Improved sampling via learned diffusions

Lorenz Richter [*1 2] **Julius Berner** [*3] **Guan-Horng Liu** [4]

Abstract

Recently, a series of papers proposed deep learning-based approaches to sample from unnormalized target densities using controlled diffusion processes. In this work, we identify these approaches as special cases of the Schrödinger bridge problem, seeking the most likely stochastic evolution between a given prior distribution and the specified target. We further generalize this framework by introducing a variational formulation based on divergences between path space measures of time-reversed diffusion processes. This abstract perspective leads to practical losses that can be optimized by gradient-based algorithms and includes previous objectives as special cases. At the same time, it allows us to consider divergences other than the reverse Kullback-Leibler divergence that is known to suffer from mode collapse. In particular, we propose the so-called *log-variance loss*, which exhibits favorable numerical properties and leads to significantly improved performance across all considered approaches.

1. Introduction

Given a function $\rho\colon \mathbb{R}^d \to [0,\infty)$, we consider the task of sampling from the density

$$p_{\text{target}} := \frac{\rho}{Z} \quad \text{with} \quad Z := \int_{\mathbb{R}^d} \rho(x)\,\mathrm{d}x,$$

where the normalizing constant Z is typically intractable. This represents a crucial and challenging problem in various scientific fields, such as Bayesian statistics, computational physics, chemistry, or biology (Liu & Liu, 2001; Stoltz et al., 2010). Fueled by the success of *denoising diffusion probabilistic modeling* (Song & Ermon, 2020; Ho et al., 2020; Kingma et al., 2021; Vahdat et al., 2021; Nichol & Dhariwal, 2021) and deep learning approaches to the *Schrödinger bridge* (SB) problem (De Bortoli et al., 2021; Chen et al., 2021a; Koshizuka & Sato, 2022), there is a significant interest in tackling the sampling problem by using stochastic differential equations (SDEs) which are controlled with learned neural networks to transport a given prior density p_{prior} to the target p_{target}.

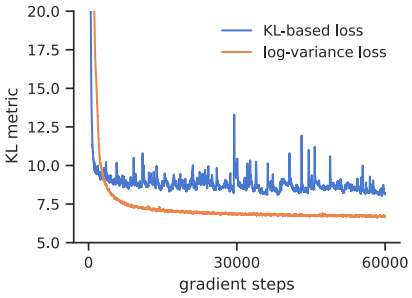

Figure 1. Improved convergence of our proposed log-variance loss for a double well problem.

Recent works include the *Path Integral Sampler* (PIS) and variations thereof (Tzen & Raginsky, 2019; Richter, 2021; Zhang & Chen, 2022; Vargas et al., 2023b), the *Time-Reversed Diffusion Sampler* (DIS) (Berner et al., 2022), as well as the *Denoising Diffusion Sampler* (DDS) (Vargas et al., 2023a). While the ideas for such sampling approaches based on controlled diffusion processes date back to earlier work, see, e.g., Dai Pra (1991); Pavon (1989), the development of corresponding numerical methods based on deep learning has become popular in the last few years.

However, up to now, more focus has been put on *generative modeling*, where samples from p_{target} are available. As a consequence, it seems that for the classical sampling problem, i.e., having only an analytical expression for $\rho \propto p_{\text{target}}$, but no samples, diffusion-based methods cannot reach state-of-the-art performance yet. Potential drawbacks might be stability issues during training, the need to differentiate through SDE solvers, or mode collapse due to the usage of objectives based on reverse Kullback-Leibler (KL) divergences, see, for instance, Zhang & Chen (2022); Vargas et al. (2023a).

[*]Equal contribution (the author order was determined by `numpy.random.rand(1)`) [1]Zuse Institute Berlin [2]dida Datenschmiede GmbH [3]University of Vienna [4]Georgia Institute of Technology. Correspondence to: Lorenz Richter <lorenz.richter@dida.do>.

Workshop on New Frontiers in Learning, Control, and Dynamical Systems at the International Conference on Machine Learning (ICML), Honolulu, Hawaii, USA, 2023. Copyright 2023 by the author(s).

In this work, we overcome these issues and advance the potential of sampling via learned diffusion processes toward more challenging problems. Our contributions can be summarized as follows:

- We provide a unifying framework for recently developed sampling methods based on learned diffusions, i.e., DIS, DDS, and PIS, from the perspective of measures on path space and time-reversals of controlled stochastic processes.
- This path space perspective, in consequence, allows us to consider arbitrary divergences for the optimization objective, whereas existing methods are solely based on minimizing a reverse KL divergence, which is prone to mode collapse.
- In particular, we propose the log-variance divergence that avoids differentiation through the SDE solver and allows to balance exploration and exploitation, resulting in significantly improved numerical stability and performance, see Figure 1.

1.1. Related work

We build our theoretical foundation on the variational formulation of SB problems proposed by Chen et al. (2021a). While the numerical treatment of SB problems has classically been approached via iterative nested schemes, the approach in Chen et al. (2021a) uses backward SDEs (BSDEs) to arrive at a single objective based on a KL divergence. This objective includes the (continuous-time) ELBO of diffusion models (Huang et al., 2021) as a special case, which can also be approached from the perspective of optimal control (Berner et al., 2022). For additional previous work on optimal control in the context of generative modeling, we refer to De Bortoli et al. (2021); Tzen & Raginsky (2019); Pavon (2022); Holdijk et al. (2022).

We extend the variational formulation of SBs to different divergences and, in particular, propose the log-variance loss that has originally been introduced in Nüsken & Richter (2021). Variants of this loss have previously only been analyzed in the context of variational inference (Richter et al., 2020) and neural solvers for partial differential equations (PDEs) (Richter & Berner, 2022). Different from previous works, our objective incorporates the path space measures of two controlled SDEs.

However, we also show that this objective, as well as the one in Chen et al. (2021a), does in general not have a unique solution as it lacks the entropy constraint of classical SB problems. Specifically, we provide a new derivation only relying on time-reversals of diffusion processes. The underlying ideas were established decades ago (Nelson, 1967; Anderson, 1982; Haussmann & Pardoux, 1986; Föllmer, 1988), however, only recently applied to diffusion models (Song et al., 2020) and SBs (Chen et al., 2021a; Vargas, 2021; Liu et al., 2022). In special cases, we recover unique objectives corresponding to recently developed sampling methods, i.e., DIS, DDS, and PIS.

The most common methods to sample from unnormalized densities and compute normalizing constants are arguably Monte Carlo (MC) techniques. Specialized variations of, e.g., *Annealed Importance Sampling* (AIS) (Neal, 2001) or *Sequential Monte Carlo* (Del Moral et al., 2006; Doucet et al., 2009) (SMC) are often regarded as the "gold standard" in the literature.

Even though MCMC methods are guaranteed to converge to the target distribution under mild assumptions, the convergence speed might be too slow in many practical settings (Robert et al., 1999). Variational methods such as mean-field approximations (Wainwright et al., 2008) and normalizing flows (Papamakarios et al., 2021; Wu et al., 2020; Midgley et al., 2022) provide an alternative. By fitting a parametric family of tractable distributions to the target density, the problem of density estimation is cast into an optimization problem.

As already observed in Berner et al. (2022); Vargas et al. (2023a), we note that one cannot directly leverage the connection of diffusion models to score matching (Hyvärinen & Dayan, 2005) for the application of sampling from densities. However, the score matching objective has been employed to approximate the extended target distribution needed in the importance sampling step of AIS methods (Doucet et al., 2022) and also in combination with importance sampling using the likelihood of the partially-trained diffusion model as proposal distribution (Jing et al., 2022). In this work, however, we want to focus on variational methods that directly fit a parametric family of tractable distributions (given by controlled SDEs) to the target density.

1.2. Outline of the article

The rest of the article is organized as follows. In Section 2 we provide an introduction to diffusion-based sampling from the perspective of path space measures and time-reversed SDEs. This can be understood as a unifying framework allowing for generalizations to divergences other than the KL divergence. We propose the *log-variance divergence* and prove that it exhibits superior properties. In Section 3, we will subsequently outline multiple connections to known methods, such as SBs in Section 3.1, diffusion-based generative modeling (i.e., DIS) in Section 3.2, and approaches based on reference processes (i.e., PIS and DDS) in Section 3.3. For all considered methods, we can find compelling numerical evidence for the superiority of the log-variance divergence, see Section 4.

2. Diffusion-based sampling

In this section, we will reformulate our sampling problem as a time-reversal of diffusion processes. This perspective can be understood as a change of measure on path space and we present two divergences that lend themselves to practical objectives. Let us first define our notation and setting.

2.1. Notation and setting

We denote the density of a random variable X by p_X. For a suitable $\mathbb{R}^d$-valued stochastic process $X = (X_t)_{t\in[0,T]}$ we define its density p_X w.r.t. to the Lebesgue measure by

$$p_X(\cdot, t) := p_{X_t}, \quad t \in [0, T].$$

For a suitable functions $f \in C(\mathbb{R}^d \times [0,T], \mathbb{R})$ and $w \in C(\mathbb{R}^d \times [0,T], \mathbb{R}^d)$, we further define

$$R_f(X) := \int_0^T f(X_s, s)\,\mathrm{d}s \tag{1}$$

and

$$S_w(X) := \int_0^T w(X_s, s) \cdot \mathrm{d}W_s, \tag{2}$$

where W is a standard d-dimensional Brownian motion. We denote by $\mathcal{P}$ the set of all probability measures on the space of continuous functions $C([0,T], \mathbb{R}^d)$ and define the path space measure $\mathbb{P}_X \in \mathcal{P}$ as the law of X. For a time-dependent function μ, we denote by $\overleftarrow{\mu}$ the time-reversal given by

$$\overleftarrow{\mu}(t) := \mu(T - t).$$

Finally, we assume that the coefficient functions of all appearing SDEs are sufficiently regular such that Novikov's condition is satisfied and such that the SDEs admit unique strong solutions with smooth and strictly positive densities p_{X_t} for $t \in (0, T)$.

2.2. Sampling as time-reversal problem

The goal of diffusion-based sampling is to sample from the density $p_{\text{target}} = \frac{\rho}{Z}$ by transporting a prior density p_{prior} via controlled stochastic processes. We consider the processes described by the generative SDE

$$\begin{cases} \mathrm{d}X_s^u = (\mu + \sigma u)(X_s^u, s)\,\mathrm{d}s + \sigma(s)\,\mathrm{d}W_s, \\ X_0^u \sim p_{\text{prior}}, \end{cases} \tag{3}$$

and the inference SDE

$$\begin{cases} \mathrm{d}Y_s^v = (-\overleftarrow{\mu} + \overleftarrow{\sigma}\overleftarrow{v})(Y_s^v, s)\,\mathrm{d}s + \overleftarrow{\sigma}(s)\,\mathrm{d}W_s, \\ Y_0^v \sim p_{\text{target}}, \end{cases} \tag{4}$$

where we aim to identify control functions $u, v \in \mathcal{U}$ in a suitable space of admissible controls

$$\mathcal{U} \subset C(\mathbb{R}^d \times [0,T], \mathbb{R}^d)$$

in order to achieve $X_T^u \sim p_{\text{target}}$ and $Y_T^v \sim p_{\text{prior}}$. Specifically, we seek controls satisfying

$$p_{\text{prior}} \underset{Y^v}{\overset{X^u}{\rightleftarrows}} p_{\text{target}}$$

in the sense that Y^v is the time-reversed process of X^u and vice versa, i.e., $\overleftarrow{p}_{X^u} = p_{Y^v}$. In this context, we recall the following well-known results on the time-reversal of stochastic processes (Nelson, 1967; Anderson, 1982; Haussmann & Pardoux, 1986; Föllmer, 1988).

Lemma 2.1 (Time-reversed SDEs)**.** *The time-reversed SDE $\overleftarrow{Y}^v$ given by*

$$\begin{cases} \mathrm{d}\overleftarrow{Y}_s^v = \big(\mu + \sigma\tilde{u}\big)(\overleftarrow{Y}_s^v, s)\,\mathrm{d}s + \sigma(s)\,\mathrm{d}W_s, \\ \overleftarrow{Y}_0^v \sim Y_T^v, \end{cases} \tag{5}$$

with $\tilde{u} := \sigma^\top \nabla \log \overleftarrow{p}_{Y^v} - v$ satisfies that $p_{\overleftarrow{Y}^v} = \overleftarrow{p}_{Y^v}$.

Proof. The result can be derived by comparing the Fokker-Planck equations governing $p_{\overleftarrow{Y}^v}$ and p_{Y^v}, see, e.g., Chen et al. (2021a); Huang et al. (2021); Berner et al. (2022). □

We can now view Problem 2.2 from the perspective of path space measures on the space of trajectories $C([0,T], \mathbb{R}^d)$, as detailed in the sequel.

Problem 2.2 (Time-reversal)**.** Let $\mathbb{P}_{X^u}$ be the path space measure of the process X^u defined in (3) and $\mathbb{P}_{\overleftarrow{Y}^v}$ the path space measure of $\overleftarrow{Y}^v$, the time-reversal of Y^v, given in (5). Further, let

$$D : \mathcal{P} \times \mathcal{P} \to \mathbb{R}_{\geq 0}$$

be a divergence, i.e., a non-negative function satisfying that $D(\mathbb{P}, \mathbb{Q}) = 0$ if and only if $\mathbb{P} = \mathbb{Q}$. We aim to find optimal controls u^*, v^* s.t.

$$u^*, v^* \in \underset{u,v\in\mathcal{U}\times\mathcal{U}}{\arg\min}\, D\left(\mathbb{P}_{X^u} | \mathbb{P}_{\overleftarrow{Y}^v}\right). \tag{6}$$

Let us note that Problem 2.2 aims to reverse the processes X^u and Y^v with respect to each other while obeying the respective initial values $X_0^u \sim p_{\text{prior}}$ and $Y_0^v \sim p_{\text{target}}$, as defined in (3) and (4). For the actual computation of suitable divergences, the following formula will be helpful.

Proposition 2.3 (Likelihood of path measures)**.** *Let X^w be a process as defined in (3) with u being replaced by $w \in \mathcal{U}$. We can compute the Radon-Nikodym derivative as*

$$\frac{\mathrm{d}\mathbb{P}_{X^u}}{\mathrm{d}\mathbb{P}_{\overleftarrow{Y}^v}}(X^w) = Z \exp\left(R_{f^{\text{SB}}_{u,v,w}} + S_{u+v} + B\right)(X^w) \tag{7}$$

with

$$B(X^w) := \log \frac{p_{\text{prior}}(X_0^w)}{\rho(X_T^w)}$$

and

$$f^{\mathrm{SB}}_{u,v,w} := (u+v)\cdot\left(w+\frac{v-u}{2}\right)+\nabla\cdot(\sigma v-\mu),$$

where S and R are defined as in (1) and (2).

Proof. The proof mainly relies on Girsanov's theorem, Itô's lemma, and the HJB equation governing $\log p_{\overleftarrow{Y}^v}$, see Appendix A.1. □

Using the representation of the Radon-Nikodym derivative in Proposition 2.3, we may, in principle, choose any suitable divergence in order to approach Problem 2.2. In the following we will analyze the default setting, i.e., a KL divergence, and propose an alternative, the so-called log-variance divergence, which offers several numerical advantages.

2.3. Comparison of the KL and log-variance divergence

Most works in the context of diffusion-based sampling rely on the KL divergence. Choosing $D = D_{\mathrm{KL}}$, which implies $w = u$ in (7), we can readily compute

$$D_{\mathrm{KL}}(\mathbb{P}_{X^u}|\mathbb{P}_{\overleftarrow{Y}^v}) = \mathbb{E}\left[\left(R_{f^{\mathrm{SB}}_{u,v,u}}+B\right)(X^u)\right]+\log Z$$

with

$$f^{\mathrm{SB}}_{u,v,u} = \frac{\|u+v\|^2}{2}+\nabla\cdot(\sigma v-\mu)\,,$$

where we used the fact that the stochastic integral S_{u+v} has vanishing expectation. Note that in practice we minimize the objective

$$\mathcal{L}_{\mathrm{KL}}(u,v) := D_{\mathrm{KL}}(\mathbb{P}_{X^u}|\mathbb{P}_{\overleftarrow{Y}^v})-\log Z. \tag{8}$$

This objective is identical to the one derived in Chen et al. (2021a) for the SB problem, see also Section 3.1 and Appendix A.2. However, the KL divergence is known to have some evident drawbacks, such as mode collapse (Minka et al., 2005; Midgley et al., 2022) or a potentially high variance of Monte Carlo estimators (Roeder et al., 2017). To address those issues, we propose another divergence that has been originally suggested in Nüsken & Richter (2021).

Definition 2.4 (Log-variance divergence)**.** Let $\widetilde{\mathbb{P}}$ be a reference measure. The log-variance divergence between the measures $\mathbb{P}$ and $\mathbb{Q}$ w.r.t. the reference $\widetilde{\mathbb{P}}$ is defined as

$$D^{\widetilde{\mathbb{P}}}_{\mathrm{LV}}(\mathbb{P},\mathbb{Q}) := \mathbb{V}_{\widetilde{\mathbb{P}}}\left[\log\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}}\right].$$

Note that the log-variance divergence is symmetric in $\mathbb{P}$ and $\mathbb{Q}$ and actually corresponds to a family of divergences, parametrized by the reference measure $\widetilde{\mathbb{P}}$. Obvious choices in our setting are

$$\widetilde{\mathbb{P}} := \mathbb{P}_{X^w}, \quad \mathbb{P} := \mathbb{P}_{X^u}, \quad \text{and} \quad \mathbb{Q} := \mathbb{P}_{\overleftarrow{Y}^v},$$

resulting in the log-variance loss

$$\begin{aligned}\mathcal{L}^w_{\mathrm{LV}}(u,v) &:= D^{\mathbb{P}_{X^w}}_{\mathrm{LV}}(\mathbb{P}_{X^u},\mathbb{P}_{\overleftarrow{Y}^v})\\ &= \mathbb{V}\left[\left(R_{f^{\mathrm{SB}}_{u,v,w}}+S_{u+v}+B\right)(X^w)\right].\end{aligned} \tag{9}$$

Since the variance is shift-invariant, we can omit $\log Z$ in the above objective.

Compared to the KL-based loss (8), the log-variances loss (9) exhibits the following beneficial properties. First, by the choice of the reference measure $\mathbb{P}_{X^w}$, one can balance exploitation and exploration. To exploit the current control u, one can set

$$w = u,$$

but one can also deviate from this control to prevent mode collapse. Next, note that the log-variance loss in (9) does not require the derivative of the process X^w w.r.t. the control w (which, for the case $w = u$, is implemented by detaching or stopping the gradient, see Appendix A.4). In contrast, the KL-based loss in (8) demands to differentiate X^u w.r.t. the control u, requiring to differentiate through the SDE solver and resulting in higher computational costs. Particularly interesting is the following property, sometimes referred to as *sticking-the-landing* (Roeder et al., 2017). It states that the gradients of the log-variance loss have zero variance at the optimal solution. This property does, in general, not hold for the KL-based loss, such that variants of gradient descent might oscillate around the optimum.

Proposition 2.5 (Robustness at the solution)**.** *Let $\widehat{\mathcal{L}}_{\mathrm{LV}}$ be the Monte Carlo estimator of the log-variance loss in (9) and let the controls $u = u_\theta$ and $v = v_\gamma$ be parametrized by θ and γ. The variances of the respective derivatives vanish at the optimal solution $(u^*, v^*) = (u_{\theta^*}, v_{\gamma^*})$, i.e.*

$$\mathbb{V}\left[\nabla_\theta\widehat{\mathcal{L}}^w_{\mathrm{LV}}(u_{\theta^*},v_{\gamma^*})\right] = 0$$

and

$$\mathbb{V}\left[\nabla_\gamma\widehat{\mathcal{L}}^w_{\mathrm{LV}}(u_{\theta^*},v_{\gamma^*})\right] = 0,$$

for all $w \in \mathcal{U}$. For the Monte Carlo estimator $\widehat{\mathcal{L}}_{\mathrm{KL}}$ of the KL-based loss in (8) the above variances are, in general, not vanishing.

Proof. The derivative and its variance can be calculated using Proposition 2.3, see Appendix A.1 and Nüsken & Richter (2021). □

3. Connections and equivalences of diffusion-based sampling approaches

In general, there are infinitely many solutions to Problem 2.2 and, in particular, to our objectives (8) and (9). In fact,

Girsanov's theorem shows that the objectives only enforce Nelson's identity (Nelson, 1967), i.e.,

$$u^* + v^* = \sigma^\top \nabla \log p_{X^{u^*}} = \sigma^\top \nabla \log \overleftarrow{p}_{Y^{v^*}}, \tag{10}$$

see also the proof of Proposition 2.3. In this section, we show how our setting generalizes existing diffusion-based sampling approaches which in turn ensure unique solutions to Problem 2.2. For each approach, we will derive the corresponding versions of the log-variance loss (9).

3.1. Schrödinger bridge problem (SB)

Out of all possible solutions u^* fulfilling (10), the *Schrödinger bridge problem* considers the solution u^* that minimizes the KL divergence

$$D_{\mathrm{KL}}(\mathbb{P}_{X^{u^*}} | \mathbb{P}_{X^r})$$

to a given reference process X^r, defined as in (3) with u replaced by $r \in \mathcal{U}$. Traditionally, the choice $r = 0$, i.e., the uncontrolled process X^0 is considered. Defining

$$f^{\mathrm{ref}}_{u,r,w} := (u - r) \cdot \Big(w - \frac{u + r}{2}\Big), \tag{11}$$

Girsanov's theorem shows that

$$\frac{\mathrm{d}\mathbb{P}_{X^u}}{\mathrm{d}\mathbb{P}_{X^r}}(X^w) = \exp\big(R_{f^{\mathrm{ref}}_{u,r,w}} + S_{u-r}\big)(X^w), \tag{12}$$

which implies that

$$D_{\mathrm{KL}}(\mathbb{P}_{X^u} | \mathbb{P}_{X^r}) = \mathbb{E}\left[R_{f^{\mathrm{ref}}_{u,r,u}}(X^u)\right], \tag{13}$$

see, e.g., Nüsken & Richter (2021, Lemma A.1) and the proof of Proposition 2.3. The SB objective can thus be written as

$$\min_{u \in \mathcal{U}} \mathbb{E}\left[R_{f^{\mathrm{ref}}_{u,r,u}}(X^u) \middle| X^u_T \sim p_{\mathrm{target}}\right], \tag{14}$$

see De Bortoli et al. (2021); Caluya & Halder (2021); Pavon & Wakolbinger (1991); Benamou & Brenier (2000); Chen et al. (2021b); Bernton et al. (2019). We note that the above can also be interpreted as an entropy-regularized *optimal transport* problem (Léonard, 2014). The entropy constraint in (13) can also be combined with our objective in (6) by considering, for instance,

$$\min_{u,v \in \mathcal{U} \times \mathcal{U}} \left\{\mathbb{E}\left[R_{f^{\mathrm{ref}}_{u,r,u}}(X^u)\right] + \lambda D\left(\mathbb{P}_{X^u} | \mathbb{P}_{\overleftarrow{Y}^v}\right)\right\},$$

where $\lambda \in (0, \infty)$ is a sufficiently large Lagrange multiplier. In Appendix A.2 we show how the SB problem (14) can be reformulated as a system of coupled PDEs or BSDEs, which can alternatively be used to regularize Problem 2.2, see also Liu et al. (2022); Koshizuka & Sato (2022). Interestingly, the BSDE system recovers our KL-based objective in (8), as originally derived in Chen et al. (2021a).

Note that via Nelson's identity (10), an optimal solution u^* to the SB problem uniquely defines an optimal control v^* and vice versa. For special cases of SBs, we can calculate such v^* or an approximation

$$\bar{v} \approx v^*.$$

Fixing $v = \bar{v}$ in (6) and only optimizing for u appearing in the generative process (3) then allows us to attain unique solutions to (an approximation of) Problem 2.2. We note that the approximation $\bar{v} \approx v^*$ incurs an irreducible loss given by

$$\frac{\mathrm{d}\mathbb{P}_{X^{u^*}}}{\mathrm{d}\mathbb{P}_{\overleftarrow{Y}^{\bar{v}}}}(X^w) = \frac{\mathrm{d}\mathbb{P}_{\overleftarrow{Y}^{v^*}}}{\mathrm{d}\mathbb{P}_{\overleftarrow{Y}^{\bar{v}}}}(X^w), \tag{15}$$

thus requiring an informed choice of $\bar{v}$ and p_{prior}, such that $Y^{\bar{v}} \approx Y^{v^*}$. We will consider two such settings in the following sections.

3.2. Diffusion-based generative modeling (DIS)

We may set

$$\bar{v} := 0,$$

which can be interpreted as a SB with

$$u^* = r = \sigma^\top \nabla \log \overleftarrow{p}_{Y^0}$$

and $p_{\mathrm{prior}} = p_{Y^0_T}$, such that the entropy constraint (13) can be minimized to zero. Note though, that this only leads to feasible sampling approaches if the functions μ and σ in the SDEs are chosen such that the distribution of $p_{Y^0_T}$ is (approximately) known and such that we can easily sample from it. In practice, one chooses functions μ and σ such that

$$p_{Y^0_T} \approx p_{\mathrm{prior}} := \mathcal{N}(0, \nu^2 \mathrm{I}),$$

see Section A.4. Related approaches are often called *diffusion-based generative modeling* or *denoising diffusion probabilistic modeling* since the (optimally controlled) generative process X^{u^*} can be understood as the time-reversal of the process Y^0 that moves samples from the target density to Gaussian noise (Ho et al., 2020; Pavon, 1989; Huang et al., 2021; Song et al., 2020).

Let us recall the notation from Proposition 2.3 and define

$$f^{\mathrm{DIS}}_{u,w} := f^{\mathrm{SB}}_{u,0,w} = u \cdot w - \frac{\|u\|^2}{2} - \nabla \cdot \mu.$$

Setting $v = 0$ in (8), we directly get the loss

$$\mathcal{L}_{\mathrm{KL}}(u) = \mathbb{E}\left[\big(R_{f^{\mathrm{DIS}}_{u,u}} + B\big)(X^u)\right],$$

which corresponds to the *Time-Reversed Diffusion Sampler* (DIS) derived in Berner et al. (2022). From (9), we analogously obtain the related log-variance loss

$$\mathcal{L}^w_{\mathrm{LV}}(u) = \mathbb{V}\left[\big(R_{f^{\mathrm{DIS}}_{u,w}} + S_u + B\big)(X^w)\right].$$

3.3. Time-reversal of reference processes (PIS & DDS)

In general, we may also set

$$\bar{v} := \sigma^\top \nabla \log p_{X^r} - r.$$

Via Lemma 2.1 this choice implies

$$\mathbb{P}_{X^r} = \mathbb{P}_{\overleftarrow{Y}^{\bar{v},\mathrm{ref}}}, \tag{16}$$

where $Y^{v,\mathrm{ref}}$ is the process Y^v as in (4), however, with p_{target} replaced by the density $p_{\mathrm{ref}} := p_{X_T^r}$, i.e.,

$$\begin{cases} \mathrm{d}Y^{v,\mathrm{ref}} = (-\bar{\mu} + \bar{\sigma}\bar{v})(Y^{v,\mathrm{ref}}, s)\,\mathrm{d}s + \bar{\sigma}(s)\,\mathrm{d}W_s, \\ Y^{v,\mathrm{ref}} \sim p_{\mathrm{ref}}. \end{cases}$$

Since $Y^{\bar{v},\mathrm{ref}}$ is the time-reversal of the reference process X^r, we note that the optimal control v^*, corresponding to the solution u^* of the SB problem in (14), minimizes

$$D_{\mathrm{KL}}(\mathbb{P}_{Y^v} | \mathbb{P}_{\overleftarrow{Y}^{\bar{v},\mathrm{ref}}})$$

among all controls v with $Y_T^v \sim p_{\mathrm{prior}}$.

Using (7) with p_{ref} instead of $p_{\mathrm{target}} = \frac{\rho}{Z}$, we obtain that

$$\begin{aligned} 1 &= \frac{\mathrm{d}\mathbb{P}_{X^r}}{\mathrm{d}\mathbb{P}_{\overleftarrow{Y}^{\bar{v},\mathrm{ref}}}}(X^w) \\ &= \frac{p_{\mathrm{prior}}(X_0^w)}{p_{\mathrm{ref}}(X_T^w)} \exp\left(R_{f^{\mathrm{SB}}_{r,\bar{v},w}} + S_{r+\bar{v}}\right)(X^w). \end{aligned} \tag{17}$$

This leads to the following alternative representation of Proposition 2.3.

Lemma 3.1 (Likelihood w.r.t. reference process)**.** *Assuming* (16)*, it holds that*

$$\frac{\mathrm{d}\mathbb{P}_{X^u}}{\mathrm{d}\mathbb{P}_{\overleftarrow{Y}^{\bar{v}}}}(X^w) = Z \exp\left(R_{f^{\mathrm{ref}}_{u,r,w}} + S_{u-r} + B^{\mathrm{ref}}\right)(X^w),$$

where $f^{\mathrm{ref}}_{u,r,w}$ *is defined as in* (11) *and*

$$B^{\mathrm{ref}}(X^w) := \log \frac{p_{\mathrm{ref}}}{\rho}(X_T^w).$$

Proof. The result follows from dividing (7) by (17). □

Note that computing the Radon-Nikodym derivative in Lemma 3.1 requires to choose r, p_{prior}, μ, and σ such that $p_{\mathrm{ref}} = p_{X_T^r}$ is tractable[1]. For suitable choices of r (see below), one can, for instance, use the SDEs with tractable densities stated in Appendix A.3 with $p_{\mathrm{prior}} = \delta_{x_0}$, $p_{\mathrm{prior}} = \mathcal{N}(0, \nu^2 \mathrm{I})$, or a mixture of such distributions. Recalling (15) and the choice

$$\bar{v} := \sigma^\top \nabla \log p_{X^r} - r,$$

we also need to guarantee that $Y^{\bar{v}} \approx Y^{v^*}$. Let us now outline two such cases in the following.

[1] In general, it suffices to be able to compute $p_{X_T^r}$ up to its normalizing constant.

PIS: We first consider the case $r := 0$. Lemma 3.1 and choosing $D = D_{\mathrm{KL}}$ in Problem 2.2 then yield the objective

$$\begin{aligned} \mathcal{L}_{\mathrm{KL}}(u) &= D_{\mathrm{KL}}(\mathbb{P}_{X^u} | \mathbb{P}_{\overleftarrow{Y}^{\bar{v}}}) - \log Z \\ &= \mathbb{E}\left[\left(R_{f^{\mathrm{ref}}_{u,0,u}} + B^{\mathrm{ref}}\right)(X^u)\right]. \end{aligned}$$

This objective has previously been considered by Tzen & Raginsky (2019); Dai Pra (1991) and corresponding numerical algorithms, referred to as *Path Integral Sampler* (PIS) in Zhang & Chen (2022), have been independently presented in Richter (2021); Zhang & Chen (2022); Vargas et al. (2023b). Choosing $D = D_{\mathrm{LV}}$, we get the corresponding log-variance loss

$$\mathcal{L}^w_{\mathrm{LV}}(u) = \mathbb{V}\left[\left(R_{f^{\mathrm{ref}}_{u,0,w}} + S_u + B^{\mathrm{ref}}\right)(X^w)\right],$$

which has already been stated in Richter (2021). Typically, this objective is used with

$$p_{\mathrm{prior}} := \delta_{x_0},$$

since Doob's h-transform guarantees that $\bar{v} = v^*$, i.e., we can solve the SB exactly, see Rogers & Williams (2000) and also Appendix A.2.1. In this special case, the SB is often referred to as a *Schrödinger half-bridge*.

DDS: Next, we consider the choices

$$r := \sigma^\top \nabla \log \overleftarrow{p}_{Y^{0,\mathrm{ref}}}, \quad \bar{v} := 0, \quad \text{and} \quad p_{\mathrm{prior}} := p_{Y_T^{0,\mathrm{ref}}},$$

which yields a special case of the setting from Section 3.2. Using Lemma 3.1, we obtain the objective

$$\mathcal{L}_{\mathrm{KL}}(u) = \mathbb{E}\left[\left(R_{f^{\mathrm{ref}}_{u,r,u}} + B^{\mathrm{ref}}\right)(X^u)\right].$$

This corresponds to the *Denoising Diffusion Sampler* (DDS) objective stated by Vargas et al. (2023a) when choosing μ and σ such that Y^0 is a VP SDE, see Appendix A.3. Choosing the invariant distribution

$$p_{\mathrm{ref}} := \mathcal{N}(0, \nu^2 \mathrm{I})$$

of the VP SDE, see (33) in the appendix, we have that

$$p_{X^r}(\cdot, t) = \overleftarrow{p}_{Y^{0,\mathrm{ref}}}(\cdot, t) = p_{\mathrm{ref}} = p_{\mathrm{prior}}, \quad t \in [0, T],$$

and, in particular,

$$r(x, t) = -\frac{\sigma^\top x}{\nu^2}.$$

The corresponding log-variance loss can now readily be computed as

$$\mathcal{L}^w_{\mathrm{LV}}(u) = \mathbb{V}\left[\left(R_{f^{\mathrm{ref}}_{u,r,w}} + S_{u-r} + B^{\mathrm{ref}}\right)(X^w)\right].$$

We refer to Table 1 for a comparison of our objectives.

Table 1. Comparison of the objectives with $f^{\mathrm{SB}}_{u,v,w} := (u+v)\cdot\left(w+\frac{v-u}{2}\right)+\nabla\cdot(\sigma v-\mu)$, $f^{\mathrm{ref}}_{u,r,w} := (u-r)\cdot\left(w-\frac{u+r}{2}\right)$, $B(X^w) := \log\frac{p_{\mathrm{prior}}(X^w_0)}{\rho(X^w_T)}$, $B^{\mathrm{ref}}(X^w) := \log\frac{p_{X^r_T}}{\rho}(X^w_T)$, and $p_{\mathrm{target}} := \frac{\rho}{Z}$.

	$\mathcal{L}_{\mathrm{KL}}$	$\mathcal{L}^w_{\mathrm{LV}}$ (ours)	p_{prior}	v	r
SB	$\mathbb{E}\left[\left(R_{f^{\mathrm{SB}}_{u,v,u}}+B\right)(X^u)\right]$	$\mathbb{V}\left[\left(R_{f^{\mathrm{SB}}_{u,v,w}}+S_{u+v}+B\right)(X^w)\right]$	arbitrary	learned	–
DIS	$\mathbb{E}\left[\left(R_{f^{\mathrm{SB}}_{u,0,u}}+B\right)(X^u)\right]$	$\mathbb{V}\left[\left(R_{f^{\mathrm{SB}}_{u,0,w}}+S_{u}+B\right)(X^w)\right]$	$\approx p_{Y^0_T}$	0	–
PIS	$\mathbb{E}\left[\left(R_{f^{\mathrm{ref}}_{u,0,u}}+B^{\mathrm{ref}}\right)(X^u)\right]$	$\mathbb{V}\left[\left(R_{f^{\mathrm{ref}}_{u,0,w}}+S_{u}+B^{\mathrm{ref}}\right)(X^w)\right]$	δ_{x_0}	$\sigma^\top\nabla\log p_{X^0}$	0
DDS	$\mathbb{E}\left[\left(R_{f^{\mathrm{ref}}_{u,r,u}}+B^{\mathrm{ref}}\right)(X^u)\right]$	$\mathbb{V}\left[\left(R_{f^{\mathrm{ref}}_{u,r,w}}+S_{u-r}+B^{\mathrm{ref}}\right)(X^w)\right]$	$p_{Y^{0,\mathrm{ref}}_T}$	0	$\sigma^\top\nabla\log\overleftarrow{p}_{Y^{0,\mathrm{ref}}}$

Remark 3.2 (Alternative derivations of Lemma 3.1)**.** The expression in Lemma 3.1 can also be derived via

$$\begin{aligned}\frac{\mathrm{d}\mathbb{P}_{X^u}}{\mathrm{d}\mathbb{P}_{\overleftarrow{Y}^{\bar{v}}}}(X^w) &= \frac{\mathrm{d}\mathbb{P}_{X^u}}{\mathrm{d}\mathbb{P}_{X^r}}(X^w)\frac{\mathrm{d}\mathbb{P}_{\overleftarrow{Y}^{\bar{v},\mathrm{ref}}}}{\mathrm{d}\mathbb{P}_{\overleftarrow{Y}^{\bar{v}}}}(X^w)\\ &= \frac{\mathrm{d}\mathbb{P}_{X^u}}{\mathrm{d}\mathbb{P}_{X^r}}(X^w)\frac{p_{X^r_T}}{p_{\mathrm{target}}}(X^w_T),\end{aligned}$$

where the first factor can be computed as in (12). Yet another viewpoint is based on importance sampling in path space, see, e.g., Hartmann et al. (2017). Since our goal is to find an optimal control u^* such that we get samples $X^{u^*}_T \sim p_{\mathrm{target}}$, we may define our target path space measure $\mathbb{P}_{X^{u^*}}$ via

$$\frac{\mathrm{d}\mathbb{P}_{X^{u^*}}}{\mathrm{d}\mathbb{P}_{X^r}}(X^w) = \frac{p_{\mathrm{target}}}{p_{X^r_T}}(X^w_T).$$

We can then compute

$$\frac{\mathrm{d}\mathbb{P}_{X^u}}{\mathrm{d}\mathbb{P}_{X^{u^*}}}(X^w) = \frac{\mathrm{d}\mathbb{P}_{X^u}}{\mathrm{d}\mathbb{P}_{X^r}}(X^w)\frac{\mathrm{d}\mathbb{P}_{X^r}}{\mathrm{d}\mathbb{P}_{X^{u^*}}}(X^w),$$

which, together with (12), is equivalent to the expression in Lemma 3.1. Note that in the importance sampling perspective we do not need the concept of time-reversals.

4. Numerical experiments

In this section, we compare the KL-based loss with the log-variance loss on the three different approaches, SB, PIS, and DIS, introduced in Sections 2.3, 3.2, and 3.3. As DDS can be seen as a special case of DIS (both with $\bar{v} = 0$), we do not consider it separately. For our PIS experiments, we follow the implementation of Zhang & Chen (2022) and only change the objective to $\mathcal{L}^w_{\mathrm{LV}}$ with $w = u$. For DIS, we also use the models from Zhang & Chen (2022), incorporating the density p_{prior} and a variance-preserving SDE (see Appendix A.4) similar to Berner et al. (2022). For SB, we analogously approach the general loss (8), which corresponds to the setting in Chen et al. (2021a) adapted to unnormalized densities.

We can demonstrate that the appealing properties of the log-variance loss can indeed lead to significant performance improvements for all approaches. Note that we always compare the same settings, in particular, the same number of target evaluations, for both the log-variance and KL-based losses and use sufficiently many gradient steps to reach convergence, see Appendix A.4 for details. Still, we observe that qualitative differences between the two losses are consistent across various hyperparameter settings.

4.1. Benchmark problems

We shall evaluate the different methods on the following three numerical benchmark examples.

Gaussian mixture model (GMM): We consider the density

$$\rho(x) = p_{\mathrm{target}}(x) = \frac{1}{m}\sum_{i=1}^{m}\mathcal{N}(x;\mu_i,\Sigma_i).$$

Specifically, we choose $m = 9$, $\Sigma_i = 0.3\,\mathrm{I}$, and

$$(\mu_i)_{i=1}^9 = \{-5,0,5\}\times\{-5,0,5\}\subset\mathbb{R}^2$$

to obtain well-separated modes, see also Figure 2.

Funnel: The 10-dimensional *Funnel distribution* (Neal, 2003) is a challenging example often used to test MCMC methods. It is given by the density

$$\rho(x) = p_{\mathrm{target}}(x) = \mathcal{N}(x_1;0,\eta^2)\prod_{i=2}^{d}\mathcal{N}(x_i;0,e^{x_1})$$

for $x = (x_i)_{i=1}^{10} \in \mathbb{R}^{10}$ with $\eta = 3$.

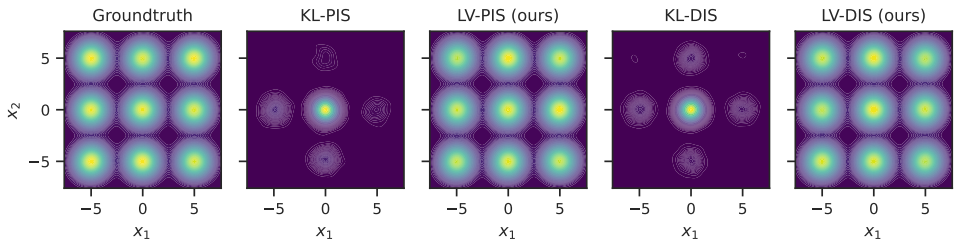

Figure 2. KDE plots of (1) samples from the groundtruth distribution, (2 & 3) PIS with KL divergence and log-variance loss, and (4 & 5) DIS with KL divergence and log-variance loss for the GMM problem (from left to right). One can clearly see that the log-variance loss does not suffer from the mode collapse of the reverse KL divergence, which only recovers the mode of $p_{\text{prior}} = \mathcal{N}(0, \mathrm{I})$.

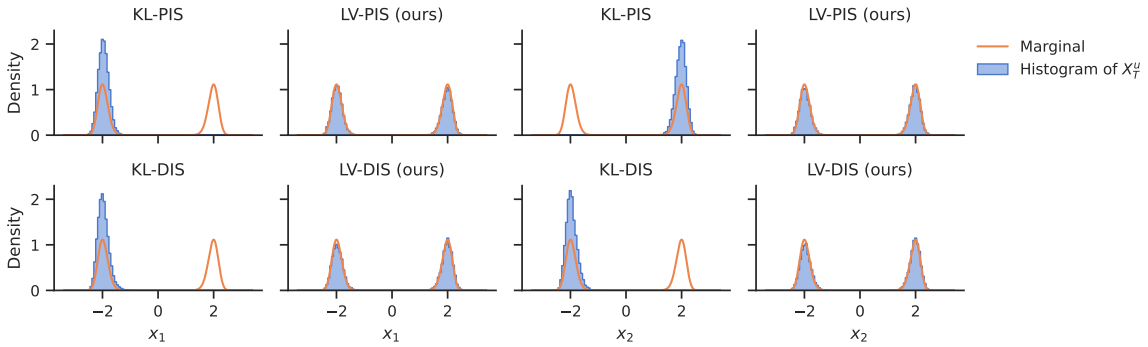

Figure 3. Marginals of the first two coordinates (left and right) of samples from PIS and DIS (top and bottom) for the DW problem with $d = 5$, $m = 5$, $\delta = 4$. Again one observes the mode coverage of the log-variance loss as compared to the reverse KL divergence.

Double well (DW): A typical problem in molecular dynamics considers sampling from the stationary distribution of a Langevin dynamics. In our example we shall consider a d-dimensional *double well* potential, corresponding to the (unnormalized) density

$$\rho(x) = \exp\left(-\sum_{i=1}^{m}(x_i^2 - \delta)^2 - \frac{1}{2}\sum_{i=m+1}^{d} x_i^2\right)$$

with $m \in \mathbb{N}$ combined double wells and a separation parameter $\delta \in (0, \infty)$, see also Wu et al. (2020) and Figure 3. Note that, due to the double well structure of the potential, the density contains 2^m modes. For these multimodal examples we can compute a reference solutions by numerical integration since ρ factorizes in the dimensions.

4.2. Results

Let us start with the SB approach and the general losses in (8) and (9). Table 2 shows that the log-variance loss can improve our considered metrics. However, we note that choosing an appropriate prior p_{prior} (with sufficient mass at the modes of the target) was necessary to make our algorithms converge, see Appendix A.4. While the general setting of SB enables us to incorporate such prior knowledge, the framework suffers from reduced efficiency and numerical instabilities. These issues are commonly observed in the context of SBs (De Bortoli et al., 2021; Chen et al., 2021a; Fernandes et al., 2021) and might be rooted in the non-uniqueness of the optimal control (cf. Section 3.1) and the fact that two controls need to be optimized. For the more challenging problems, the SB approach did not achieve satisfying results. Since such pathologies do not appear in the special cases of DIS and PIS, we shall focus on them separately in the sequel.

We observe that the log-variance loss significantly improves both DIS and PIS across our considered benchmark problems and metrics, see Table 3. The improvements are quite remarkable considering that we only replaced the KL-based loss $\mathcal{L}_{\text{KL}}$ by the log-variance loss $\mathcal{L}_{\text{LV}}$ without tuning the hyperparameter for the latter loss. In the few cases, where the KL divergence performs better, the difference seems rather insignificant. In particular, Figures 2 and 3 show that the log-variance loss successfully counteracts mode-collapse, leading to quite substantial improvements.

5. Conclusion

In this work, we provide a novel unifying perspective on diffusion-based generative modeling that is based on path space measures of time-reversed diffusion processes. In

Table 2. SB metrics for selected benchmark problems of two dimensions d. We report (average) errors for estimating the log-normalizing constant $\log Z$ as well the standard deviations (std) of the marginals. Furthermore, we report the normalized effective sample size (ESS) and, for problems where we can compute reference samples (i.e., GMM and Funnel), we report the Sinkhorn distance $\mathcal{W}_\gamma^2$ (Cuturi, 2013). The arrows ↑ and ↓ indicate whether we want to maximize or minimize a given metric and the better loss is formatted bold.

Problem	Method	Loss	$\Delta \log Z \downarrow$	$\mathcal{W}_\gamma^2 \downarrow$	ESS ↑	Δstd ↓
GMM ($d = 2$)	SB	KL (Chen et al., 2021a)	**0.174**	0.062	0.7416	0.221
		LV (ours)	0.266	**0.058**	**0.9427**	**0.148**
DW ($d = 5, m = 5, \delta = 4$)	SB	KL (Chen et al., 2021a)	14.204	-	0.0157	0.506
		LV (ours)	**12.538**	-	**0.3035**	**0.050**

Table 3. PIS and DIS metrics for the benchmark problems of various dimension d. We report the median over five independent runs, see Figure 4 for a corresponding boxplot. Note that we omit the Sinkhorn distance $\mathcal{W}_\gamma^2$ on DW as ground-truth samples are not available.

Problem	Method	Loss	$\Delta \log Z \downarrow$	$\mathcal{W}_\gamma^2 \downarrow$	ESS ↑	Δstd ↓
GMM ($d = 2$)	PIS	KL (Zhang & Chen, 2022)	2.146	1.587	0.0009	3.425
		LV (ours)	**0.049**	**0.020**	**0.9021**	**0.032**
	DIS	KL (Berner et al., 2022)	1.708	0.080	0.0085	2.718
		LV (ours)	**0.060**	**0.020**	**0.8590**	**0.017**
Funnel ($d = 10$)	PIS	KL (Zhang & Chen, 2022)	0.389	5.947	0.0333	6.880
		LV (ours)	**0.357**	**5.819**	**0.0732**	**6.628**
	DIS	KL (Berner et al., 2022)	0.547	**5.111**	0.1233	5.361
		LV (ours)	**0.499**	5.170	**0.1998**	**5.143**
DW ($d = 5, m = 5, \delta = 4$)	PIS	KL (Zhang & Chen, 2022)	3.672	-	0.0001	1.793
		LV (ours)	**0.228**	-	**0.6537**	**0.003**
	DIS	KL (Berner et al., 2022)	3.989	-	0.0152	1.712
		LV (ours)	**0.401**	-	**0.4335**	**0.002**
DW ($d = 50, m = 5, \delta = 2$)	PIS	KL (Zhang & Chen, 2022)	0.192	-	0.6739	0.004
		LV (ours)	**0.110**	-	**0.8099**	**0.002**
	DIS	KL (Berner et al., 2022)	23.405	-	0.0000	**0.185**
		LV (ours)	**22.889**	-	0.0000	0.186

principle, this perspective allows to consider arbitrary divergences between such measures as objectives for the corresponding task of interest.

While the KL divergence yields already known objectives, we find that choosing the log-variance divergence leads to novel algorithms which are particularly useful for the task of sampling from (unnormalized) densities. Specifically, this divergence exhibits beneficial properties, such as lower variance, computational efficiency, and exploration-exploitation trade-offs. We can demonstrate in multiple numerical examples that the the log-variance loss greatly improves sampling quality across a range of metrics. We believe that problem and approach-specific finetuning might further enhance the performance of the log-variance loss, thereby paving the way for competitive diffusion-based sampling approaches.

Based on our work, one could also explore more divergences, e.g., the family of α-divergences, see Minka et al. (2005). Finally, we anticipate further performance improvements by combining diffusion-based samplers with neural solvers for the optimality PDEs or MCMC methods, as has been successfully done for normalizing flows (Wu et al., 2020; Midgley et al., 2022; Máté & Fleuret, 2023).

Acknowledgements

The research of L.R. was funded by Deutsche Forschungsgemeinschaft (DFG) through the grant CRC 1114 "Scaling Cascades in Complex Systems" (project A05, project number 235221301). The computational results have been achieved in part using the Vienna Scientific Cluster (VSC).

References

Anderson, B. D. Reverse-time diffusion equation models. *Stochastic Processes and their Applications*, 12(3):313–326, 1982.

Benamou, J.-D. and Brenier, Y. A computational fluid mechanics solution to the monge-kantorovich mass transfer problem. *Numerische Mathematik*, 84(3):375–393, 2000.

Berner, J., Richter, L., and Ullrich, K. An optimal control perspective on diffusion-based generative modeling. *arXiv preprint arXiv:2211.01364*, 2022.

Bernton, E., Heng, J., Doucet, A., and Jacob, P. E. Schrödinger Bridge Samplers. *arXiv preprint arXiv:1912.13170*, 2019.

Caluya, K. F. and Halder, A. Wasserstein proximal algorithms for the Schrödinger bridge problem: Density control with nonlinear drift. *IEEE Transactions on Automatic Control*, 67(3):1163–1178, 2021.

Chen, T., Liu, G.-H., and Theodorou, E. A. Likelihood training of Schrödinger Bridge using Forward-Backward SDEs theory. *arXiv preprint arXiv:2110.11291*, 2021a.

Chen, Y., Georgiou, T. T., and Pavon, M. Stochastic control liaisons: Richard Sinkhorn meets Gaspard Monge on a Schrödinger bridge. *SIAM Review*, 63(2):249–313, 2021b.

Cuturi, M. Sinkhorn distances: Lightspeed computation of optimal transport. *Advances in neural information processing systems*, pp. 2292–2300, 2013.

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

De Bortoli, V., Thornton, J., Heng, J., and Doucet, A. Diffusion Schrödinger bridge with applications to score-based generative modeling. *Advances in Neural Information Processing Systems*, 34:17695–17709, 2021.

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.

Doucet, A., Johansen, A. M., et al. A tutorial on particle filtering and smoothing: Fifteen years later. *Handbook of nonlinear filtering*, 12(656-704):3, 2009.

Doucet, A., Grathwohl, W. S., Matthews, A. G., and Strathmann, H. Score-based diffusion meets annealed importance sampling. In *Advances in Neural Information Processing Systems*, 2022.

Fernandes, D. L., Vargas, F., Ek, C. H., and Campbell, N. D. Shooting Schrödinger's cat. In *Fourth Symposium on Advances in Approximate Bayesian Inference*, 2021.

Fleming, W. H. and Soner, H. M. *Controlled Markov processes and viscosity solutions*, volume 25. Springer Science & Business Media, 2006.

Föllmer, H. Random fields and diffusion processes. In *École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87*, pp. 101–203. Springer, 1988.

Hartmann, C., Richter, L., Schütte, C., and Zhang, W. Variational characterization of free energy: Theory and algorithms. *Entropy*, 19(11):626, 2017.

Haussmann, U. G. and Pardoux, E. Time reversal of diffusions. *The Annals of Probability*, pp. 1188–1205, 1986.

Ho, J., Jain, A., and Abbeel, P. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33:6840–6851, 2020.

Holdijk, L., Du, Y., Hooft, F., Jaini, P., Ensing, B., and Welling, M. Path integral stochastic optimal control for sampling transition paths. *arXiv preprint arXiv:2207.02149*, 2022.

Huang, C.-W., Lim, J. H., and Courville, A. C. A variational perspective on diffusion-based generative models and score matching. *Advances in Neural Information Processing Systems*, 34, 2021.

Hyvärinen, A. and Dayan, P. Estimation of non-normalized statistical models by score matching. *Journal of Machine Learning Research*, 6(4), 2005.

Jing, B., Corso, G., Chang, J., Barzilay, R., and Jaakkola, T. Torsional diffusion for molecular conformer generation. *arXiv preprint arXiv:2206.01729*, 2022.

Kingma, D., Salimans, T., Poole, B., and Ho, J. Variational diffusion models. *Advances in Neural Information Processing Systems*, 34:21696–21707, 2021.

Koshizuka, T. and Sato, I. Neural Lagrangian Schrödinger bridge. *arXiv preprint arXiv:2204.04853*, 2022.

Léonard, C. Some properties of path measures. In *Séminaire de Probabilités XLVI*, pp. 207–230. Springer, 2014.

Liu, G.-H., Chen, T., So, O., and Theodorou, E. A. Deep generalized Schrödinger bridge. *arXiv preprint arXiv:2209.09893*, 2022.

Liu, J. S. and Liu, J. S. *Monte Carlo strategies in scientific computing*, volume 10. Springer, 2001.

Máté, B. and Fleuret, F. Learning deformation trajectories of Boltzmann densities. *arXiv preprint arXiv:2301.07388*, 2023.

Midgley, L. I., Stimper, V., Simm, G. N., Schölkopf, B., and Hernández-Lobato, J. M. Flow annealed importance sampling bootstrap. In *NeurIPS 2022 AI for Science: Progress and Promises*, 2022.

Minka, T. et al. Divergence measures and message passing. Technical report, Citeseer, 2005.

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

Neal, R. M. Slice sampling. *The Annals of Statistics*, 31(3): 705–767, 2003.

Nelson, E. Dynamical theories of Brownian motion. *Press, Princeton, NJ*, 1967.

Nichol, A. Q. and Dhariwal, P. Improved denoising diffusion probabilistic models. In *International Conference on Machine Learning*, pp. 8162–8171. PMLR, 2021.

Nüsken, N. and Richter, L. Solving high-dimensional Hamilton–Jacobi–Bellman PDEs using neural networks: perspectives from the theory of controlled diffusions and measures on path space. *Partial Differential Equations and Applications*, 2(4):1–48, 2021.

Papamakarios, G., Nalisnick, E. T., Rezende, D. J., Mohamed, S., and Lakshminarayanan, B. Normalizing flows for probabilistic modeling and inference. *J. Mach. Learn. Res.*, 22(57):1–64, 2021.

Pavon, M. Stochastic control and nonequilibrium thermodynamical systems. *Applied Mathematics and Optimization*, 19(1):187–202, 1989.

Pavon, M. On local entropy, stochastic control and deep neural networks. *arXiv preprint arXiv:2204.13049*, 2022.

Pavon, M. and Wakolbinger, A. On free energy, stochastic control, and Schrödinger processes. In *Modeling, Estimation and Control of Systems with Uncertainty*, pp. 334–348. Springer, 1991.

Pham, H. *Continuous-time Stochastic Control and Optimization with Financial Applications*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2009.

Richter, L. *Solving high-dimensional PDEs, approximation of path space measures and importance sampling of diffusions*. PhD thesis, BTU Cottbus-Senftenberg, 2021.

Richter, L. and Berner, J. Robust SDE-based variational formulations for solving linear PDEs via deep learning. In *International Conference on Machine Learning*, pp. 18649–18666. PMLR, 2022.

Richter, L., Boustati, A., Nüsken, N., Ruiz, F., and Akyildiz, O. D. VarGrad: low-variance gradient estimator for variational inference. *Advances in Neural Information Processing Systems*, 33:13481–13492, 2020.

Robert, C. P., Casella, G., and Casella, G. *Monte Carlo statistical methods*, volume 2. Springer, 1999.

Roeder, G., Wu, Y., and Duvenaud, D. K. Sticking the landing: Simple, lower-variance gradient estimators for variational inference. *Advances in Neural Information Processing Systems*, 30, 2017.

Rogers, L. C. G. and Williams, D. *Diffusions, Markov processes and martingales: Volume 2, Itô calculus*, volume 2. Cambridge university press, 2000.

Siddiqi, A. H. and Nanda, S. *Functional analysis with applications*. Springer, 1986.

Song, Y. and Ermon, S. Improved techniques for training score-based generative models. *Advances in neural information processing systems*, 33:12438–12448, 2020.

Song, Y., Sohl-Dickstein, J., Kingma, D. P., Kumar, A., Ermon, S., and Poole, B. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2020.

Stoltz, G., Rousset, M., et al. *Free energy computations: A mathematical perspective*. World Scientific, 2010.

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

Vahdat, A., Kreis, K., and Kautz, J. Score-based generative modeling in latent space. *Advances in Neural Information Processing Systems*, 34:11287–11302, 2021.

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

Vargas, F., Grathwohl, W., and Doucet, A. Denoising diffusion samplers. In *International Conference on Learning Representations*, 2023a.

Vargas, F., Ovsianas, A., Fernandes, D., Girolami, M., Lawrence, N. D., and Nüsken, N. Bayesian learning via neural Schrödinger–Föllmer flows. *Statistics and Computing*, 33(1):1–22, 2023b.

Wainwright, M. J., Jordan, M. I., et al. Graphical models, exponential families, and variational inference. *Foundations and Trends in Machine Learning*, 1(1–2):1–305, 2008.

Wu, H., Köhler, J., and Noé, F. Stochastic normalizing flows. *Advances in Neural Information Processing Systems*, 33: 5933–5944, 2020.

Zhang, Q. and Chen, Y. Path Integral Sampler: a stochastic control approach for sampling. In *International Conference on Learning Representations*, 2022.

A. Appendix

A.1. Proofs

Proof of Proposition 2.3. Let us define the path space measures $\mathbb{P}_{X^{u,x}}$ and $\mathbb{P}_{\overleftarrow{Y}^{v,x}}$ as the measures of X^u and $\overleftarrow{Y}^v$ conditioned on $X_0^u = x$ and $\overleftarrow{Y}_0^v = x$ with $x \in \mathbb{R}^d$, respectively. We can then compute

$$\log \frac{\mathrm{d}\mathbb{P}_{X^u}}{\mathrm{d}\mathbb{P}_{\overleftarrow{Y}^v}}(X^w) = \log \frac{\mathrm{d}\mathbb{P}_{X^{u,x}}}{\mathrm{d}\mathbb{P}_{\overleftarrow{Y}^{v,x}}}(X^w) + \log \frac{\mathrm{d}\mathbb{P}_{X_0^u}}{\mathrm{d}\mathbb{P}_{\overleftarrow{Y}_0^v}}(X_0^w) = \log \frac{\mathrm{d}\mathbb{P}_{X^{u,x}}}{\mathrm{d}\mathbb{P}_{\overleftarrow{Y}^{v,x}}}(X^w) + \log \frac{p_{\mathrm{prior}}(X_0^w)}{p_{\overleftarrow{Y}^v}(X_0^w, 0)}. \tag{18}$$

We follow Liu et al. (2022) and first note that the time-reversal of the process Y^v defined in (4) is given by

$$\mathrm{d}\overleftarrow{Y}_s^v = (\mu + \sigma\sigma^\top \nabla g - \sigma v)(\overleftarrow{Y}_s^v, s)\,\mathrm{d}s + \sigma(s)\,\mathrm{d}W_s,$$

where we abbreviate $g := \log \overleftarrow{p}_{Y^v}$, see Lemma 2.1. Let us further define the short-hand notations $h := u + v - \sigma^\top \nabla g$ and $b := \mu + \sigma(u - h)$. Then, we can write the SDEs in (3) and (4) as

$$\begin{cases} \mathrm{d}X_s^u = (b + \sigma h)(X_s^u, s)\,\mathrm{d}s + \sigma(s)\,\mathrm{d}W_s, \\ \mathrm{d}\overleftarrow{Y}_s^v = b(\overleftarrow{Y}_s^v, s)\,\mathrm{d}s + \sigma(s)\,\mathrm{d}W_s. \end{cases}$$

We can now apply Girsanov's theorem (see, e.g., Nüsken & Richter, 2021, Lemma A.1) to rewrite the logarithm of the Radon-Nikodym derivative in (18) as

$$\begin{aligned} \log \frac{\mathrm{d}\mathbb{P}_{X^{u,x}}}{\mathrm{d}\mathbb{P}_{\overleftarrow{Y}^{v,x}}}(X^w) &= \int_0^T \left(\sigma^{-\top} h\right)(X_s^w, s) \cdot \mathrm{d}X_s^w - \int_0^T \left(\sigma^{-1} b \cdot h\right)(X_s^w, s)\,\mathrm{d}s - \frac{1}{2}\int_0^T \|h(X_s^w, s)\|^2\,\mathrm{d}s \\ &= \int_0^T \left((w - u) \cdot h + \frac{1}{2}\|h\|^2\right)(X_s^w, s)\;\mathrm{d}s + S_h(X^w) \\ &= \int_0^T \left((w - u) \cdot \left(u + v - \sigma^\top \nabla g\right) + \frac{1}{2}\|u + v - \sigma^\top \nabla g\|^2\right)(X_s^w, s)\,\mathrm{d}s + S_h(X^w) \\ &= R_{f^{\mathrm{SB}}_{u,v,w}} - \int_0^T \left(\nabla \cdot (\sigma v - \mu) + (v + w) \cdot \sigma^\top \nabla g - \frac{1}{2}\|\sigma^\top \nabla g\|^2\right)(X_s^w, s)\,\mathrm{d}s + S_h(X^w), \end{aligned} \tag{19}$$

Further, we may apply Itô's lemma to the function g to get

$$g(X_T^w, T) - g(X_0^w, 0) = \int_0^T \left(\partial_s g + \nabla g \cdot (\mu + \sigma w) + \frac{1}{2}\operatorname{Tr}\left(\sigma\sigma^\top \nabla^2 g\right)\right)(X_s^w, s)\,\mathrm{d}s + \int_0^T \sigma^\top \nabla g(X_s^w, s) \cdot \mathrm{d}W_s.$$

Noting that $g = \log p_{\overleftarrow{Y}^v}$ fulfills the Hamilton-Jacobi-Bellman equation (see, e.g., Berner et al., 2022)

$$\partial_s g = -\frac{1}{2}\operatorname{Tr}\left(\sigma\sigma^\top \nabla^2 g\right) + (\sigma v - \mu) \cdot \nabla g + \nabla \cdot (\sigma v - \mu) - \frac{1}{2}\|\sigma^\top g\|^2,$$

we get

$$g(X_T^w, T) - g(X_0^w, 0) = \int_0^T \left(\nabla \cdot (\sigma v - \mu) + (v + w) \cdot \sigma^\top \nabla g - \frac{1}{2}\|\sigma^\top g\|^2\right)(X_s^w, s) + \int_0^T \sigma^\top \nabla g(X_s^w, s) \cdot \mathrm{d}W_s.$$

Finally, combining this with (18) and (19) and noting that

$$g(X_T^w, T) = \log p_{\overleftarrow{Y}^v}(X_T^w, T) = \log p_{Y^v}(X_T^w, 0) = p_{\mathrm{target}}(X_T^w),$$

yields the desired expression. □

Proof of Proposition 2.5. Let us first recall the notion of Gâteaux derivatives, see Siddiqi & Nanda (1986, Section 5.2). We say that $\mathcal{L}\colon \mathcal{U} \times \mathcal{U} \to \mathbb{R}_{\geq 0}$ is *Gâteaux differentiable* at $u \in \mathcal{U}$ if for all $v, \phi \in \mathcal{U}$ the mapping

$$\varepsilon \mapsto \mathcal{L}(u + \varepsilon\phi, v)$$

is differentiable at $\varepsilon = 0$. The Gâteaux derivative of $\mathcal{L}$ w.r.t. u in direction ϕ is then defined as

$$\frac{\delta}{\delta u}\mathcal{L}(u, v; \phi) := \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big|_{\varepsilon=0}\mathcal{L}(u + \varepsilon\phi, v).$$

The derivative of $\mathcal{L}$ w.r.t. v is defined analogously. Let now $u = u_\theta$ and $v = v_\gamma$ be parametrized[2] by $\theta \in \mathbb{R}^p$ and $\gamma \in \mathbb{R}^p$. Relating the Gâteaux derivatives to partial derivatives w.r.t. θ and γ, respectively, let us note that we are particularly interested in the directions $\phi = \partial_{\theta_i} u_\theta$ and $\phi = \partial_{\gamma_i} v_\gamma$ for $i \in \{1, \dots, p\}$. This choice is motivated by the chain rule of the Gâteaux derivative, which, under suitable assumptions, states that

$$\partial_{\theta_i}\mathcal{L}(u_\theta, v_\gamma) = \frac{\delta}{\delta u}\Big|_{u=u_\theta}\mathcal{L}\left(u, v_\gamma; \partial_{\theta_i} u_\theta\right) \quad \text{and} \quad \partial_{\gamma_i}\mathcal{L}(u_\theta, v_\gamma) = \frac{\delta}{\delta v}\Big|_{v=v_\gamma}\mathcal{L}\left(u_\theta, v; \partial_{\gamma_i} v_\gamma\right).$$

Analogous to the computations in Nüsken & Richter (2021), the Gâteaux derivatives of the Monte Carlo estimator $\widehat{\mathcal{L}}^w_{\mathrm{LV}}$ of the log-variance loss $\mathcal{L}^w_{\mathrm{LV}}$ in (9) with $K \in \mathbb{N}$ samples is given by

$$\frac{\delta}{\delta u}\widehat{\mathcal{L}}^w_{\mathrm{LV}}(u, v; \phi) = \frac{2}{K}\sum_{k=1}^{K}\mathcal{A}^{u,v,w,(k)}\left(\left(R_{f^{\mathrm{gen}}_{u,w,\phi}} + S^{(k)}_\phi\right)(X^{w,(k)}) - \frac{1}{K}\sum_{i=1}^{K}\left(R_{f^{\mathrm{gen}}_{u,w,\phi}} + S^{(i)}_\phi\right)(X^{w,(i)})\right), \tag{20}$$

where the superscript (k) denotes the index of the k-th i.i.d. sample in the Monte Carlo estimator $\widehat{\mathcal{L}}^w_{\mathrm{LV}}$ and we define the short-hand notations

$$\mathcal{A}^{u,v,w,(k)} := \left(R_{f^{\mathrm{SB}}_{u,v,w}} + S^{(k)}_{u+v} + B\right)(X^{w,(k)}) + \log Z \quad \text{and} \quad f^{\mathrm{gen}}_{u,w,\phi} = (w - u)\cdot\phi.$$

Now, note that the definition of the log-variance loss and Proposition 2.3 imply that for the optimal choices $u = u^*, v = v^*$ it holds that

$$\mathcal{A}^{u^*,v^*,w,(k)} = 0$$

almost surely for every $k \in \{1, \dots, K\}$ and $w \in \mathcal{U}$. This readily implies the statement for the derivative w.r.t. the control u_γ. The analogous statement holds true for the derivative w.r.t. v_γ, as we can compute

$$\frac{\delta}{\delta v}\widehat{\mathcal{L}}^w_{\mathrm{LV}}(u, v; \phi) = \frac{2}{K}\sum_{k=1}^{K}\mathcal{A}^{u,v,w,(k)}\left(\left(R_{f^{\mathrm{inf}}_{v,w,\phi}} + S^{(k)}_\phi\right)(X^{w,(k)}) - \frac{1}{K}\sum_{i=1}^{K}\left(R_{f^{\mathrm{inf}}_{v,w,\phi}} + S^{(i)}_\phi\right)(X^{w,(i)})\right),$$

where

$$f^{\mathrm{inf}}_{v,w,\phi} = (v + w)\cdot\phi + \nabla\cdot(\sigma\phi).$$

For the derivative of the Monte Carlo version of the loss $\mathcal{L}_{\mathrm{KL}}$ as defined in (8) w.r.t. to v we may compute

$$\frac{\delta}{\delta v}\widehat{\mathcal{L}}_{\mathrm{KL}}(u, v; \phi) = \frac{1}{K}\sum_{k=1}^{K}\int_0^T \left((u + v)\cdot\phi + \nabla\cdot(\sigma\phi)\right)(X^{u,(k)}_s, s)\,\mathrm{d}s.$$

We note that even for $u = u^*$ and $v = v^*$ we can usually not expect the variance of the corresponding Monte Carlo estimator to be zero. For the computation of the derivative w.r.t. u we refer to Nüsken & Richter (2021, Proposition 5.3). □

Remark A.1 (Control variate interpretation)**.** For the gradient of the loss $\mathcal{L}_{\mathrm{KL}}$ w.r.t. to u we may compute

$$\frac{\delta}{\delta u}\mathcal{L}_{\mathrm{KL}}(u, v; \phi) = \mathbb{E}\left[\int_0^T \left((u + v)\cdot\phi\right)(X^u_s, s)\,\mathrm{d}s + \left(R_{f^{\mathrm{SB}}_{u,v,u}}(X^u) + B(X^u)\right)S_\phi(X^u)\right] = \mathbb{E}\left[\mathcal{A}^{u,v,u}S_\phi(X^u)\right],$$

where we used Girsanov's theorem and the Itô isometry. Comparing with (20), we realize that the derivative of $\mathcal{L}_{\mathrm{LV}}$ w.r.t. u for the choice $w = u$ can be interpreted as a control variate version of the derivative of $\mathcal{L}_{\mathrm{KL}}$, thereby promising reduced variance of the corresponding Monte Carlo estimators, cf. Nüsken & Richter (2021); Richter et al. (2020).

[2] We only assume that θ and γ are in the same space $\mathbb{R}^p$ for notational simplicity.

A.2. The Schrödinger bridge problem

In the following, we will formulate optimality conditions for the Schrödinger bridge problem defined in (14) for the standard case $r = 0$. Moreover, we outline how the associated system of BSDE system leads to the same losses as given in (8) and (9), respectively. The ideas are based on Chen et al. (2021a); Vargas (2021); Liu et al. (2022); Caluya & Halder (2021).

First, we can define the

$$\phi(x,t) \coloneqq \min_{u\in\mathcal{U}} \mathbb{E}\left[\int_t^T \frac{1}{2}\|u(X_s^u,s)\|^2 \,\mathrm{d}s \middle| X_t^u = x,\ X_T^u \sim p_{\text{target}}\right].$$

By the *dynamic programming principle* it holds that ϕ solves the *Hamilton-Jacobi-Bellman* (HJB) equation

$$\partial_t\phi = -\mu\cdot\nabla\phi - \frac{1}{2}\operatorname{Tr}\left(\sigma\sigma^\top\nabla^2\phi\right) + \frac{1}{2}\left\|\sigma^\top\nabla\phi\right\|^2 \tag{21}$$

(with unknown boundary conditions) and that the optimal control satisfies

$$u^* = -\sigma^\top\nabla\phi.$$

Together with the corresponding Fokker-Planck equation for X^{u^*}, this yields necessary and sufficient conditions for the solution to (13). Now, we can transform the Fokker-Planck equation and the HJB equation (21) into a system of linear equations, using the exponential transform

$$\psi \coloneqq \exp(-\phi) \quad \text{and} \quad \widehat{\psi} \coloneqq p_{X^{u^*}}\exp(\phi) = \frac{p_{X^{u^*}}}{\psi}, \tag{22}$$

often referred to as the *Hopf-Cole transform*. This yields the following well-known optimality conditions of the Schrödinger Bridge problem defined in (14).

Theorem A.2 (Optimality PDEs)**.** *The solution u^* to the Schrödinger Bridge problem* (14) *is equivalently given by*

1. *$u^* \coloneqq -\sigma^\top\nabla\phi$, where $p_{X^{u^*}}$ and ϕ are the unique solutions to the coupled PDEs*

$$\begin{cases}\partial_t p_{X^{u^*}} = -\nabla\cdot\left(p_{X^{u^*}}(\mu - \sigma\sigma^\top\nabla\phi)\right) + \frac{1}{2}\operatorname{Tr}\left(\sigma\sigma^\top\nabla^2 p_{X^{u^*}}\right)\\ \partial_t\phi = -\mu\cdot\nabla\phi - \frac{1}{2}\operatorname{Tr}\left(\sigma\sigma^\top\nabla^2\phi\right) + \frac{1}{2}\left\|\sigma^\top\nabla\phi\right\|^2,\end{cases}$$

   with boundary conditions

$$\begin{cases}p_{X^{u^*}}(\cdot,0) = p_{\text{prior}},\\ p_{X^{u^*}}(\cdot,T) = p_{\text{target}}.\end{cases}$$

2. *$u^* \coloneqq \sigma^\top\nabla\log\psi$, where ψ and $\widehat{\psi}$ are the the unique solutions to the PDEs*

$$\begin{cases}\partial_t\psi = -\nabla\psi\cdot\mu - \frac{1}{2}\operatorname{Tr}\left(\sigma\sigma^\top\nabla^2\psi\right),\\ \partial_t\widehat{\psi} = -\nabla\cdot\left(\widehat{\psi}\mu\right) + \frac{1}{2}\operatorname{Tr}\left(\sigma\sigma^\top\nabla^2\widehat{\psi}\right),\end{cases} \tag{23}$$

   with coupled boundary conditions

$$\begin{cases}\psi(\cdot,0)\widehat{\psi}(\cdot,0) = p_{\text{prior}},\\ \psi(\cdot,T)\widehat{\psi}(\cdot,T) = p_{\text{target}}.\end{cases} \tag{24}$$

The optimal control v^* is given by Nelson's identity (10)*, i.e.,*

$$v^* = \sigma^\top\nabla\log p_{X^{u^*}} - u^* = \sigma^\top\nabla\log\widehat{\psi}. \tag{25}$$

Using Itô's lemma, we now derive a BSDE system corresponding to the PDE system in (23).

Proposition A.3 (BSDEs for the SB problem)**.** *Let us assume ψ and $\widehat{\psi}$ fulfill the PDEs (23) with boundary conditions (24) and let us define the processes*

$$\begin{cases} \mathcal{Y}_s^w = \log \psi(X_s^w, s), \\ \widehat{\mathcal{Y}}_s^w = \log \widehat{\psi}(X_s^w, s), \\ \mathcal{Z}_s^w = \sigma^\top \nabla \log \psi(X_s^w, s) = u^*(X_s^w, s), \\ \widehat{\mathcal{Z}}_s^w = \sigma^\top \nabla \log \widehat{\psi}(X_s^w, s) = v^*(X_s^w, s), \end{cases}$$

where the process X^w is given by

$$\mathrm{d}X_s^w = (\mu + \sigma w)(X_s^w, s)\,\mathrm{d}s + \sigma(s)\,\mathrm{d}W_s$$

with $w \in \mathcal{U}$ being an arbitrary control function. We then get the BSDE system

$$\begin{cases} \mathrm{d}\mathcal{Y}_s^w = \left(\mathcal{Z}_s^w \cdot w(X_s^w, s) - \frac{1}{2}\|\mathcal{Z}_s^w\|^2\right)\,\mathrm{d}s + \mathcal{Z}_s^w \cdot \mathrm{d}W_s, \\ \mathrm{d}\widehat{\mathcal{Y}}_s^w = \left(\frac{1}{2}\|\widehat{\mathcal{Z}}_s^w\|^2 + \nabla \cdot (\sigma \widehat{\mathcal{Z}}_s^w - \mu(X_s^w, s)) + \widehat{\mathcal{Z}}_s^w \cdot w(X_s^w, s)\right)\,\mathrm{d}s + \widehat{\mathcal{Z}}_s^w \cdot \mathrm{d}W_s. \end{cases}$$

Furthermore, it holds

$$\mathcal{Y}_s^w + \widehat{\mathcal{Y}}_s^w = \log p_{X^{u^*}}(X_s^w, s) = \log \bar{p}_{Y^{v^*}}(X_s^w, s). \tag{26}$$

Proof. The proof is similar to the one in Chen et al. (2021a). For brevity, we define $D = \frac{1}{2}\sigma\sigma^\top$. We can apply Itô's lemma to the stochastic process $\mathcal{Y}_s^w = \log \psi(X_s^w, s)$ and get

$$\mathrm{d}\mathcal{Y}_s^w = \left(\partial_s \log \psi + \nabla \log \psi \cdot (\mu + \sigma w) + \mathrm{Tr}\left(D\nabla^2 \log \psi\right)\right)(X_s^w, s)\,\mathrm{d}s + \sigma^\top \nabla \log \psi(X_s^w, s) \cdot \mathrm{d}W_s. \tag{27}$$

Further, via (23) it holds

$$\partial_s \log \psi = \frac{1}{\psi}\left(-\nabla\psi \cdot \mu - \mathrm{Tr}\left(D\nabla^2\psi\right)\right) = -\nabla \log \psi \cdot \mu - \mathrm{Tr}\left(\frac{D\nabla^2\psi}{\psi}\right), \tag{28}$$

and we note the identity

$$\nabla^2 \log \psi = \frac{\nabla^2\psi}{\psi} - \frac{\nabla\psi\,(\nabla\psi)^\top}{\psi^2}. \tag{29}$$

Combining (27), (28), and (29), we get

$$\begin{aligned} \mathrm{d}\mathcal{Y}_s^w &= \left(\sigma^\top \nabla \log \psi \cdot w - \mathrm{Tr}\left(D\frac{\nabla\psi\,(\nabla\psi)^\top}{\psi^2}\right)\right)(X_s^w, s)\,\mathrm{d}s + \sigma^\top \nabla \log \psi(X_s^w, s) \cdot \mathrm{d}W_s \\ &= \left(\mathcal{Z}_s^w \cdot w(X_s^w, s) - \frac{1}{2}\|\mathcal{Z}_s^w\|^2\right)\,\mathrm{d}s + \mathcal{Z}_s^w \cdot \mathrm{d}W_s. \end{aligned}$$

Similarly, we may apply Itô's lemma to $\widehat{\mathcal{Y}}_s^w = \log \widehat{\psi}(X_s^w, s)$ and get

$$\mathrm{d}\widehat{\mathcal{Y}}_s^w = \left(\partial_s \log \widehat{\psi} + \nabla \log \widehat{\psi} \cdot (\mu + \sigma w) + \mathrm{Tr}\left(D\nabla^2 \log \widehat{\psi}\right)\right)(X_s^w, s)\,\mathrm{d}s + \sigma^\top \nabla \log \widehat{\psi}(X_s^w, s) \cdot \mathrm{d}W_s. \tag{30}$$

Now, via (23) it holds that

$$\partial_s \log \widehat{\psi} = \frac{1}{\widehat{\psi}}\left(-\nabla \cdot \left(\widehat{\psi}\mu\right) + \mathrm{Tr}\left(D\nabla^2\widehat{\psi}\right)\right) = -\nabla \log \widehat{\psi} \cdot \mu - \nabla \cdot \mu + \mathrm{Tr}\left(\frac{D\nabla^2\widehat{\psi}}{\widehat{\psi}}\right). \tag{31}$$

Combining (30) and (31), we get

$$\mathrm{d}\log \widehat{\psi}(X_s^w, s) = \left(\mathrm{Tr}\left(D\frac{\nabla^2\widehat{\psi}}{\widehat{\psi}} + D\nabla^2 \log \widehat{\psi}\right) - \nabla \cdot \mu + \sigma^\top \nabla \log \widehat{\psi} \cdot w\right)(X_s^w, s)\,\mathrm{d}s + \sigma^\top \nabla \log \widehat{\psi}(X_s^w, s) \cdot \mathrm{d}W_s.$$

Now, noting the identity

$$\mathrm{Tr}\left(D\frac{\nabla^2\widehat{\psi}}{\widehat{\psi}} + D\nabla^2 \log \widehat{\psi}\right) = 2\,\mathrm{Tr}\left(D\frac{\nabla^2\widehat{\psi}}{\widehat{\psi}}\right) - \frac{1}{2}\|\sigma^\top \nabla \log \widehat{\psi}\|^2 = \frac{1}{2}\|\sigma^\top \nabla \log \widehat{\psi}\|^2 + \nabla \cdot \left(\sigma\sigma^\top \nabla \log \widehat{\psi}\right),$$

we can get the relation

$$\begin{aligned}\mathrm{d}\widehat{\mathcal{Y}}_s^w &= \Big(\frac{1}{2}\|\sigma^\top \nabla \log \widehat{\psi}\|^2 + \nabla \cdot \big(\sigma\sigma^\top \nabla \log \widehat{\psi} - \mu\big) + \sigma^\top \nabla \log \widehat{\psi} \cdot w\Big)(X_s^w, s)\,\mathrm{d}s + \sigma^\top \nabla \log \widehat{\psi}(X_s^w, s) \cdot \mathrm{d}W_s \\ &= \Big(\frac{1}{2}\|\widehat{\mathcal{Z}}_s^w\|^2 + \nabla \cdot (\sigma \widehat{\mathcal{Z}}_s^w - \mu) + \widehat{\mathcal{Z}}_s^w \cdot w\Big)(X_s^w, s)\,\mathrm{d}s + \widehat{\mathcal{Z}}^w \cdot \mathrm{d}W_s,\end{aligned}$$

which concludes the proof. □

Note that the BSDE system is slightly more general than the one introduced in Chen et al. (2021a), which can be recovered with the choice $w(X_s^w, s) = \mathcal{Z}_s^w$. Also, the roles of p_{prior} and p_{target} are interchanged in Chen et al. (2021a) since they consider generative modeling instead of sampling from densities.

A valid loss can now be derived by adding the two BSDEs and recalling relation (26), which yields

$$\log \frac{p_{\text{target}}(X_T^w)}{p_{\text{prior}}(X_0^w)} = \int_0^T \Big((\mathcal{Z}_s^w + \widehat{\mathcal{Z}}_s^w) \cdot \Big(w + \frac{\widehat{\mathcal{Z}}_s^w - \mathcal{Z}_s^w}{2}\Big) + \nabla \cdot \big(\sigma \widehat{\mathcal{Z}}_s^w - \mu\big)\Big)(X_s^w, s)\,\mathrm{d}s + \int_0^T \Big(\mathcal{Z}_s^w + \widehat{\mathcal{Z}}_s^w\Big) \cdot \mathrm{d}W_s$$

almost surely. Analogous to Berner et al. (2022); Huang et al. (2021) in generative modeling, the above equality suggests a parameterized lower bound of the log-likelihood $\log p_{\text{prior}}$ when replacing the optimal controls in $\mathcal{Z}_s^w = u^*(X^w, s)$ and $\widehat{\mathcal{Z}}_s^w = v^*(X_s^w, s)$ with their approximations u and v, see Chen et al. (2021a). This lower bound exactly recovers the loss given in (8). Further, note that variance of the left-hand minus the right-hand side is zero, which readily yields our log-variance loss as defined in (9).

A.2.1. Schrödinger half-bridges (PIS)

For the Schrödinger half-bridge, also referred to as PIS, introduced in Section 3.3, we can find an alternative derivation, motivated by the PDE perspective outlined in Appendix A.2. For this derivation it is crucial that we assume the prior density to be concentrated at a single point, i.e., $p_{\text{prior}} := \delta_{x_0}$ for some $x_0 \in \mathbb{R}^d$ (typically $x_0 = 0$), see Tzen & Raginsky (2019); Dai Pra (1991). We can recover the corresponding objectives by noting that, in the case $p_{\text{prior}} = \delta_{x_0}$, the system of PDEs in (23) can be decoupled. More precisely, we observe that the second equation in (23) is the Fokker-Planck equation of X^0 and we have that

$$\widehat{\psi} = p_{X^{u^*}} \exp(\phi) = p_{X^0} \quad \text{and} \quad \widehat{\psi}(\cdot, 0) = p_{X_0^0} = \delta_{x_0}.$$

In view of (25), we note that this defines $v^* = \sigma^\top \nabla \log p_{X^0}$. By (22), we observe that $\psi = \frac{p_{X^{u^*}}}{p_{X^0}}$, which yields the boundary condition

$$\phi(\cdot, T) = -\log \psi(\cdot, T) = \log \frac{p_{X_T^0}}{p_{\text{target}}} = \log \frac{Z p_{X_T^0}}{\rho}$$

to the HJB equation in (21). By the *verification theorem* (Dai Pra, 1991; Pavon, 1989; Nüsken & Richter, 2021; Fleming & Soner, 2006; Pham, 2009), we thus obtain the PIS objective

$$\mathcal{L}_{\text{KL}}(u) = \mathbb{E}\left[\int_0^T \frac{1}{2}\|u(X_s^u, s)\|^2\,\mathrm{d}s + \log \frac{p_{X_T^0}(X_T^u)}{\rho(X_T^u)}\right] = \mathbb{E}\left[\big(R_{f_{u,0,u}^{\text{ref}}} + B^{\text{ref}}\big)(X^u)\right].$$

Moreover, the optimal control is given by $u^* = -\sigma^\top \nabla \phi = \sigma^\top \nabla \log \psi$. We can also derive this objective from the BSDE system in Proposition A.3. Since $\widehat{\psi}(\cdot, 0) = \delta_{x_0}$, we may focus on the process $\mathcal{Y}_s^w = \log \psi(X_s^w, s)$ only and get

$$\mathcal{Y}_T^w - \mathcal{Y}_0^w = \int_0^T \mathcal{Z}_s^w \cdot w(X_s^w, s) - \frac{1}{2}\|\mathcal{Z}_s^w\|^2\,\mathrm{d}s + \int_0^T \mathcal{Z}_s^w \cdot \mathrm{d}W_s.$$

The PIS objective now follows by choosing $w(X_s^w, s) = \mathcal{Z}_s^w$ and noting that

$$\mathcal{Y}_T^w = \log \psi(X_T^w, T) = \log \frac{p_{\text{target}}}{p_{X_T^0}}(X_T^w).$$

Recalling our notation in (1) and (2), this also shows that the log-variance loss can be written as

$$\mathcal{L}_{\text{LV}}^w(u) = \mathbb{V}\left[\big(R_{f_{u,0,w}^{\text{ref}}} + S_u + B^{\text{ref}}\big)(X^w)\right].$$

A.3. Tractable SDEs

Let us present some commonly used SDEs of the form

$$\mathrm{d}X_s^u = \mu(X_s^u, s)\,\mathrm{d}s + \sigma(s)\,\mathrm{d}W_s$$

with affine drifts that have tractable marginals conditioned on their initial value, see Song et al. (2020). For notational convenience, let us define

$$\alpha(t) \coloneqq \int_0^t \beta(s)\mathrm{d}s$$

with suitable $\beta \in C([0,T],(0,\infty))$.

Variance-Preserving (VP) SDE: This *Ornstein-Uhlenbeck* process is given by

$$\sigma(t) \coloneqq \nu\sqrt{2\beta(t)}\,\mathrm{I} \quad \text{and} \quad \mu(x,t) \coloneqq -\beta(t)x.$$

with $\nu \in (0,\infty)$. Then, we have that

$$X_t|X_0 \sim \mathcal{N}\left(e^{-\alpha(t)}X_0, \nu^2\left(1-e^{-2\alpha(t)}\right)\mathrm{I}\right).$$

This shows that for $\alpha(T)$ sufficiently large it holds that $X_T \approx \mathcal{N}\left(0,\nu^2\mathrm{I}\right)$. For $X_0 \sim \mathcal{N}(m,\Sigma)$, we further have that

$$X_t \sim \mathcal{N}\left(e^{-\alpha(t)}m, e^{-2\alpha(t)}\left(\Sigma-\nu^2\mathrm{I}\right)+\nu^2\mathrm{I}\right). \tag{33}$$

Variance-exploding (VE) SDE / scaled Brownian motion: This SDE is given by a scaled Brownian motion, i.e., $\mu \coloneqq 0$ and σ as defined above. It holds that

$$X_t|X_0 \sim \mathcal{N}\left(X_0, 2\nu^2\alpha(t)\mathrm{I}\right).$$

For $X_0 \sim \mathcal{N}(m,\Sigma)$, we thus have that

$$X_t \sim \mathcal{N}\left(m, 2\nu^2\alpha(t)\mathrm{I}+\Sigma\right).$$

A.4. Computational details

In our implementations, we generally follow the settings and hyperparameters of PIS in Zhang & Chen (2022). The main difference is that we observed better performance (for all considered methods and losses) by choosing more steps for the SDE solver, larger batch sizes, and more gradient steps during training. We thus always used 200 steps for the Euler-Maruyama scheme, a batch size of 2048, and 60000 gradient steps for the experiments with $d \leq 10$ and 120000 gradient steps otherwise. However, we observed that the differences between the losses are already visible before convergence, see, e.g., Figure 1.

For DIS, we replace the pinned Brownian motion of Zhang & Chen (2022) by the VP-SDE in Song et al. (2020). Specifically, we use $\nu \coloneqq 1$ and

$$\beta(t) \coloneqq (1-t)\,\beta_{\min} + t\beta_{\max}, \quad t \in [0,1],$$

with $\beta_{\min} = 0.05$ and $\beta_{\max} = 5$, see Appendix A.3. Similar to PIS, we also use the score of the density $\nabla \log \rho$ (typically given in closed-form or evaluated via automatic differentiation) for the parametrization of the control u.

For the SB examples reported in Table 2, we chose the same setting as in the PIS and DIS experiments. One difference is, however, that we are free to choose the prior density p_{prior} as well as the drift function μ in the SDEs (3) and (4). We choose $\mu = 0$, noting that more sophisticated, potentially problem-specific choices might be investigated in future studies. We observed that for the double well example with $d = 5$ it was sufficient to choose $p_{\text{prior}} = \mathcal{N}(0,\mathrm{I})$. For the GMM and the high-dimensional double well example, on the other hand, the experiments did not properly converge using the Gaussian prior. In case of the GMM example, choosing the uniform density $p_{\text{prior}} = \frac{1}{256}\mathbb{1}_{[-8,8]^2}$, helped the model to converge while detecting all nine modes.

For the log-variance loss, we used the default choice of $w \coloneqq u$, i.e., $X^w \coloneqq X^u$. We emphasize that we do not need to differentiate w.r.t. w, which results in reduced training times, see Figure 4. In practice, we detach X^w from the computational graph, which can be achieved by the `detach` and `stop_gradient` operations in PyTorch and TensorFlow, respectively.

We leave other choices of w to future research and anticipate that choosing noisy versions of u in the initial phase of training might lead to even better exploration and performance. Furthermore, we use the same hyperparameters for the log-variance loss as for the KL-based loss. As these settings originate form Zhang & Chen (2022) and have been tuned for the KL-based loss, we suspect that optimizing the hyperparameters for the log-variance loss can lead to further improvements.

To evaluate our metrics, we consider $n = 10^5$ samples $(x^{(i)})_{i=1}^n$ and use the ELBO as an approximation to the log-normalizing constant $\log Z$, see Appendix A.4.1. We further compute the (normalized) *effective sample size*

$$\mathrm{ESS} := \frac{\left(\sum_{i=1}^n w^{(i)}\right)^2}{n \sum_{i=1}^n \left(w^{(i)}\right)^2},$$

where $(w^{(i)})_{i=1}^n$ are the importance weights of the samples $(x^{(i)})_{i=1}^n$ in path space. Finally, we estimate the Sinkhorn distance[3] $\mathcal{W}_\gamma^2$ (Cuturi, 2013) and report the error for estimating the average standard deviation across the marginals, i.e.,

$$\mathrm{std} := \frac{1}{d} \sum_{k=1}^d \sqrt{\mathbb{V}[G_k]}, \quad \text{where} \quad G \sim p_{\text{target}}.$$

A.4.1. Computation of log-normalizing constant

For the computation of the log-normalizing constant in the general SB setting, Proposition 2.3 ensures that for the optimal u^* and v^* it holds that

$$\log Z = -\big(R_{f^{\mathrm{SB}}_{u^*,v^*,u^*}} + S_{u^*+v^*} + B\big)(X^{u^*})$$

Using approximations of u^* and v^*, the ELBO yields a lower bound to $\log Z$. For PIS and DIS, the log-normalizing constants can be computed analogously, see also Zhang & Chen (2022); Berner et al. (2022).

A.5. Further experiments

In Figure 4 we present boxplots to show that our results from Table 3 are robust w.r.t. different seeds.

[3] Our implementation is based on https://github.com/fwilliams/scalable-pytorch-sinkhorn with the default parameters.

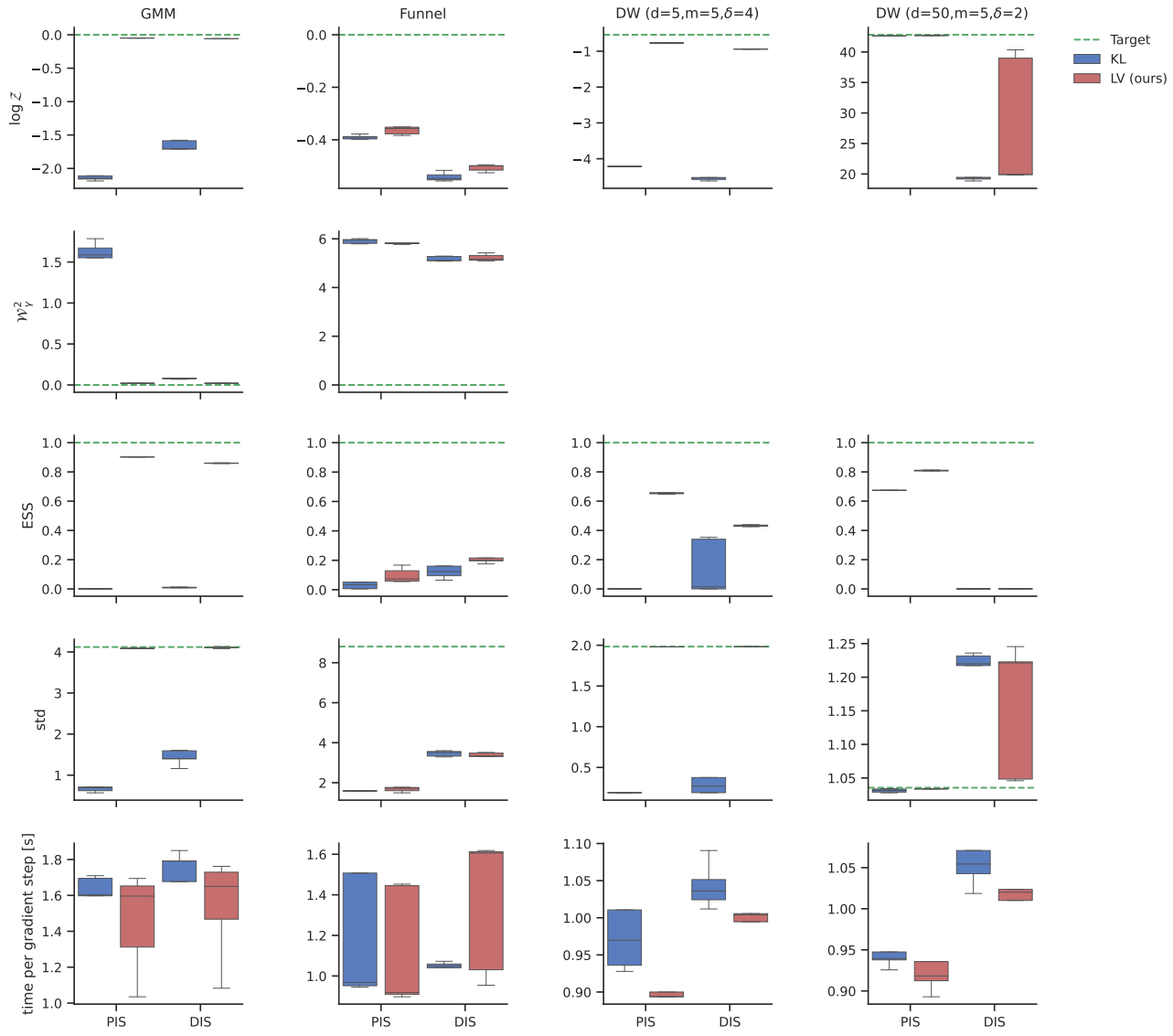

Figure 4. Boxplots for five independent runs for each problem and method (KL-PIS, LV-PIS (ours), KL-DIS, LV-DIS (ours) from left to right in each plot) in the settings of Table 3 and corresponding ground truth or optimal values (dashed lines). It can be seen that the performance improvements of the log-variance loss are robust across different seeds. At the same time, the log-variance loss reduces the average time per gradient step by circumventing differentiation through the SDE solver.