^{K}\ln\Big(\frac{q^\theta_{k-1|k}(x_{k-1}|x_k)}{p_{k|k-1}(x_k|x_{k-1})}\Big)\Big]$$

Now using the identity for $k \geq 2$

$$\begin{aligned}p_{k-1,k|0}(x_{k-1},x_k|x_0) &= p_{k-1|0}(x_{k-1}|x_0)p_{k|k-1}(x_k|x_{k-1}) \\ &= p_{k-1|0,k}(x_{k-1}|x_0,x_k)p_{k|0}(x_k|x_0),\end{aligned}$$

we can rewrite $L(\theta)$ as

$$\begin{aligned}L(\theta) &= \mathbb{E}_{q^\theta}\Big[\ln q_K(x_K) + \sum_{k=2}^{K}\ln\Big(\frac{q^\theta_{k-1|k}(x_{k-1}|x_k)}{p_{k-1|0,k}(x_{k-1}|x_0,x_k)}.\frac{p_{k-1|0}(x_{k-1}|x_0)}{p_{k|0}(x_k|x_0)}\Big) + \ln\Big(\frac{q^\theta_{0|1}(x_0|x_1)}{p_{1|0}(x_1|x_0)}\Big)\Big] \\ &= \mathbb{E}_{q^\theta}\Big[\ln\Big(\frac{q_K(x_K)}{p_{K|0}(x_K|x_0)}\Big) + \sum_{k=2}^{K}\ln\Big(\frac{q^\theta_{k-1|k}(x_{k-1}|x_k)}{p_{k-1|0,k}(x_{k-1}|x_0,x_k)}\Big) + \ln q^\theta_{0|1}(x_0|x_1)\Big] \\ &= \mathbb{E}_{q^\theta}[\mathrm{KL}(q_K(x_K)||p_{K|0}(x_K|x_0))] + \sum_{k=2}^{K}\mathrm{KL}(q^\theta_{k-1|k}(x_{k-1}|x_k)||p_{k-1|0,k}(x_{k-1}|x_0,x_k)) \\ &\quad + \ln q^\theta_{0|1}(x_0|x_1)\Big]\end{aligned}$$

The result now follows directly. □

B EXPERIMENTS

In this section we incorporate additional experiments and ablations as well as further experimental detail.

B.1 TUNING HYPER-PARAMETERS

For both DDS and PIS we explore a grid of 25 hyper-parameter values. For PIS we search for the best performing value of the volatility coefficient over 25 values, depending on the task we vary the end points of the drift however we noticed that PIS leads to numerical instabilities for $\gamma > 4$ thus we never search for values larger than this. For DDS we searched over 5 values for σ and 5 values for $\beta_{\max}$, this led to a total of 25 combinations we explored for each experiment. Finally for SMC we searched over 3 of its different step sizes leading to a total of $5^3 = 125$.

B.2 BROWNIAN MOTION MODEL WITH GAUSSIAN OBSERVATION NOISE

The model is given by:

$$\alpha_{\text{inn}} \sim \text{LogNormal}(0,2), \tag{41}$$
$$\alpha_{\text{obs}} \sim \text{LogNormal}(0,2), \tag{42}$$
$$x_1 \sim \mathcal{N}(0,\alpha_{\text{inn}}), \tag{43}$$
$$x_i \sim \mathcal{N}(x_{i-1},\alpha_{\text{inn}}), \quad i = 2,\dots 20 \tag{44}$$
$$x_i \sim \mathcal{N}(x_i,\alpha_{\text{obs}}), \quad i = 1,\dots 30 \tag{45}$$

The goal is to perform inference over the variables $\alpha_{\text{inn}}, \alpha_{\text{obs}}$ and $x_i{}_{i=1}^{3}0$ given the observations $\{y_i\}_{i=1}^{10} \cup \{y_i\}_{i=20}^{30}$.

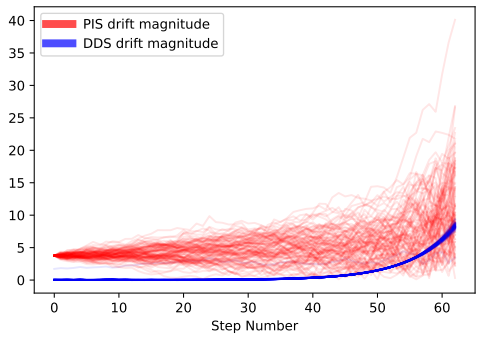

Figure 4: Magnitude comparison of trained networks.

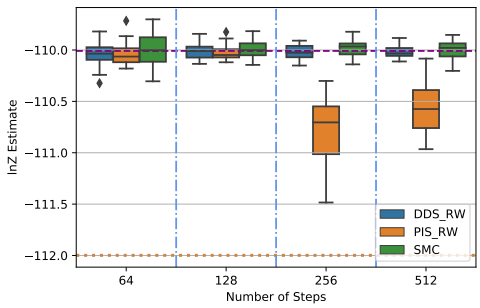

Figure 5: Results on pretrained VAE from Arbel et al. (2021).

B.3 Drift Magnitude

In Figure 4 we compare the magnitudes of the NN drifts learned by PIS and DDS on a $\mathcal{N}(6, \sigma)$ target. The drifts where sampled on randomly evaluated trajectories from each of the learned samplers. We observe the PIS drift to have large magnitudes with higher variance more prone to numerical instabilities, whilst the OU drift has a notably lower magnitude with significantly less variance. We conjecture this is the reason that PIS becomes unstable at training time for a large number of steps across several of our PIS experiments, many of which did not converge due numerical instabilities. For simplicity in this experiment we set β_k to be uniform.

B.4 Variational Encoder in Latent Space

Following Arbel et al. (2021) we explore estimating the posterior of a pre-trained Variational Auto Encoder (VAE) (Kingma & Welling, 2013). We found all approaches to reach a very small error for $\ln Z$ very quickly however as seen with PIS across experiments results became unstable for large T and thus many hyper-parameters when tuning the volatility for PIS led to the loss exploding..

B.4.1 Detached Gradient Ablations

In this section we perform an ablation over our proposed modification of the PIS-Grad network architecture. We compare the same architecture with and without the gradient attached. We found that detaching gradient led to a favourable performance as well as more stable training. Results are presented in table B.4.1 and were carried out for $K = 128$ and using the best found diffusion coefficient from each task.

Method	Funnel	LGCP	VAE	Sonar	Ionosphere	Brownian	NICE
PIS Grad Detach	-0.268 ± 0.07	502.002 ± 0.957	-110.025 ± 0.0704	-109.349 ± 0.356	-111.602 ± 0.145	0.960 ± 0.145	-3.985 ± 0.349
PIS Grad	-0.268 ± 0.09	501.273 ± 0.798	-110.033 ± 0.0647	-109.377 ± 0.361	-111.705 ± 0.190	-5.793 ± 0.499	-3.927 ± 0.666

Table 2: Results for PIS grad with and without detaching the gradient.