Transport, Variational Inference and Diffusions: with Applications to Annealed Flows and Schrödinger Bridges

Francisco Vargas [1] **Nikolas Nüsken** [2]

Abstract

This paper explores the connections between optimal transport and variational inference, with a focus on forward and reverse time stochastic differential equations and Girsanov transformations.We present a principled and systematic framework for sampling and generative modelling centred around divergences on path space. Our work culminates in the development of a novel score-based annealed flow technique (with connections to Jarzynski and Crooks identities from statistical physics) and a regularised iterative proportional fitting (IPF)-type objective, departing from the sequential nature of standard IPF. Through a series of generative modelling examples and a double-well-based rare event task, we showcase the potential of the proposed methods.

1. Introduction

Optimal transport (Villani et al., 2009) and variational inference (Blei et al., 2017) have for a long time been separate fields of research. In recent years, many fruitful connections have been established (Liu et al., 2019), in particular based on dynamical formulations (Tzen and Raginsky, 2019), and in conjunction with time reversals (Huang et al., 2021a; Song et al., 2021). In this paper, we enhance those relationships based on forward and reverse time stochastic differential equations, and associated Girsanov transformations.

To set the stage, we recall a classical approach (Kingma and Welling, 2014; Rezende and Mohamed, 2015) towards generating samples from a target distribution $\mu(\boldsymbol{x})$, which is the goal both in generative modelling and sampling:

*Equal contribution [1]Department of Computer Science, Cambridge University, Cambridge, CB3 0FD, UK [2]Department of Mathematics, King's College London, London, WC2R 2LS. Correspondence to: Francisco Vargas <fav25@cam.ac.uk>, Nikolas Nüsken <nikolas.nusken@kcl.ac.uk>.

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

Generative processes, encoders and decoders. We consider methodologies which can be implemented via the following generative process,

$$\boldsymbol{z} \sim \nu(\boldsymbol{z}), \quad \boldsymbol{x}|\boldsymbol{z} \sim p^\theta(\boldsymbol{x}|\boldsymbol{z}), \tag{1}$$

transforming a sample $\boldsymbol{z} \sim \nu(\boldsymbol{z})$ into a sample $\boldsymbol{x} \sim \int p^\theta(\boldsymbol{x}|\boldsymbol{z})\nu(\mathrm{d}\boldsymbol{z})$. Traditionally, $\nu(\boldsymbol{z})$ is a simple auxiliary distribution, and the family of transitions $p^\theta(\boldsymbol{x}|\boldsymbol{z})$ is parameterised flexibly and in such a way that sampling according to (1) is tractable. Then we can frame the tasks of generative modelling and sampling as finding transition densities such that the marginal in $\boldsymbol{x}$ matches the target distribution,

$$\mu(\boldsymbol{x}) = \int p^\theta(\boldsymbol{x}|\boldsymbol{z})\nu(\mathrm{d}\boldsymbol{z}). \tag{2}$$

To learn such a transition, it is helpful to introduce a reversed process

$$\boldsymbol{x} \sim \mu(\boldsymbol{x}), \quad \boldsymbol{z}|\boldsymbol{x} \sim q^\phi(\boldsymbol{z}|\boldsymbol{x}), \tag{3}$$

relying on an appropriately parameterised backward transition $q^\phi(\boldsymbol{z}|\boldsymbol{x})$. We will say that (1) and (3) are *reversals* of each other in the case when the joint distributions associated to both coincide, that is, when

$$q^\phi(\boldsymbol{z}|\boldsymbol{x})\mu(\boldsymbol{x}) = p^\theta(\boldsymbol{x}|\boldsymbol{z})\nu(\boldsymbol{z}). \tag{4}$$

To appreciate the significance of the reversed process in (3), notice that if the reversal relation in (4) holds, then (2) is implied by integrating both sides with respect to $\boldsymbol{z}$. Building on this observation, it is natural to define the loss function

$$\mathcal{L}_D(\phi, \theta) := D\left(q^\phi(\boldsymbol{z}|\boldsymbol{x})\mu(\boldsymbol{x})\middle|\middle|p^\theta(\boldsymbol{x}|\boldsymbol{z})\nu(\boldsymbol{z})\right), \tag{5}$$

where D is a divergence[1] between distributions yet to be specified. Along the lines of Bengio et al. (2021); Sohl-Dickstein et al. (2015); Wu et al. (2020); Liu et al. (a), we summarise the discussion so far as follows, laying the foundation for algorithmic approaches that aim at sampling from $\mu(\boldsymbol{x})$ by minimising $\mathcal{L}_D(\phi, \theta)$:

Framework 1. Let D be an arbitrary divergence, and assume that $\mathcal{L}_D(\phi, \theta) = 0$. Then we have

$$\mu(\boldsymbol{x}) = \int p^\theta(\boldsymbol{x}|\boldsymbol{z})\nu(\mathrm{d}\boldsymbol{z}) \text{ and } \nu(\boldsymbol{z}) = \int q^\phi(\boldsymbol{z}|\boldsymbol{x})\mu(\mathrm{d}\boldsymbol{x}), \tag{6}$$

[1]As usual, divergences are characterised by the requirement that $D(\alpha||\beta) \geq 0$, with equality if and only if $\alpha = \beta$.

that is, $\nu(\boldsymbol{z})$ is transformed into $\mu(\boldsymbol{x})$ by $p^\theta(\boldsymbol{x}|\boldsymbol{z})$, and $\mu(\boldsymbol{x})$ is transformed into $\nu(\boldsymbol{z})$ by $q^\phi(\boldsymbol{z}|\boldsymbol{x})$.

KL-divergence, ELBO and variational inference. Choosing $D = D_{\mathrm{KL}}$ in (5), variational inference (VI) and latent variable model (LVM) based approaches (Dempster et al., 1977; Blei et al., 2017; Kingma and Welling, 2014) can elegantly be placed within Framework 1. Indeed, direct computation shows that

$$\mathcal{L}_{D_{\mathrm{KL}}}(\phi,\theta) = -\mathbb{E}_{\boldsymbol{x}\sim\mu(\boldsymbol{x})}\overbrace{\left[\int \ln \frac{p^\theta(\boldsymbol{x}|\boldsymbol{z})\nu(\boldsymbol{z})}{q^\phi(\boldsymbol{z}|\boldsymbol{x})} q^\phi(\mathrm{d}\boldsymbol{z}|\boldsymbol{x})\right]}^{=\mathrm{ELBO}_x(\phi,\theta)} + \int \ln \mu(\boldsymbol{x})\mu(\mathrm{d}\boldsymbol{x}), \tag{7}$$

so that minimising $\mathcal{L}_{D_{\mathrm{KL}}}(\phi,\theta)$ is equivalent to maximising the expected evidence lower bound (ELBO), also known as the negative free energy (Blei et al., 2017). This derivation is alternative to the standard approach via maximum likelihood and convex duality (or Jensen's inequality) (Kingma et al., 2021, Section 2.2), and straightforwardly accomodates various modifications by replacing the D_{KL}-divergence (see Appendix B). Other popular choices for D include the f and α divergences (Csiszár, 1967; Ali and Silvey, 1966; Rényi, 1961).

Couplings, (optimal) transport and nonuniqueness. Assuming that (4) holds, it is natural to define the joint distribution $\pi(\boldsymbol{x},\boldsymbol{z}) := q^\phi(\boldsymbol{z}|\boldsymbol{x})\mu(\boldsymbol{x}) = p^\theta(\boldsymbol{x}|\boldsymbol{z})\nu(\boldsymbol{z})$, which is then a coupling between $\mu(\boldsymbol{x})$ and $\nu(\boldsymbol{z})$ by virtue of (6). Viewed from this angle, it is clear that the set of minimisers associated to $\mathcal{L}(\phi,\theta)$ stands in one-to-one correspondence with the set of couplings between $\mu(\boldsymbol{x})$ and $\nu(\boldsymbol{z})$, provided that the parameterisations are chosen flexibly enough. Under the latter assumption, the objective in (5) hence generically admits an infinite number of minimisers, rendering algorithmic approaches solely based on Framework 1 potentially unstable and their output hard to interpret. In the language of *optimal transport* (Villani, 2003), minimising $\mathcal{L}(\phi,\theta)$ enforces the marginal (*'transport'*) constraints in (6) without a selection principle based on an appropriate cost function (*'optimal'*).

In practice, methodologies such as VAEs (Kingma and Welling, 2014) parameterise the transition densities $p^\theta(\boldsymbol{x}|\boldsymbol{z})$ and $q^\phi(\boldsymbol{z}|\boldsymbol{x})$ with a restricted family of distributions (such as Gaussians), thus restricting the set of couplings. Expectation maximisation (EM) carries out the minimisation of $\mathcal{L}(\phi,\theta)$ in a component-wise fashion, resolving nonquniqueness in a procedural manner (see Section 4.1 for a connection to Schrödinger bridges). In this paper, we develop principled approaches towards resolving the nonuniqueness inherent in Framework 1, relying on a hierarchical extension of $p^\theta(\boldsymbol{x}|\boldsymbol{z})$ and $q^\phi(\boldsymbol{z}|\boldsymbol{x})$ and proceeding to an infinite-depth limit. In particular, we proceed as follows:

Outline and contributions. In Section 2 we recall hierarchical VAEs (Rezende et al., 2014) and, following Tzen and Raginsky (2019), proceed to the infinite-depth limit described by the SDEs in (13). Readers more familiar with VI and discrete time might want to take the development in Section 2.1 as an explanation of (13); readers with background in stochastic analysis might take Framework 1′ as their starting point. In Proposition 2.2 we provide a generalised form of the Girsanov theorem for forward-reverse time SDEs, crucially incorporating the choice of a reference process that allows us to reason about sampling and generation in a systematic and principled way. In Section 3, we utilise the flexibility in this formulation to recover recent approaches from the literature, allowing for slight variations and generalisations, and develop a novel score-based annealed flow technique. Section 4 reinforces the connection between optimal transport and VI by establishing a correspondence between expectation-maximisation (EM) and iterative proportional fitting (IPF). Motivated by this result and building on the set-up of Section 2, we suggest a regularised IPF-type objective that allows us to circumvent the sequential nature of standard IPF. Finally, we explore our proposed regularised IPF objective across a series of toy generative modelling examples and a double-well-based rare event (Hartmann et al., 2013) task.

2. From hierarchical VAEs to forward-reverse SDEs

2.1. Hierarchical VAEs (Rezende et al., 2014)

A particularly flexible choice of implicitly parameterising $p^\theta(\boldsymbol{x}|\boldsymbol{z})$ and $q^\phi(\boldsymbol{z}|\boldsymbol{x})$ can be achieved via a hierarchical model with intermediate latents: We identify $\boldsymbol{x} =: \boldsymbol{y}_0$ and $\boldsymbol{z} =: \boldsymbol{y}_L$ with the 'endpoints' of the layered augmentation $(\boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{L-1}, \boldsymbol{y}_L) =: \boldsymbol{y}_{0:L}$, and define

$$\begin{aligned} q^\phi(\boldsymbol{y}_L, \boldsymbol{y}_{L-1}, \ldots, \boldsymbol{y}_1|\boldsymbol{y}_0) &:= \prod_{l=1}^{L} q^{\phi_{l-1}}(\boldsymbol{y}_l|\boldsymbol{y}_{l-1}), \\ p^\theta(\boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{L-1}|\boldsymbol{y}_L) &:= \prod_{l=1}^{L} p^{\theta_l}(\boldsymbol{y}_{l-1}|\boldsymbol{y}_l), \end{aligned} \tag{8}$$

so that $q^\phi(\boldsymbol{z}|\boldsymbol{x})$ and $p^\theta(\boldsymbol{x}|\boldsymbol{z})$ can be obtained from (8) by marginalising over the auxiliary variables $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{L-1}$. Here, $\phi = (\phi_0, \ldots, \phi_{L-1})$ and $\theta = (\theta_1, \ldots, \theta_L)$ refer to sets of parameters to be specified in more detail below. Further introducing notation, we write $q^{\mu,\phi}(\boldsymbol{y}_{0:L}) := q^\phi(\boldsymbol{y}_{1:L}|\boldsymbol{y}_0)\mu(\boldsymbol{y}_0)$ as well as $p^{\nu,\theta}(\boldsymbol{y}_{0:L}) := p^\theta(\boldsymbol{y}_{0:L-1}|\boldsymbol{y}_L)\nu(\boldsymbol{y}_L)$ and think of those implied joint distributions as emanating from $\mu(\boldsymbol{x}) = \mu(\boldsymbol{y}_0)$ and $\nu(\boldsymbol{z}) = \nu(\boldsymbol{y}_L)$, respectively, moving 'forwards' or 'backwards' according to the specific choices for ϕ and θ. In the regime when L is large, the models in (8) are very expressive, even if the intermediate transition kernels are parameterised in a simple manner. We hence proceed by assuming Gaussian

distributions,

$$q^{\phi_{l-1}}(\boldsymbol{y}_l|\boldsymbol{y}_{l-1}) = \mathcal{N}(\boldsymbol{y}_l|\boldsymbol{y}_{l-1} + \delta a_{l-1}^{\phi}(\boldsymbol{y}_{l-1}), \delta\sigma^2 I),$$
$$p^{\theta_l}(\boldsymbol{y}_{l-1}|\boldsymbol{y}_l) = \mathcal{N}(\boldsymbol{y}_{l-1}|\boldsymbol{y}_l + \delta b_l^{\theta}(\boldsymbol{y}_l), \delta\sigma^2 I), \tag{9}$$

where $\sigma > 0$ controls the standard deviation, and $\delta > 0$ is a small parameter, anticipating the limits $L \to \infty$, $\delta \to 0$ to be taken in Section 2.2 below. The vector fields $a_l^{\phi}(\boldsymbol{y}_l)$ and $b_l^{\theta}(\boldsymbol{y}_l)$ introduced in (9) should be thought of as parameterised by ϕ and θ, but we will henceforth suppress this dependence for notational efficiency (in fact, those vector fields form a natural parameterisation in their own right).

The models (8)-(9) could equivalently be defined via the Markov chains

$$\boldsymbol{y}_{l+1} = \boldsymbol{y}_l + \delta a_l(\boldsymbol{y}_l) + \sqrt{\delta}\sigma\xi_l, \quad \boldsymbol{y}_0 \sim \mu$$
$$\implies \boldsymbol{y}_{0:L} \sim q^{\mu,\phi}(\boldsymbol{y}_{0:L}), \tag{10a}$$

$$\boldsymbol{y}_{l-1} = \boldsymbol{y}_l + \delta b_l(\boldsymbol{y}_l) + \sqrt{\delta}\sigma\xi_l, \quad \boldsymbol{y}_L \sim \nu$$
$$\implies \boldsymbol{y}_{0:L} \sim p^{\nu,\theta}(\boldsymbol{y}_{0:L}), \tag{10b}$$

where $(\xi_l)_{l=1}^L$ is an iid sequence of standard Gaussian random variables. As indicated, the forward process in (10a) may serve to define the distribution $q^{\mu,\phi}(\boldsymbol{y}_{0:L})$, whilst the backward process in (10b) induces $p^{\nu,\theta}(\boldsymbol{y}_{0:L})$.

Note that the transition densities $p^{\theta}(\boldsymbol{x}|\boldsymbol{z})$ and $q^{\phi}(\boldsymbol{z}|\boldsymbol{x})$ obtained as the marginals of (8) will in general not be available in closed form. However, generalising slightly from Framework 1, we may set out to minimise the extended loss

$$\mathcal{L}_D^{\text{ext}}(\phi,\theta) = D(q^{\mu,\phi}(\boldsymbol{y}_{0:L})||p^{\nu,\theta}(\boldsymbol{y}_{0:L})), \tag{11}$$

where D refers to a divergence on the 'discrete path space' $\{\boldsymbol{y}_{0:L}\}$. Clearly, $\mathcal{L}_D^{\text{ext}}(\phi,\theta) = 0$ still implies (6), but is no longer equivalent. More specifically, in the case when $D = D_{\text{KL}}$, the data processing inequality yields

$$D_{\text{KL}}(q^{\mu,\phi}(\boldsymbol{y}_{0:L})||p^{\nu,\theta}(\boldsymbol{y}_{0:L})) \geq$$
$$D_{\text{KL}}\left(q^{\phi}(\boldsymbol{z}|\boldsymbol{x})\mu(\boldsymbol{x})\middle||p^{\theta}(\boldsymbol{x}|\boldsymbol{z})\nu(\boldsymbol{z})\right), \tag{12}$$

so that $\mathcal{L}_{D_{\text{KL}}}^{\text{ext}}(\phi,\theta)$ provides an upper bound for $\mathcal{L}_{D_{\text{KL}}}(\phi,\theta)$ as defined in (5).

2.2. SDE transition densities – hierarchical VAEs in the infinite depth limit

In this section, we take inspiration from Section 2.1 and Tzen and Raginsky (2019); Li et al. (2020); Huang et al. (2021a) to develop the infinite-depth limit ($L \to \infty$), using the machinery of stochastic differential equations (SDEs). To this end, we think of $l = 0, \dots, L$ as discrete instances in a fixed time interval $[0, T]$, equidistant with time step δ, that is, we set $\delta = TL^{-1}$. The discrete paths $\boldsymbol{y}_{0:L}$ give rise to continuous paths $(\boldsymbol{Y}_t)_{0\leq t\leq T} \in C([0,T];\mathbb{R}^d)$ by setting $\boldsymbol{Y}_{\delta l} = \boldsymbol{y}_l$ and interpolating linearly between $\boldsymbol{Y}_{\delta l}$ and $\boldsymbol{Y}_{\delta(l+1)}$. To complete the set-up, we think of $a^{\phi} = (a_0^{\phi}, \dots, a_{L-1}^{\phi})$ and $b^{\theta} = (b_1^{\theta}, \dots, b_L^{\theta})$ in (9) as arising from time-dependent vector fields $a, b \in C^{\infty}([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$ via $a_l^{\phi}(\boldsymbol{y}_l) = a_{t\delta^{-1}}(\boldsymbol{Y}_{\delta l})$ and $b_l^{\theta}(\boldsymbol{y}_l) = b_{t\delta^{-1}}(\boldsymbol{Y}_{\delta l})$.

We are now prepared to invoke the limit $\delta \to 0$ (equivalently, $L \to \infty$), keeping $T > 0$ fixed. The proposed relabelling and rescaling transforms the Markov chains in (10) into continuous-time dynamics described by the SDEs

$$\mathrm{d}\boldsymbol{Y}_t = a_t(\boldsymbol{Y}_t)\,\mathrm{d}t + \sigma\,\overrightarrow{\mathrm{d}}\,\boldsymbol{W}_t, \quad \boldsymbol{Y}_0 \sim \mu$$
$$\implies (\boldsymbol{Y}_t)_{0\leq t\leq T} \sim \mathbb{Q}^{\mu,a} \equiv \overrightarrow{\mathbb{P}}^{\mu,a}, \tag{13a}$$

$$\mathrm{d}\boldsymbol{Y}_t = b_t(\boldsymbol{Y}_t)\,\mathrm{d}t + \sigma\,\overleftarrow{\mathrm{d}}\,\boldsymbol{W}_t, \quad \boldsymbol{Y}_T \sim \nu$$
$$\implies (\boldsymbol{Y}_t)_{0\leq t\leq T} \sim \mathbb{P}^{\nu,b} \equiv \overleftarrow{\mathbb{P}}^{\nu,b}, \tag{13b}$$

where $\overrightarrow{\mathrm{d}}$ and $\overleftarrow{\mathrm{d}}$ denote forward and backward Itô integration (see Appendix A for more details and remarks on the notation), and $(\boldsymbol{W}_t)_{0\leq t\leq T}$ is a standard Brownian motion. In complete analogy with (10), the SDEs in (13) induce the distributions $\mathbb{Q}^{\mu,a}$ and $\mathbb{P}^{\nu,b}$ on the path space $C([0,T];\mathbb{R}^d)$. Relating back to the discussion in the introduction, note that we maintain the relations $\boldsymbol{Y}_0 = \boldsymbol{x}$ and $\boldsymbol{Y}_T = \boldsymbol{z}$, and the transitions are parameterised by the vector fields a, b, in the sense that $p^{\theta}(\boldsymbol{x}|\boldsymbol{z}) = \mathbb{P}_0^{\nu,b^{\theta}}(\boldsymbol{x}|\boldsymbol{Y}_T = \boldsymbol{z}) = \mathbb{P}_0^{\delta_{\boldsymbol{z}},b^{\theta}}(\boldsymbol{x})$ and $q^{\phi}(\boldsymbol{z}|\boldsymbol{x}) = \mathbb{Q}_T^{\mu,a^{\phi}}(\boldsymbol{z}|\boldsymbol{Y}_0 = \boldsymbol{x}) = \mathbb{Q}_T^{\delta_{\boldsymbol{x}},a^{\phi}}(\boldsymbol{z})$.

The following well-known result (Anderson, 1982; Haussmann and Pardoux, 1985; Nelson, 1967) allows us to relate forward and backward path measures via a local (score-matching) condition for the reversal relation in (4). [2]

Proposition 2.1 (Nelson's relation)**.** *For fixed μ and a, denote the time-marginals of the corresponding path measure by $\mathbb{P}_t^{\mu,a} =: \rho_t^{\mu,a}$. Assuming that for all $t \in (0,T]$, ρ_t admits a strictly positive and smooth density with respect to the Lebesgue measure, we have that $\overrightarrow{\mathbb{P}}^{\mu,a} = \overleftarrow{\mathbb{P}}^{\nu,b}$ if and only if $\nu = \overrightarrow{\mathbb{P}}_T^{\mu,a}$ and*

$$b_t = a_t - \sigma^2\nabla\ln\rho_t^{\mu,a}, \quad \textit{for all } t \in (0,T]. \tag{14}$$

Remark 1. A similarly clean characterisation of equality between forward and backward path measures is not available for the discrete-time setting as presented in (10). In particular, Gaussianity of the intermediate transitions is not preserved under time-reversal, and thus the parameterisations in (9) are sub-optimal without the limit $\delta \to 0$.

A recurring theme in this work and related literature is the interplay between the score-matching condition in (14) and the global condition $D(\overrightarrow{\mathbb{P}}^{\mu,a}|\overleftarrow{\mathbb{P}}^{\nu,b}) = 0$, invoking Framework 1. To enable calculations involving the latter, we will rely on the following result:

[2] The global condition $\overrightarrow{\mathbb{P}}^{\mu,a} = \overleftarrow{\mathbb{P}}^{\nu,b}$ can be captured by the local condition (14) due to the Markovian nature of (13).

Proposition 2.2 (forward-backward Radon-Nikodym derivatives)**.** *Let $\overrightarrow{\mathbb{P}}^{\Gamma_0,\gamma^+} = \overleftarrow{\mathbb{P}}^{\Gamma_T,\gamma^-}$ be a reference path measure (that is, Γ_0, Γ_T and $\gamma^\pm$ define diffusions as in (13) and are related as in Proposition 2.1), absolutely continuous with respect to both $\overrightarrow{\mathbb{P}}^{\mu,a}$ and $\overleftarrow{\mathbb{P}}^{\nu,b}$. Then, $\overrightarrow{\mathbb{P}}^{\mu,a}$-almost surely, the corresponding Radon-Nikodym derivative (RND) can be expressed as follows,*

$$\ln\left(\frac{\mathrm{d}\overrightarrow{\mathbb{P}}^{\mu,a}}{\mathrm{d}\overleftarrow{\mathbb{P}}^{\nu,b}}\right)(\boldsymbol{Y}) = \ln\left(\frac{\mathrm{d}\mu}{\mathrm{d}\Gamma_0}\right)(\boldsymbol{Y}_0) - \ln\left(\frac{\mathrm{d}\nu}{\mathrm{d}\Gamma_T}\right)(\boldsymbol{Y}_T) \tag{15a}$$

$$+ \tfrac{1}{\sigma^2}\int_0^T \left(a_t - \gamma_t^+\right)(\boldsymbol{Y}_t)\cdot\left(\overrightarrow{\mathrm{d}}\boldsymbol{Y}_t - \tfrac{1}{2}\left(a_t + \gamma_t^+\right)(\boldsymbol{Y}_t)\,\mathrm{d}t\right) \tag{15b}$$

$$- \tfrac{1}{\sigma^2}\int_0^T \left(b_t - \gamma_t^-\right)(\boldsymbol{Y}_t)\cdot\left(\overleftarrow{\mathrm{d}}\boldsymbol{Y}_t - \tfrac{1}{2}\left(b_t + \gamma_t^-\right)(\boldsymbol{Y}_t)\,\mathrm{d}t\right). \tag{15c}$$

Proof. The proof relies on Girsanov's theorem (Üstünel and Zakai, 2013), using the reference to relate the forward and backward processes. For details, see Appendix E. □

Remark 2 (Role of the reference process)**.** According to Proposition 2.2, the Radon-Nikodym derivative between $\overrightarrow{\mathbb{P}}^{\mu,a}$ and $\overleftarrow{\mathbb{P}}^{\nu,b}$ can be decomposed into boundary terms involving the marginals in (15a), as well as forward and backward path integrals in (15b) and (15c). Note that the left-hand side of (15a) does not depend on the reference $\Gamma_{0,T}$, $\gamma^\pm$. Hence, the expressions in (15) are in principle equivalent for all choices of reference, and close inspection shows that the freedom in $\Gamma_{0,T}$ and $\gamma^\pm$ allows us to 'reweight' between (15a), (15b) and (15c), or even cancel terms.[3] A canonical choice is the Lebesgue measure for Γ_0 and Γ_T, and $\gamma^\pm = 0$, see Appendix C.1. For constructing algorithms, we may however choose the reference diffusion in such a way that intractable quantities are suppressed, or so that the variance of associated estimators is reduced. We will come back to these points in Section 3.

Remark 3 (Discretisation and conversion formulae)**.** The distinction between forward and backward integration in (15) is related to the time points at which the integrands $\left(a_t - \gamma_t^+\right)(\boldsymbol{Y}_t)$ and $\left(b_t - \gamma_t^-\right)(\boldsymbol{Y}_t)$ would be evaluated in discrete-time approximations, e.g.,

$$\int_0^T a_t(\boldsymbol{Y}_t)\cdot\overrightarrow{\mathrm{d}}\boldsymbol{Y}_t \approx \sum_i a_{t_i}(\boldsymbol{Y}_{t_i})\cdot(\boldsymbol{Y}_{t_{i+1}} - \boldsymbol{Y}_{t_i}),$$

$$\int_0^T a_t(\boldsymbol{Y}_t)\cdot\overleftarrow{\mathrm{d}}\boldsymbol{Y}_t \approx \sum_i a_{t_{i+1}}(\boldsymbol{Y}_{t_{i+1}})\cdot(\boldsymbol{Y}_{t_{i+1}} - \boldsymbol{Y}_{t_i}).$$

[3] Note that not all terms can be cancelled at the same time, since $\Gamma_{0,T}$, $\gamma_\pm$ cannot be chosen independently (they need to satisfy the conditions imposed by Proposition 2.1).

Alternatively, forward and backward integrals can be transformed into each other using the conversion formula

$$\int_0^T a_t(\boldsymbol{Y}_t)\cdot\overrightarrow{\mathrm{d}}\boldsymbol{Y}_t = \int_0^T a_t(\boldsymbol{Y}_t)\cdot\overleftarrow{\mathrm{d}}\boldsymbol{Y}_t - \sigma^2\int_0^T(\nabla\cdot a_t)(\boldsymbol{Y}_t)\,\mathrm{d}t. \tag{16}$$

We refer to Kunita (2019) and Appendix A for further details. In passing, we note that (16) allows us to eliminate the Hutchinson estimator (Hutchinson, 1989) from a variety of objectives, potentially reducing the variance of gradient estimators, see Appendix C.1.

Let us conclude this section by translating Framework 1 into the setting of (13), noting that (12) continues to hold with appropriate modifications.

Framework 1′. For a divergence D on path space, minimise $D(\overrightarrow{\mathbb{P}}^{\mu,a}|\overleftarrow{\mathbb{P}}^{\nu,b})$. If $D(\overrightarrow{\mathbb{P}}^{\mu,a}|\overleftarrow{\mathbb{P}}^{\nu,b}) = 0$, then (13a) transports μ to ν, and (13b) transports ν to μ.

At optimality, $D(\overrightarrow{\mathbb{P}}^{\mu,a}|\overleftarrow{\mathbb{P}}^{\nu,b}) = 0$, Proposition 2.1 allows us to obtain the scores associated to the learned diffusion via $\sigma^2\nabla\ln\rho_t^{\mu,a} = a_t - b_t$. In this way, Framework 1′ is closely connected to (and in some ways extends) score-matching ideas (Song and Ermon, 2019; Song et al., 2021).

3. Forward-reverse SDEs as a unifying perspective for generative modeling and sampling

We begin by showing that recent approaches towards generative modeling and sampling can be recovered from Framework 1′ by making specific choices for the divergence D, the parameterisations for a and b, as well as for the reference diffusion $\overrightarrow{\mathbb{P}}^{\Gamma_0,\gamma^+} = \overleftarrow{\mathbb{P}}^{\Gamma_T,\gamma^-}$ in Proposition 2.2. Recall from the introduction that complete flexibility in a and b will render the minima of $D(\overrightarrow{\mathbb{P}}^{\mu,a}|\overleftarrow{\mathbb{P}}^{\nu,b})$ highly nonunique, and so it is natural to fix either $\overrightarrow{\mathbb{P}}^{\mu,a}$ or $\overleftarrow{\mathbb{P}}^{\nu,b}$, or to impose further constraints relating the two (see Section 3.1).

Score-based generative modeling: Letting μ be the target and fixing the forward drift a_t, and, motivated by Proposition 2.1, parameterising the backward drift as $b_t = a_t - s_t$, we recover the SGM objectives in Hyvärinen and Dayan (2005); Song and Ermon (2019); Song et al. (2021) from $D = D_{\mathrm{KL}}$; when $\overrightarrow{\mathbb{P}}^{\mu,a} = \overleftarrow{\mathbb{P}}^{\nu,b}$, the variable drift component s_t will represent the score $\sigma^2\nabla\ln\rho_t^{\mu,a}$. Modifications can be obtained from the conversion formula (16), see Appendix C.2.

Score-based sampling – ergodic drift: In this setting, ν becomes the target and we fix b_t to be the drift of an ergodic (backward) process. Then choosing $\Gamma_{0,T} = \mu$, $\gamma^\pm = b$ allows us to recover the approaches in Vargas et al. (2023a); Berner et al. (2022). Possible generalisations based on Framework 1′ include IWAE-type objectives, see Appendix C.3.

Score-based sampling – Föllmer drift: Finally choosing $b_t(x) = x/t$ we can recover the Föllmer drift approach to sampling (Follmer, 1984; Vargas et al., 2023b; Zhang and Chen, 2022; Huang et al., 2021b), see Appendix C.3.

Domain adaptation and stochastic filtering: Setting $\overrightarrow{\mathbb{P}}^{\Gamma_0,\gamma^+}$ equal to the signal process in stochastic filtering (Reich and Cotter, 2015), we can target the corresponding Bayesian posterior with log-likelihood $\ln(\nu/\Gamma_T)$. Note that to obtain a tractable objective, we would need in a preliminary step to learn γ^- using score matching techniques. Similarly, we can design a reference process with $\Gamma_T = \nu$ in case the log-density of ν is not available, as in domain adaptation (Courty et al., 2017).

3.1. Score-based annealed flows

Departing from the discussion so far, we now attempt to vary both $\overrightarrow{\mathbb{P}}^{\mu,a}$ and $\overleftarrow{\mathbb{P}}^{\nu,b}$ simultaneously. To retain uniqueness, and guided by Proposition 2.1, we impose the additional constraint $b_t = a_t - \sigma^2 \nabla \ln \pi_t$, so that when $\overrightarrow{\mathbb{P}}^{\mu,a} = \overleftarrow{\mathbb{P}}^{\nu,b}$, the time marginals necessarily satisfy $\overrightarrow{\mathbb{P}}^{\mu,a}_t = \overleftarrow{\mathbb{P}}^{\nu,b}_t = \pi_t$, for all $t \in [0,T]$. In this construction, $(\pi_t)_{t\in[0,T]}$ is a prescribed curve of distributions whose scores (and unnormalised densities $\hat{\pi}_t$) are assumed to be available in tractable form; this is the scenario typically considered in annealed importance sampling and related approaches towards computing posterior expectations (Neal, 2001; Reich, 2011; Heng et al., 2021; 2020; Arbel et al., 2021; Doucet et al., 2022). To complete the setting, we fix the parameterisation $a_t = \sigma^2 \nabla \ln \pi_t + \nabla \phi_t$ in terms of a scalar function ϕ_t,[4] so that the forward process is given by the controlled time-inhomogeneous Langevin dynamics

$$\boldsymbol{Y}_0 \sim \pi_0,$$
$$\mathrm{d}\boldsymbol{Y}_t = \big(\sigma^2 \nabla \ln \pi_t(\boldsymbol{Y}_t) + \nabla \phi_t(\boldsymbol{Y}_t)\big)\,\mathrm{d}t + \sigma\sqrt{2}\,\overrightarrow{\mathrm{d}}\boldsymbol{W}_t. \quad (17)$$

Using the score constraint, we obtain $b_t = -\sigma^2 \nabla \ln \pi_t + \nabla \phi_t$ for the backward drift. Notice that for $\phi = 0$, the dynamics is 'locally at equilibrium', that is, $\boldsymbol{Y}_t \sim \pi_0$ for all $t \in [0,T]$, if $\pi_t = \pi_0$, for all $t \in [0,T]$. Consequently, $\nabla \phi_t$ may be thought of as a control enabling the transition between neighbouring densities π_t and $\pi_{t+\delta t}$.

In order to learn the drift $\nabla \phi_t$ so that (17) reproduces $(\pi_t)_{t\in[0,T]}$, we can define the loss

$$\mathcal{L}(\phi) := D\left(\overrightarrow{\mathbb{P}}^{\pi_0,\sigma^2\nabla \ln \pi + \nabla\phi}, \overleftarrow{\mathbb{P}}^{\pi_T,-\sigma^2\nabla \ln \pi + \nabla\phi}\right), \quad (18)$$

as per Framework 1′, and choose for instance the KL or the log-variance divergence (see Appendix B) for D, where the RND can be computed and discretised as per equation (73) in Appendix F.1. Additionally, we can simplify the RND to observe the following connection to normalising flows:

[4] Restricting a_t further to be of gradient form ensures uniqueness, see Proposition 3.1 below.

Remark 4. Consider the log-variance loss, which can be re-expressed as

$$\mathcal{L}(\phi) := \mathrm{Var}\left(\ln \frac{\mathrm{d}\,\overrightarrow{\mathbb{P}}^{\pi_0,\sigma^2\nabla \ln \pi + \nabla\phi}}{\mathrm{d}\,\overleftarrow{\mathbb{P}}^{\pi_T,-\sigma^2\nabla \ln \pi + \nabla\phi}}\right) =$$
$$= \mathrm{Var}\Bigg(\ln \pi_T(\boldsymbol{Y}_T) - \ln \pi_0(\boldsymbol{Y}_0) + \int_0^T \Delta\phi_t(\boldsymbol{Y}_t)\,\mathrm{d}t$$
$$-\sigma\sqrt{2}\int_0^T \nabla \ln \pi_t(\boldsymbol{Y}_t) \circ \mathrm{d}\boldsymbol{W}_t - \sigma^2 \int_0^T |\nabla \ln \pi_t(\boldsymbol{Y}_t)|^2 \mathrm{d}t\Bigg), \quad (19)$$

see Appendix D. Here, $\circ\mathrm{d}\boldsymbol{W}_t$ denotes Stratonovich integration, and the variance is taken with respect to samples from (17). We note that in the limit $\sigma \to 0$, the objective in (19) enforces an integrated version of the instantaneous change of density formula $\partial_t \ln \pi_t(\boldsymbol{Y}_t) = -\Delta\phi_t(\boldsymbol{Y}_t)$ for continuous time normalising flows of the form $\dot{\boldsymbol{Y}}_t = \nabla\phi_t(\boldsymbol{Y}_t)$, see Papamakarios et al. (2021, Section 4). Notwithstanding the intuitive appeal of the term involving $\Delta\phi_t$, it may be favourable computationally to replace it by a combination of forward and backward integrals using (16), see Remark 13.

The additional constraints detailed above resolve the nonuniqueness in Frameworks 1 and 1′:

Proposition 3.1. *Under mild conditions on $(\pi_t)_{t\in[0,T]}$, (18) admits a (π_t-a.e.)unique minimiser, up to additive constants.*

We provide a proof in Appendix D; a numerical evaluation is postponed to future work.

Remark 5 (Detaching the gradient)**.** Since the log-variance divergence is valid regardless of the base distribution according to which the variance is computed, it is possible to detach the gradient with respect to $\boldsymbol{Y}_t$ when optimising (19) (Nüsken and Richter, 2021; Richter et al., 2020), avoiding differentiation through the score $\nabla \ln \pi_t$ from which a D_{KL}-based objective would suffer.

Remark 6 (Action matching)**.** Neklyudov et al. similarly propose to learn ϕ_t in such a way that (17) reproduces a fixed curve of distributions $(\pi_t)_{t\in[0,T]}$, albeit from samples rather than from available scores. We show in Appendix C.4 that their objective can be recovered from Framework 1′ using a diffusion with marginals $(\pi_t)_{t\in[0,T]}$ as the reference process.

Finally following Zhang and Chen (2022); Vargas et al. (2023a), we can obtain an unbiased estimate of the normalising Z in $\pi_T = \hat{\pi}_T/Z$ from the RND in (19) via the following importance sampling identity:

$$S(\boldsymbol{Y}) = \ln \pi_T(\boldsymbol{Y}_T) - \ln \pi_0(\boldsymbol{Y}_0) + \int_0^T \Delta\phi_t(\boldsymbol{Y}_t)\,\mathrm{d}t$$
$$-\sigma\sqrt{2}\int_0^T \nabla \ln \pi_t(\boldsymbol{Y}_t) \circ \mathrm{d}\boldsymbol{W}_t - \sigma^2 \int_0^T |\nabla \ln \pi_t(\boldsymbol{Y}_t)|^2 \mathrm{d}t,$$

$$\hat{Z} = e^{-S(\boldsymbol{Y})}, \tag{20}$$

where, as mentioned previously, we can use the more computationally friendly expression for the RND in equation (73), which as in Vargas et al. (2023a) provides an ELBO for $\hat{Z}$ when discretised (see Remark 17).

3.2. Fluctuation theorem and the Jarzynski equality

To shed further light on score-based annealed flows, we examine the RND involved in (19) when $\phi = 0$. Connecting (20) to well-established importance sampling (IS) estimators for partition functions (Jarzynski, 1997; Neal, 2001; Chopin, 2002) we will motivate how our approach in Section 3.1 allows us to improve on said estimators. At the same time, we showcase an application of Proposition 2.2.

To establish the aforementioned connection we present a novel derivation of the Jarzynski equality (Jarzynski, 1997) and Crooks' fluctuation theorem (Crooks, 1999) for diffusion processes which is more in line with Crooks' original proof. Our derivation appears to be more succinct and 'pathwise', in contrast to previous approaches which combine the Feynman-Kac formula with the FPK equation (Chen et al., 2019; Stoltz et al., 2010; Hartmann et al., 2019; Hummer and Szabo, 2001).

Proposition 3.2 (Crooks' fluctuation theorem)**.** *Given the annealed flow $\pi_t = \hat{\pi}_t/\mathcal{Z}_t$ interpolating between π_0 and π_T, we define* work done *and* free energy *as*

$$\mathcal{W}_T(\boldsymbol{Y}) = -\int_0^T \sigma^2 \partial_t \ln \hat{\pi}_t(\boldsymbol{Y}_t)\,\mathrm{d}t, \quad \mathcal{F}_t = -\sigma^2 \ln \mathcal{Z}_t,$$

see Jarzynski (1997); Chen et al. (2019). Then we have Crooks' identity*,*

$$\left(\frac{\mathrm{d}\overrightarrow{\mathbb{P}}^{\pi_0,\sigma^2\nabla\ln\pi}}{\mathrm{d}\overleftarrow{\mathbb{P}}^{\pi_T,-\sigma^2\nabla\ln\pi}}\right)(\boldsymbol{Y}) = e^{-\frac{1}{\sigma^2}(\mathcal{F}_T-\mathcal{F}_0)} e^{\frac{1}{\sigma^2}\mathcal{W}_T(\boldsymbol{Y})},$$

$\overrightarrow{\mathbb{P}}^{\pi_0,\sigma^2\nabla\ln\pi}$*-a.s., which implies* Jarzynski's equality

$$\mathbb{E}_{\overrightarrow{\mathbb{P}}^{\pi_0,\sigma^2\nabla\ln\pi}}\left[e^{-\frac{1}{\sigma^2}\mathcal{W}_T}\right] = \frac{\mathcal{Z}_T}{\mathcal{Z}_0} \tag{21}$$

by taking expectations.

The proof directly uses Proposition 2.2 and the conversion formula (16) to compute the RND $\mathrm{d}\overrightarrow{\mathbb{P}}^{\pi_0,\sigma^2\nabla\ln\pi}/\mathrm{d}\overleftarrow{\mathbb{P}}^{\pi_T,-\sigma^2\nabla\ln\pi}$, followed by an application of Itô's formula to $t \mapsto \ln\hat{\pi}_t(\boldsymbol{Y}_t)$, see Appendix E.2.

Prior works (Vaikuntanathan and Jarzynski, 2008; Neal, 2001; Chopin, 2002) have made use of the Jarzynski equality (21) to estimate partition functions via importance sampling. However, such choices for forward and backward processes do not yield optimal variance for the estimator (Del Moral et al., 2006), and indeed standard Monte Carlo estimators for (21) might exhibit high variance, see Stoltz et al. (2010, Section 4.1.4). In contrast, our proposed approach allows us to optimise the vector field $\nabla\phi_t$ which if optimised fully would make the IS estimator (20) optimal (zero variance).

Remark 7 (Related work)**.** The task of learning the vector field $\nabla\phi_t$ so that (17) reproduces $(\pi_t)_{t\in[0,T]}$ has been approached from various directions. Reich (2011); Heng et al. (2021); Reich (2022); Vaikuntanathan and Jarzynski (2008) explore methodologies that exploit the characterisation of $\nabla\phi_t$ in terms of the elliptic PDE (55). Arbel et al. (2021) propose to leverage normalising flows sequentially to minimise KL divergences between implied neighboring densities. In an appropriate limiting regime, they recover the SDE (17), see Remark 14. More closely related to the objective in (19) are physics-inspired attempts to optimise the IS-variance in (21), see (Hartmann et al., 2019; Doucet et al., 2022; Zhang, 2021), however, these methods are based on the time reversal of ULA diffusions and unlike ours do not reproduce the interpolating distribution $(\pi_t)_{t\in[0,T]}$.

4. Entropic optimal transport

Another way of selecting a particular transition between μ and ν is by imposing an entropic penalty, encouraging the dynamics to stay close to a prescribed, oftentimes physically or biologically motivated, reference process. Using the notation employed in Framework 1, the *static* Schrödinger problem (Schrödinger, 1931) is given by

$$\pi^*(\boldsymbol{x},\boldsymbol{z}) \in \underset{\pi(\boldsymbol{x},\boldsymbol{z})}{\arg\min}\Big\{D_{\mathrm{KL}}(\pi(\boldsymbol{x},\boldsymbol{z})||r(\boldsymbol{x},\boldsymbol{z})) : \pi_{\boldsymbol{x}}(\boldsymbol{x}) = \mu(\boldsymbol{x}), \pi_{\boldsymbol{z}}(\boldsymbol{z}) = \nu(\boldsymbol{z})\Big\}, \tag{22}$$

where $r(\boldsymbol{x},\boldsymbol{z})$ is the Schrödinger prior encoding additional domain-specific information. In the context of Framework 1′, the *dynamical* Schrödinger problem can be presented in the form

$$\mathbb{P}^* \in \underset{\overrightarrow{\mathbb{P}}_T^{\mu,a}=\nu}{\arg\min}\, \mathbb{E}_{\boldsymbol{Y}\sim\overrightarrow{\mathbb{P}}^{\mu,a}}\left[\frac{1}{2\sigma^2}\int_0^T \|a_t - f_t\|^2(\boldsymbol{Y}_t)\,\mathrm{d}t\right], \tag{23}$$

that is, the driving vector field a_t determining $\mathbb{P}^*$ should be chosen in such a way that (i), the corresponding diffusion transitions from μ to ν, and (ii), among such diffusions, the vector field a_t remains close to the prescribed vector field f_t, in mean square sense. Under mild conditions, it can be shown that solutions to the problems in (22) and (23) exist and are unique, and that the static and dynamic viewpoints are related through a mixture-of-bridges construction (assuming that the conditionals $r(\boldsymbol{z}|\boldsymbol{x})$ correspond to the transitions induced by f_t),[5] see (Léonard, 2014a, Section

[5]This correspondence shows that the solution in (22) only depends on the conditionals $r(\boldsymbol{z}|\boldsymbol{x})$, and not on the full distribution $r(\boldsymbol{x},\boldsymbol{z})$.

2). To make this relationship plausible, we remark that the D_{KL}-divergence in (22) can be translated into the time integral in (23) using Girsanov's theorem.

4.1. Iterative proportional fitting (IPF) and the EM algorithm

It is well known that approximate solutions for $\pi^*(\boldsymbol{x}, \boldsymbol{z})$ and $\mathbb{P}^*$ can be obtained using alternating D_{KL}-projections, keeping one of the marginals fixed in each iteration: Under mild conditions, the sequence defined by

$$\pi^{2n+1}(\boldsymbol{x}, \boldsymbol{z}) = \tag{24a}$$
$$\arg\min_{\pi(\boldsymbol{x},\boldsymbol{z})} \left\{ D_{\mathrm{KL}}(\pi(\boldsymbol{x}, \boldsymbol{z}) || \pi^{2n}(\boldsymbol{x}, \boldsymbol{z})) : \ \pi_{\boldsymbol{x}}(\boldsymbol{x}) = \mu(\boldsymbol{x}) \right\},$$

$$\pi^{2n+2}(\boldsymbol{x}, \boldsymbol{z}) = \tag{24b}$$
$$\arg\min_{\pi(\boldsymbol{x},\boldsymbol{z})} \left\{ D_{\mathrm{KL}}(\pi(\boldsymbol{x}, \boldsymbol{z}) || \pi^{2n+1}(\boldsymbol{x}, \boldsymbol{z})) : \ \pi_{\boldsymbol{z}}(\boldsymbol{z}) = \nu(\boldsymbol{z}) \right\},$$

for $n \geq 0$, with initialisation $\pi^0(\boldsymbol{x}, \boldsymbol{z}) = r(\boldsymbol{x}, \boldsymbol{z})$, is known to converge to $\pi^*(\boldsymbol{x}, \boldsymbol{z})$ as $n \to \infty$ (De Bortoli et al., 2021), and this procedure is commonly referred to as iterative proportional fitting (IPF) (Fortet, 1940; Kullback, 1968; Ruschendorf, 1995; Reich, 2019) or Sinkhorn updates (Cuturi, 2013). IPF can straightforwardly be modified to the path space setting of (23), and the resulting updates coincide with the Föllmer drift updates discussed in Section C.3, see (Vargas et al., 2021a). In applications, however, IPF is faced with the following challenges:

1. The sequential nature of IPF, coupled with the need for each iteration to undergo comprehensive training as outlined in Section C.3, results in significant computational demands.
2. The reference distribution $r(\boldsymbol{x}, \boldsymbol{z})$ (or the reference vector field f_t) enters the iterations in (24) only through the initialisation. As a consequence, numerical errors accumulate, and it is often observed that the Schrödinger prior is 'forgotten' as IPF proceeds (Vargas et al., 2021a; Fernandes et al., 2021; Shi et al., 2023).

To address these issues, we first establish a connection to expectation-maximization (EM) (Dempster et al., 1977), originally devised for finding maximum likelihood estimates in models with latent (or hidden) variables. According to Neal and Hinton (1998), the EM-algorithm can be described in the setting from the introduction, and written in the form

$$\theta_{n+1} = \arg\min_{\theta} \mathcal{L}_{D_{\mathrm{KL}}}(\phi_n, \theta),$$
$$\phi_{n+1} = \arg\min_{\phi} \mathcal{L}_{D_{\mathrm{KL}}}(\phi, \theta_{n+1}), \tag{25}$$

with $\mathcal{L}_{D_{\mathrm{KL}}}$ defined as in (5). If the initialisations are matched appropriately, the following result establishes an exact correspondence between the IPF updates in (24) and the EM updates in (25):

Proposition 4.1 (EM $\iff$ IPF)**.** *Assume that the transition densities $p^\theta(\boldsymbol{x}|\boldsymbol{z})$ and $q^\phi(\boldsymbol{z}|\boldsymbol{x})$ are parameterised with perfect flexibility,[6] and furthermore that the EM-scheme (25) is initialised at ϕ_0 in such a way that $q^{\phi_0}(\boldsymbol{z}|\boldsymbol{x}) = r(\boldsymbol{z}|\boldsymbol{x})$. Then the IPF iterations in (24) agree with the EM iterations in (25) for all $n \geq 1$, in the sense that*

$$\pi^n(\boldsymbol{x}, \boldsymbol{z}) = q^{\phi_{(n-1)/2}}(\boldsymbol{z}|\boldsymbol{x})\mu(\boldsymbol{x}), \quad \textit{for } n \textit{ odd},$$
$$\pi^n(\boldsymbol{x}, \boldsymbol{z}) = p^{\theta_{n/2}}(\boldsymbol{x}|\boldsymbol{z})\nu(\boldsymbol{z}), \quad \textit{for } n \textit{ even}. \tag{26}$$

Proof. See Appendix E. □

Corollary 4.2 (Path space EM)**.** *For the initialisation $\phi_0 = 0$ the alternating scheme*

$$\theta_{n+1} = \arg\min_{\theta} D_{\mathrm{KL}}(\overrightarrow{\mathbb{P}}^{\mu, f+\sigma^2\nabla\phi_n}, \overleftarrow{\mathbb{P}}^{\nu, f+\sigma^2\nabla\theta}),$$
$$\phi_{n+1} = \arg\min_{\phi} D_{\mathrm{KL}}(\overrightarrow{\mathbb{P}}^{\mu, f+\sigma^2\nabla\phi}, \overleftarrow{\mathbb{P}}^{\nu, f+\sigma^2\nabla\theta_{n+1}})$$

agrees with the path space IPF iterations in (Bernton et al., 2019; Vargas et al., 2021a; De Bortoli et al., 2021).

The key observation in the proof of Proposition 4.1 is that replacing forward-D_{KL} by reverse-D_{KL} in one or both of (24a) and (24b) does not – in theory – change the sequence of minimisers. In practice, however, optimisation with respect to forward-D_{KL} and backward-D_{KL} encourages moment-matching and mode-seeking behaviour, respectively, and so an alternating scheme as defined in (25) might present a suitable compromise. For comparisons of IPF using different directions of D_{KL} we refer the reader to Vargas et al. (2021a); Zhang et al. (2014).

4.2. Optimality conditions and HJB-regularisers

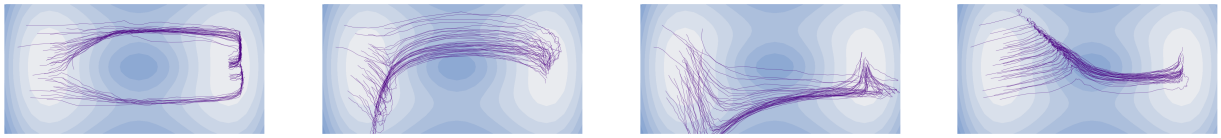
As per Section 4.1, IPF resolves the nonquniqueness in minimising $\mathcal{L}_D(\phi, \theta)$ by performing the coordinate-wise updates (25) starting from an initialisation informed by the Schrödinger prior. On the basis of this observation, the joint updates $(\phi_{n+1}, \theta_{n+1}) \leftarrow (\phi_n, \theta_n) - h\nabla_{\phi,\theta}\mathcal{L}_D(\phi, \theta)$ suggest themselves, in the spirit of VAEs (Kingma et al., 2019) and as already proposed by Neal and Hinton (1998). However, as is clear from the introduction, the limit $\lim_{n\to\infty}(\phi_n, \theta_n)$, can merely be expected to respect the marginals in (6), and no optimality in the sense of (23) is obtained. As a remedy, we recall the following result:

Proposition 4.3 (Mean-field game formulation)**.** *Assume that $\phi \in C^{1,2}([0, T]\times\mathbb{R}^d;\mathbb{R})$ satisfies the conditions:*

[6] In precise terms, we assume that for any transition density $p(\boldsymbol{x}|\boldsymbol{z})$, there exists θ_* such that $p(\boldsymbol{x}|\boldsymbol{z}) = p^{\theta_*}(\boldsymbol{x}|\boldsymbol{z})$. Likewise, we assume that for any transition density $q(\boldsymbol{z}|\boldsymbol{x})$, there exists ϕ_* such that $q(\boldsymbol{z}|\boldsymbol{x}) = q^{\phi_*}(\boldsymbol{z}|\boldsymbol{x})$.

Target	Method	KL		SBP Loss		PINN Loss		Cross Ent	
		Val	Train	Val	Train	Val	Train	Val	Train
tree	$\lambda = 0.5$	1.67±0.02	1.40±0.01	**47.84±1.58**	**42.31±1.52**	**0.06±0.00**	**0.05±0.00**	2.87±0.01	2.80±0.01
	DNF (EM Init)	1.63±0.02	1.39±0.01	55.33±1.79	49.60±1.68	1.74±0.04	1.64±0.04	2.88±0.01	2.80±0.01
olympics	$\lambda = 0.5$	2.95±0.06	0.12±0.01	39.30±0.90	**25.24±0.62**	**0.26±0.01**	**0.10±0.00**	2.49±0.01	2.77±0.02
	DNF (EM Init)	**2.70±0.05**	0.02±0.01	40.20±0.77	38.30±1.53	1.64±0.04	2.05±0.08	**2.54±0.01**	2.77±0.02
sierpinski	$\lambda = 0.5$	2.31±0.01	2.20±0.01	28.54±1.49	**26.67±0.90**	**0.04±0.00**	**0.03±0.00**	2.82±0.01	2.83±0.01
	DNF (EM Init)	2.30±0.01	2.20±0.00	30.87±1.93	29.53±1.18	7.25±0.14	7.22±0.14	2.80±0.01	2.82±0.02
swirl	$\lambda = 0.5$	15.67±0.29	1.95±0.03	**121.81±1.94**	**40.24±1.74**	**1.01±0.03**	**0.14±0.00**	2.97±0.01	2.69±0.02
	DNF (EM Init)	**13.77±0.38**	1.92±0.04	151.67±3.68	67.55±1.86	5.89±0.15	2.63±0.08	2.95±0.02	2.74±0.03
checkerboard	$\lambda = 0.5$	4.79±0.01	4.70±0.01	**34.47±0.80**	**33.70±0.91**	**0.03±0.00**	**0.02±0.00**	2.82±0.00	2.81±0.01
	DNF (EM Init)	4.78±0.01	4.70±0.02	39.76±0.83	39.20±1.10	3.66±0.07	3.68±0.06	2.81±0.01	2.81±0.02

Table 1. Generative Modelling Results comparing DNF (Zhang and Chen, 2021) ($\lambda = 0$) to our PINN regualirsed approach with $\lambda = 0.5$. We observe that PINN regularisation obtains similar KL and Cross entropy losses to DNF whilst achieving lower distances to the prior.



(a) $\lambda = 2$ (b) $\lambda = 0$ (EM Init) (c) $\lambda = 0$ (Random Init) (d) DNF (No Prior)

Figure 1. (a) our proposed regularised objective, (b) λ set to 0 but using clever EM motivated initialisation, (c) λ set to 0 with random initialisation of the forward drift, (d) for reference DNF with $f_t = 0$ (uninformative Schrödinger prior).

1. *The forward SDE*

$$\mathrm{d}\boldsymbol{Y}_t = f_t(\boldsymbol{Y}_t)\,\mathrm{d}t + \sigma^2\nabla\phi_t(\boldsymbol{Y}_t)\,\mathrm{d}t + \sigma\,\overrightarrow{\mathrm{d}}\boldsymbol{W}_t,\ \boldsymbol{Y}_0 \sim \mu \tag{27}$$

admits a unique strong solution on $[0,T]$*, satisfying moreover the terminal constraint* $\boldsymbol{Y}_T \sim \nu$*.*

2. *The Hamilton-Jacobi-Bellmann (HJB) equation*

$$\partial_t\phi + f\cdot\nabla\phi + \frac{\sigma^2}{2}\Delta\phi + \tfrac{\sigma^2}{2}|\nabla\phi|^2 = 0 \tag{28}$$

holds for all $(t,x) \in [0,T]\times\mathbb{R}^d$*.*

Then $a = \sigma^2\nabla\phi$ *provides the unique solution to the dynamical Schrödinger problem as posed in (23).*

For a proof, see Chen et al. (2021, Proposition 5.1) or Appendix E. Proposition 4.3 identifies the HJB-equation (28) as the missing link that renders joint minimisation of objectives of the form $D(\overrightarrow{\mathbb{P}}^{\mu,a}|\overleftarrow{\mathbb{P}}^{\nu,b})$ theoretically sound for solving (23):

Corollary 4.4. *For* $\lambda > 0$ *and* $\phi, \psi \in C^{1,2}([0,T]\times\mathbb{R}^d;\mathbb{R})$*, let*

$$\mathcal{L}_{\mathrm{Schr}}(\phi,\theta) := D_{\mathrm{KL}}(\overrightarrow{\mathbb{P}}^{\mu,f+\sigma^2\nabla\phi}, \overleftarrow{\mathbb{P}}^{\nu,f+\sigma^2\nabla\theta}) + \lambda\mathrm{Reg}(\phi), \tag{29}$$

where D *is a divergence on path space, and* $\mathrm{Reg}(\phi) = 0$ *if and only if (28) is satisfied. Then* $\mathcal{L}_{\mathrm{Schr}}(\phi,\theta) = 0$ *implies that the drift* $a_t = \sigma^2\nabla\phi_t$ *solves (23).*

In the experiments in Section 5 we choose $D = D_{\mathrm{KL}}$ and the PINN (Raissi et al., 2019) regulariser

$$\mathrm{Reg}(\phi) = \int_0^T \mathbb{E}\left|\partial_t\phi + f\cdot\nabla\phi + \frac{\sigma^2}{2}\Delta\phi + \tfrac{\sigma^2}{2}|\nabla\phi|^2\right|(\boldsymbol{Y}_t)\,\mathrm{d}t, \tag{30}$$

but other choices are possible, see Remarks 8 and 15 below. The objective (29) can now be optimised, jointly in (ϕ,θ), without reverting to the block-form IPF updates in (25). Procedures based on (29) hold the promise of alleviating the drawbacks of IPF mentioned in Section 4.1:

1. The objective $\mathcal{L}_{\mathrm{Schr}}$ allows for joint updates in (ϕ,θ), without the need to follow the iterative IPF schedule.
2. Since the vector field f_t enters the objective directly through $\mathrm{Reg}(\phi)$, see (30), optimality in the sense of (23) is enforced throughout training.

Remark 8 (Relationship to previous work)**.** For $\lambda = 0$, coordinate-wise updates of $\mathcal{L}_{\mathrm{Schr}}(\phi,\theta)$ recover the IPF updates from De Bortoli et al. (2021); Vargas et al. (2021a) according to Corollary 4.1. Note that $\mathcal{L}_{\mathrm{Schr}}$ is an unconstrained objective, in contrast to (23); previous works (Koshizuka and Sato, 2023; Zhang and Katsoulakis, 2023) have suggested incorporating the marginal constraints softly by adding penalising terms to the running cost in (23). Those approaches require a limiting argument (from an algorithmic standpoint, adaptive tuning of a weight parameter) to recover the solution to (23). In contrast, the conclusion of

Corollary 4.4 holds for arbitrary $\lambda > 0$. Shi et al. (2023); Peluchetti (2023) suggest an algorithm involving reciprocal projections onto the reciprocal class associated to f_t. From (Clark, 1991; Thieullen, 2002; Rœlly, 2013), the HJB-equation (28) is a local characteristic ($\mathrm{Reg}(\phi) = 0$ forces (27) to be in the reciprocal class); hence $\mathrm{Reg}(\phi)$ in (29) plays a similar role as the reciprocal projection (Shi et al., 2023, Definition 3), see Remark 15. Liu et al. (b) suggest an iterative IPF-like scheme involving a temporal difference term (Sutton and Barto, 2018, Chapter 6). As in Nüsken and Richter (2023), this is a an HJB-regulariser in the sense of Corollary 4.4, see Remark 15. Finally, Albergo et al. (2023, Theorem 5.3) and Gushchin et al. (2022) develop saddle-point objectives for (23).

Remark 9 (Role of the backward drift $\nabla\theta_t$)**.** The transtion from the one-drift objective (23) to the two-drift objective (29) allows us to perform optimisation without constraints. Enforcing the backward drift to be of gradient form is not essential, and $\nabla\theta$ may be replaced by v in Corollary 4.4, allowing for more flexible NN-parameterisations. It is also possible to introduce a regulariser based on θ and a suitable HJB equation; we elaborate on this in Remark 16.

5. Experiments

Across our experiments, we use D_{KL} and let $\Gamma_0 = \Gamma_T = \mathrm{Leb}$, which can be simplified to the forward-backwards KL objective used in DNF (Zhang et al., 2014), see Appendix F.1. We use the Adam optimiser (Kingma and Ba, 2015) trained on 50,000 samples and batches of size 5000 following Zhang and Chen (2021). For the generative modelling tasks we use 30 time steps and train for 100 epochs whilst for the double well we train all experiments for 17 epochs (early stopping via the validation set) and 60 discretisation steps. Finally note we typically compare our approach with $\lambda > 0$ to DNF ($\lambda = 0$), with DNF initialised at the reference process, which we call DNF (EM Init), see Appendix E.4 for further details.

5.1. 2D toy targets – generative modelling

Here we consider the suite of standard 2D toy targets for generative modelling explored in Zhang and Chen (2021) In contrast to Zhang and Chen (2021) we consider the SDE $\mathrm{d}\boldsymbol{Y}_t = -\sigma^2 \boldsymbol{Y}_t \,\mathrm{d}t + \sigma\sqrt{2}\,\mathrm{d}\boldsymbol{W}_t$ as the Schrödinger prior across methods. We parametrise DNF and our proposed approach with the same architectures for a fair comparison. Furthermore, we incorporate the drift of the above Schrödinger prior into DNF via parameterising the forward drift as in (27), partly motivated by Corollary 4.2.

In order to assess the quality of the bridge we consider three different error metrics. Firstly we estimate D_{KL} between the Schrödinger prior and the learned forward process (i.e. $\mathbb{E}_{\boldsymbol{Y}\sim\overrightarrow{\mathbb{P}}^{\mu,a}}\left[\frac{1}{2\sigma^2}\int_0^T \|a_t - f_t\|^2(\boldsymbol{Y}_t)\,\mathrm{d}t\right]$). Secondly, we evaluate $D_{\mathrm{KL}}(\overrightarrow{\mathbb{P}}^{\mu,f+\sigma^2\nabla\phi}, \overleftarrow{\mathbb{P}}^{\nu,f+\sigma^2\nabla\theta})$ to obtain a proxy error between the learned and target marginals. Finally, we estimate the cross entropy between $\overrightarrow{\mathbb{P}}^{\mu,a}_T$ and ν to assess how well the constraint at time T is met.

In Table 1 we observe that similar values of D_{KL} are attained across both approaches in the tree, sierpinski, and checkerboard datatsets whilst achieving significantly lower values of the SBP loss across all training sets, and for tree, swirl and checkerboard validation datasets. At the same time, we can see that the cross-entropy errors are effectively the same across both approaches. Overall we can conclude that on the empirical measures over which we train our approach, we obtain a much better fit for the target Schrödinger bridge, and on the validation results we can see that we generalise to 3/5 datasets in improving the bridge quality whilst preserving the marginals to a similar quality.

5.2. Double well – rare event

In this task we consider the double well potential explored in (Vargas et al., 2021b; Hartmann et al., 2013) where the Schrödinger prior is specified via the following overdamped Langevin dynamics $\mathrm{d}\boldsymbol{Y}_t = -\nabla_{\boldsymbol{Y}_t} U(\boldsymbol{Y}_t)\,\mathrm{d}t + \sigma\,\mathrm{d}\boldsymbol{W}_t$. The potential $U(\boldsymbol{y})$ typically models a landscape for which it is difficult to transport μ into ν.

This is a notably challenging task as we are trying to sample a rare event and as noted by Vargas et al. (2021a) many runs would result in collapsing into one path rather than bifurcating. In Figure 1 we can observe how our proposed regularised approach (1a) is able to successfully transport particles across the well whilst respecting the potential, whilst both variants of DNF using the EM-Init for ϕ (1b) and random init (1c) fail to respect the prior as nicely and do not bifurcate, with the random init in particular sampling quite inconsistent trajectories. Finally for reference we train a DNF model with $f_t = 0$ and ϕ (1d) initialised at random to illustrate the significance of the initialisation of ϕ.

6. Discussion

Overall we have successfully introduced a novel variational framework bridging VI and transport using modern advances in diffusion models and processes. In particular, we have shown that many existing diffusion-based methods for generative modelling and sampling can be viewed as special instances of our proposed framework. Building on this, we have developed novel objectives for annealed flows (with connections to fluctuation theorems due to Crook and Jarzynski, rooted in statistical physics) and dynamic entropy regularised transports (based on a relationship between the EM and IPF algorithms). Finally, we have explored an instance of our proposed Schrödinger objective across a series of toy generative modelling tasks and a rare event problem where we obtain both promising and competitive results.

7. Acknowledgements

We would like to thank Jeremy Heng, Arnaud Doucet, and Valentin De Bortoli for their insightful discussions and feedback which led to the improvement of this manuscript.

References

Cédric Villani et al. *Optimal transport: old and new*, volume 338. Springer, 2009.

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

Chang Liu, Jingwei Zhuo, Pengyu Cheng, Ruiyi Zhang, and Jun Zhu. Understanding and accelerating particle-based variational inference. In *International Conference on Machine Learning*, pages 4082–4092. PMLR, 2019.

Belinda Tzen and Maxim Raginsky. Neural stochastic differential equations: Deep latent Gaussian models in the diffusion limit. *arXiv preprint arXiv:1905.09883*, 2019.

Chin-Wei Huang, Jae Hyun Lim, and Aaron C Courville. A variational perspective on diffusion-based generative models and score matching. *Advances in Neural Information Processing Systems*, 34:22863–22876, 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*, 2021. URL https://openreview.net/forum?id=PxTIG12RRHS.

Diederik P Kingma and Max Welling. Auto-encoding variational Bayes. *ICLR*, 2014.

Danilo Rezende and Shakir Mohamed. Variational inference with normalizing flows. In *International conference on machine learning*, pages 1530–1538. PMLR, 2015.

Yoshua Bengio, Salem Lahlou, Tristan Deleu, Edward J Hu, Mo Tiwari, and Emmanuel Bengio. GflowNet foundations. *arXiv preprint arXiv:2111.09266*, 2021.

Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *International Conference on Machine Learning*, pages 2256–2265. PMLR, 2015.

Hao Wu, Jonas Köhler, and Frank Noé. Stochastic normalizing flows. *Advances in Neural Information Processing Systems*, 33:5933–5944, 2020.

Xingchao Liu, Lemeng Wu, Mao Ye, et al. Let us build bridges: Understanding and extending diffusion generative models. In *NeurIPS 2022 Workshop on Score-Based Methods*, a.

Arthur P Dempster, Nan M Laird, and Donald B Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society: series B (methodological)*, 39(1):1–22, 1977.

Diederik Kingma, Tim Salimans, Ben Poole, and Jonathan Ho. Variational diffusion models. *Advances in neural information processing systems*, 34:21696–21707, 2021.

Imre Csiszár. On information-type measure of difference of probability distributions and indirect observations. *Studia Sci. Math. Hungar.*, 2:299–318, 1967.

Syed Mumtaz Ali and Samuel D Silvey. A general class of coefficients of divergence of one distribution from another. *Journal of the Royal Statistical Society: Series B (Methodological)*, 28(1):131–142, 1966.

Alfréd Rényi. On measures of entropy and information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, volume 4, pages 547–562. University of California Press, 1961.

Cédric Villani. *Topics in optimal transportation*. Number 58. American Mathematical Soc., 2003.

Danilo Jimenez Rezende, Shakir Mohamed, and Daan Wierstra. Stochastic backpropagation and approximate inference in deep generative models. In *International conference on machine learning*, pages 1278–1286. PMLR, 2014.

Carsten Hartmann, Ralf Banisch, Marco Sarich, Tomasz Badowski, and Christof Schütte. Characterization of rare events in molecular dynamics. *Entropy*, 16(1):350–376, 2013.

Xuechen Li, Ting-Kam Leonard Wong, Ricky T. Q. Chen, and David K. Duvenaud. Scalable gradients and variational inference for stochastic differential equations. In *Symposium on Advances in Approximate Bayesian Inference*, pages 1–28. PMLR, 2020.

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

UG Haussmann and E Pardoux. Time reversal of diffusion processes. In *Stochastic Differential Systems Filtering and Control*, pages 176–182. Springer, 1985.

Edward Nelson. *Dynamical theories of Brownian motion*, volume 3. Princeton university press, 1967.

A Süleyman Üstünel and Moshe Zakai. *Transformation of measure on Wiener space*. Springer Science & Business Media, 2013.

Hiroshi Kunita. *Stochastic flows and jump-diffusions*, volume 92. Springer, 2019.

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

Yang Song and Stefano Ermon. Generative modeling by estimating gradients of the data distribution. *Advances in neural information processing systems*, 32, 2019.

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

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

Julius Berner, Lorenz Richter, and Karen Ullrich. An optimal control perspective on diffusion-based generative modeling. In *NeurIPS 2022 Workshop on Score-Based Methods*, 2022.

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

Francisco Vargas, Andrius Ovsianas, David Fernandes, Mark Girolami, Neil D Lawrence, and Nikolas Nüsken. Bayesian learning via neural Schrödinger–Föllmer flows. *Statistics and Computing*, 33(1):3, 2023b.

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

Jian Huang, Yuling Jiao, Lican Kang, Xu Liao, Jin Liu, and Yanyan Liu. Schrödinger-Föllmer sampler: Sampling without ergodicity. *arXiv preprint arXiv:2106.10880*, 2021b.

Sebastian Reich and Colin Cotter. *Probabilistic forecasting and Bayesian data assimilation*. Cambridge University Press, 2015.

Nicolas Courty, Rémi Flamary, Amaury Habrard, and Alain Rakotomamonjy. Joint distribution optimal transportation for domain adaptation. *Advances in neural information processing systems*, 30, 2017.

Radford M Neal. Annealed importance sampling. *Statistics and computing*, 11:125–139, 2001.

Sebastian Reich. A dynamical systems framework for intermittent data assimilation. *BIT Numerical Mathematics*, 51(1):235–249, 2011.

Jeremy Heng, Arnaud Doucet, and Yvo Pokern. Gibbs flow for approximate transport with applications to Bayesian computation. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 83(1):156–187, 2021.

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

Michael Arbel, Alex Matthews, and Arnaud Doucet. Annealed flow transport Monte Carlo. In *International Conference on Machine Learning*, pages 318–330. PMLR, 2021.

Arnaud Doucet, Will Grathwohl, Alexander G Matthews, and Heiko Strathmann. Score-based diffusion meets annealed importance sampling. *Advances in Neural Information Processing Systems*, 35:21482–21494, 2022.

George Papamakarios, Eric Nalisnick, Danilo Jimenez Rezende, Shakir Mohamed, and Balaji Lakshminarayanan. Normalizing flows for probabilistic modeling and inference. *The Journal of Machine Learning Research*, 22(1):2617–2680, 2021.

Nikolas Nüsken and Lorenz Richter. 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.

Lorenz Richter, Ayman Boustati, Nikolas Nüsken, Francisco Ruiz, and Omer Deniz Akyildiz. Vargrad: a low-variance gradient estimator for variational inference. *Advances in Neural Information Processing Systems*, 33:13481–13492, 2020.

Kirill Neklyudov, Daniel Severo, and Alireza Makhzani. Action matching: A variational method for learning stochastic dynamics from samples. In *NeurIPS 2022 Workshop on Score-Based Methods*.

Christopher Jarzynski. Nonequilibrium equality for free energy differences. *Physical Review Letters*, 78(14):2690, 1997.

Nicolas Chopin. A sequential particle filter method for static models. *Biometrika*, 89(3):539–552, 2002.

Gavin E Crooks. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Physical Review E*, 60(3):2721, 1999.

Yongxin Chen, Tryphon T Georgiou, and Allen Tannenbaum. Stochastic control and nonequilibrium thermodynamics: Fundamental limits. *IEEE transactions on automatic control*, 65(7):2979–2991, 2019.

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

Carsten Hartmann, Christof Schütte, and Wei Zhang. Jarzynski's equality, fluctuation theorems, and variance reduction: Mathematical analysis and numerical algorithms. *Journal of Statistical Physics*, 175:1214–1261, 2019.

Gerhard Hummer and Attila Szabo. Free energy reconstruction from nonequilibrium single-molecule pulling experiments. *Proceedings of the National Academy of Sciences*, 98(7):3658–3661, 2001.

Suriyanarayanan Vaikuntanathan and Christopher Jarzynski. Escorted free energy simulations: Improving convergence by reducing dissipation. *Physical Review Letters*, 100 (19):190601, 2008.

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.

Sebastian Reich. Data assimilation: A dynamic homotopy-based coupling approach. *arXiv preprint arXiv:2209.05279*, 2022.

Wei Zhang. Some new results on relative entropy production, time reversal, and optimal control of time-inhomogeneous diffusion processes. *Journal of Mathematical Physics*, 62(4):043302, 2021.

Erwin Schrödinger. Uber die Umkehrung der Naturgesetze. *Sitzungsberichte der Preuss Akad. Wissen. Berlin, Phys. Math. Klasse*, pages 144–153, 1931.

Christian Léonard. A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete and Continuous Dynamical Systems-Series A*, 34(4):1533–1574, 2014a.

Qinsheng Zhang and Yongxin Chen. Diffusion normalizing flow. *Advances in Neural Information Processing Systems*, 34:16280–16291, 2021.

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

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

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

Ludger Ruschendorf. Convergence of the iterative proportional fitting procedure. *The Annals of Statistics*, 23(4): 1160–1174, 1995.

Sebastian Reich. Data assimilation: the Schrödinger perspective. *Acta Numerica*, 28:635–711, 2019.

Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems*, 2013.

Francisco Vargas, Pierre Thodoroff, Austen Lamacraft, and Neil Lawrence. Solving Schrödinger bridges via maximum likelihood. *Entropy*, 23(9):1134, 2021a.

David Lopes Fernandes, Francisco Vargas, Carl Henrik Ek, and Neill DF Campbell. Shooting Schrödinger's cat. In *Fourth Symposium on Advances in Approximate Bayesian Inference*, 2021.

Yuyang Shi, Valentin De Bortoli, Andrew Campbell, and Arnaud Doucet. Diffusion Schrödinger bridge matching. *arXiv preprint arXiv:2303.16852*, 2023.

Radford M Neal and Geoffrey E Hinton. A view of the EM algorithm that justifies incremental, sparse, and other variants. *Learning in graphical models*, pages 355–368, 1998.

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

Wei Zhang, Han Wang, Carsten Hartmann, Marcus Weber, and Christof Schütte. Applications of the cross-entropy method to importance sampling and optimal control of diffusions. *SIAM Journal on Scientific Computing*, 36(6): A2654–A2672, 2014.

Diederik P Kingma, Max Welling, et al. An introduction to variational autoencoders. *Foundations and Trends® in Machine Learning*, 12(4):307–392, 2019.

Yongxin Chen, Tryphon T Georgiou, and Michele Pavon. Stochastic control liaisons: Richard sinkhorn meets gaspard monge on a schrodinger bridge. *Siam Review*, 63(2): 249–313, 2021.

Maziar Raissi, Paris Perdikaris, and George E Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378:686–707, 2019.

Takeshi Koshizuka and Issei Sato. Neural Lagrangian Schrödinger bridge: Diffusion modeling for population dynamics. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=d3QNWD_pcFv.

Benjamin J Zhang and Markos A Katsoulakis. A mean-field games laboratory for generative modeling. *arXiv preprint arXiv:2304.13534*, 2023.

Stefano Peluchetti. Diffusion bridge mixture transports, Schrödinger bridge problems and generative modeling. *arXiv preprint arXiv:2304.00917*, 2023.

JMC Clark. A local characterization of reciprocal diffusions. *Applied Stochastic Analysis*, 5:45–59, 1991.

Michèle Thieullen. Reciprocal diffusions and symmetries of parabolic PDE: The nonflat case. *Potential Analysis*, 16:1–28, 2002.

Sylvie Rœlly. Reciprocal processes: a stochastic analysis approach. In *Modern stochastics and applications*, pages 53–67. Springer, 2013.

Guan-Horng Liu, Tianrong Chen, Oswin So, and Evangelos Theodorou. Deep generalized Schrödinger bridge. In *Advances in Neural Information Processing Systems*, b.

Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

Nikolas Nüsken and Lorenz Richter. Interpolating between BSDEs and PINNs: Deep learning for elliptic and parabolic boundary value problems. *Journal of Machine Learning*, 2(1):31–64, 2023.

Michael S Albergo, Nicholas M Boffi, and Eric Vanden-Eijnden. Stochastic interpolants: A unifying framework for flows and diffusions. *arXiv preprint arXiv:2303.08797*, 2023.

Nikita Gushchin, Alexander Kolesov, Alexander Korotin, Dmitry Vetrov, and Evgeny Burnaev. Entropic neural optimal transport via diffusion processes. *arXiv preprint arXiv:2211.01156*, 2022.

Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *ICLR 2015*, 2015.

Francisco Vargas, Pierre Thodoroff, Austen Lamacraft, and Neil Lawrence. Solving Schrödinger bridges via maximum likelihood. *Entropy*, 23(9), 2021b. ISSN 1099-4300. doi: 10.3390/e23091134. URL https://www.mdpi.com/1099-4300/23/9/1134.

Ioannis Karatzas, Ioannis Karatzas, Steven Shreve, and Steven E Shreve. *Brownian motion and stochastic calculus*, volume 113. Springer Science & Business Media, 1991.

Bernt Øksendal. Stochastic differential equations. In *Stochastic differential equations*, pages 65–84. Springer, 2003.

Jianfeng Zhang. *Backward stochastic differential equations*. Springer, 2017.

Tianrong Chen, Guan-Horng Liu, and Evangelos Theodorou. Likelihood training of Schrödinger bridge using forward-backward SDEs theory. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=nioAdKCEdXB.

Peter E Kloeden, Eckhard Platen, Peter E Kloeden, and Eckhard Platen. *Stochastic differential equations*. Springer, 1992.

Annie Millet, David Nualart, and Marta Sanz. Integration by parts and time reversal for diffusion processes. *The Annals of Probability*, pages 208–238, 1989.

Hans Föllmer. Time reversal on Wiener space. In *Stochastic Processes—Mathematics and Physics: Proceedings of the 1st BiBoS-Symposium held in Bielefeld, West Germany, September 10–15, 1984*, pages 119–129. Springer, 2006a.

Francesco Russo and Pierre Vallois. Itô formula for C^1-functions of semimartingales. *Probability theory and related fields*, 104:27–41, 1996.

Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293. Springer Science & Business Media, 2013.

Hans Föllmer. Calcul d'Itô sans probabilités. In *Séminaire de Probabilités XV 1979/80: Avec table générale des exposés de 1966/67 à 1978/79*, pages 143–150. Springer, 2006b.

Francesco Russo and Pierre Vallois. The generalized covariation process and Itô formula. *Stochastic Processes and their applications*, 59(1):81–104, 1995.

Hans Föllmer and Philip Protter. On Itô's formula for multidimensional Brownian motion. *Probability Theory and Related Fields*, 116:1–20, 2000.

Peter K Friz and Martin Hairer. *A course on rough paths*. Springer, 2020.

Geoffrey Roeder, Yuhuai Wu, and David Duvenaud. Sticking the landing: Simple, lower-variance gradient estimators for variational inference. *arXiv preprint arXiv:1703.09194*, 2017.

Winnie Xu, Ricky T. Q. Chen, Xuechen Li, and David Duvenaud. Infinitely deep Bayesian neural networks with stochastic differential equations. *arXiv preprint arXiv:2102.06559*, 2021.

Tarek A El Moselhy and Youssef M Marzouk. Bayesian inference with optimal maps. *Journal of Computational Physics*, 231(23):7815–7850, 2012.

Yuri Burda, Roger Grosse, and Ruslan Salakhutdinov. Importance weighted autoencoders. *arXiv preprint arXiv:1509.00519*, 2015.

Jose Hernandez-Lobato, Yingzhen Li, Mark Rowland, Thang Bui, Daniel Hernández-Lobato, and Richard Turner. Black-box alpha divergence minimization. In *International conference on machine learning*, pages 1511–1520. PMLR, 2016.

Yingzhen Li and Richard E Turner. Rényi divergence variational inference. *Advances in neural information processing systems*, 29, 2016.

Kamélia Daudel, Joe Benton, Yuyang Shi, and Arnaud Doucet. Alpha-divergence variational inference meets importance weighted auto-encoders: Methodology and asymptotics. *arXiv preprint arXiv:2210.06226*, 2022.

Yang Song, Sahaj Garg, Jiaxin Shi, and Stefano Ermon. Sliced score matching: A scalable approach to density and score estimation. In *Uncertainty in Artificial Intelligence*, pages 574–584. PMLR, 2020.

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

Mao Ye, Lemeng Wu, and Qiang Liu. First hitting diffusion models for generating manifold, graph and categorical data. In *Advances in Neural Information Processing Systems*, 2022.

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

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

Dominique Bakry, Ivan Gentil, Michel Ledoux, et al. *Analysis and geometry of Markov diffusion operators*, volume 103. Springer, 2014.

Christian Léonard. Some properties of path measures. In *Séminaire de Probabilités XLVI*, pages 207–230. Springer, 2014b.

Christian Léonard, Sylvie Rœlly, and Jean-Claude Zambrini. Reciprocal processes. a measure-theoretical point of view. 2014.

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

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33:6840–6851, 2020.

Derek Onken, Samy Wu Fung, Xingjian Li, and Lars Ruthotto. OT-flow: Fast and accurate continuous normalizing flows via optimal transport. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 9223–9232, 2021.

Will Grathwohl, Ricky T. Q. Chen, Jesse Bettencourt, and David Duvenaud. Scalable reversible generative models with free-form continuous dynamics. In *International Conference on Learning Representations*, 2019. URL https://openreview.net/forum?id=rJxgknCcK7.

A. Stochastic analysis for backward processes

In this appendix, we briefly discuss background in stochastic analysis relevant to the SDEs in (13), here repeated for convenience:

$$\mathrm{d}\boldsymbol{Y}_t = a_t(\boldsymbol{Y}_t)\,\mathrm{d}t + \sigma \overrightarrow{\mathrm{d}}\boldsymbol{W}_t, \qquad \boldsymbol{Y}_0 \sim \mu, \tag{31a}$$

$$\mathrm{d}\boldsymbol{Y}_t = b_t(\boldsymbol{Y}_t)\,\mathrm{d}t + \sigma \overleftarrow{\mathrm{d}}\boldsymbol{W}_t, \qquad \boldsymbol{Y}_T \sim \nu. \tag{31b}$$

Recall that the forward Itô differential $\overrightarrow{\mathrm{d}}$ in (31a) is far more commonly denoted simply[7] by d, and theory for the forward SDE (31a) is widely known (Karatzas et al., 1991; Øksendal, 2003). In contrast, reverse-time SDEs of the form (31b) are less common and there are fewer textbook accounts of their interactions with forward SDEs. We highlight Kunita (2019) for an in-depth treatment, and alert the reader to the fact that 'backward stochastic differential equations' as discussed in Zhang (2017); Chen et al. (2022), for instance, are largely unrelated. We therefore refer to (31b) as a 'reverse-time' SDE.

Remark 10 (Notation)**.** We deliberately depart from some of the notation employed in the recent literature (see, for instance, Huang et al. (2021a); Liu et al. (a)) by using $\boldsymbol{Y}_t$ in both (31a) and (31b), and not introducing an auxiliary process capturing the reverse-time dynamics. From a technical perspective, this is justified since $(\boldsymbol{Y}_t)_{0\le t\le T}$ merely represents a generic element in path space, and full information is encoded in the path measures $\mathbb{Q}^{\mu,a} \equiv \overrightarrow{\mathbb{P}}^{\mu,a}$ and $\mathbb{P}^{\nu,b} \equiv \overleftarrow{\mathbb{P}}^{\nu,b}$. Importantly, placing (31a) and (31b) on an equal footing seems essential for a convenient formulation of Proposition 2.2. Slightly departing from the VAE-inspired notation from Section 2.1, we equivalently refer to these path measures by $\overrightarrow{\mathbb{P}}^{\mu,a}$ and $\overleftarrow{\mathbb{P}}^{\nu,b}$, highlighting the symmetry of the setting in (31).

Intuitively, (31) can be viewed as continuous time limits of the Markov chains defined in (10), or, in other words, the Markov chains (10) are the Euler-Maruyama discretisations for (31), see Kloeden et al. (1992, Section 9.1). Throughout, we impose the following:

Assumption A.1 (Smoothness and linear growth of vector fields)**.** All (time-dependent) vector fields in this paper belong to the set

$$\mathcal{U} := \Big\{ a \in C^\infty([0,T]\times\mathbb{R}^d;\mathbb{R}^d) : \quad \text{there exists a constant } C>0 \\ \text{such that } \|a_t(\boldsymbol{x}) - a_t(\boldsymbol{y})\| \le C\|\boldsymbol{x}-\boldsymbol{y}\|, \text{ for all } t\in[0,T],\ \boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^d \Big\}.$$

The preceding assumption guarantees existence and uniqueness for (31a) and (31b), and it allows us to use Girsanov's theorem in the proof of Proposition 2.2 (Novikov's condition can be shown to be satisfied, see Øksendal (2003, Section 8.6)). Furthermore, Assumption A.1 is sufficient to conclude Nelson's relation (Proposition 2.1), see Haussmann and Pardoux (1985); Millet et al. (1989); Föllmer (2006a) and the discussion in Russo and Vallois (1996). Having said all that, it is possible to substantially weaken Assumption A.1 with more technical effort. Moreover, we can replace the constant $\sigma>0$ by $\sigma : [0,T]\times\mathbb{R}^d \to \mathbb{R}^{d\times d}$ throughout, assuming sufficient regularity, growth and invertibility properties, and amending the formulas accordingly.

The precise meaning of (31) is given by the integrated formulations

$$\boldsymbol{Y}_t = \boldsymbol{Y}_0 + \int_0^t a_s(\boldsymbol{Y}_s)\,\mathrm{d}s + \int_0^t \sigma \overrightarrow{\mathrm{d}}\boldsymbol{W}_s, \qquad \boldsymbol{Y}_0 \sim \mu, \tag{33a}$$

$$\boldsymbol{Y}_t = \boldsymbol{Y}_T - \int_t^T b_s(\boldsymbol{Y}_s)\,\mathrm{d}s - \int_t^T \sigma \overleftarrow{\mathrm{d}}\boldsymbol{W}_s, \qquad \boldsymbol{Y}_T \sim \nu, \tag{33b}$$

where the forward and backward integrals need defining. Roughly speaking, we have

$$\int_{t_0}^{t_1} \boldsymbol{X}_s \cdot \overrightarrow{\mathrm{d}}\boldsymbol{Z}_s = \lim_{`\Delta t\to 0'} \sum_i \boldsymbol{X}_{t_i}\cdot(\boldsymbol{Z}_{t_{i+1}} - \boldsymbol{Z}_{t_i}), \tag{34a}$$

[7] ...but in this paper we stick to the notation $\overrightarrow{\mathrm{d}}$ to emphasise the symmetry of the setting.

$$\int_{t_0}^{t_1} \boldsymbol{X}_s \cdot \overleftarrow{\mathrm{d}}\, \boldsymbol{Z}_s = \lim_{`\Delta t \to 0'} \sum_i \boldsymbol{X}_{t_{i+1}} \cdot (\boldsymbol{Z}_{t_{i+1}} - \boldsymbol{Z}_{t_i}), \tag{34b}$$

see Remark 3, for 'appropriate' processes $(\boldsymbol{X}_t)_{0\le t\le T}$ and $(\boldsymbol{Z}_t)_{0\le t\le T}$, and where the limit $\Delta t \to 0$ of vanishing step sizes needs careful analysis (see Remark 11 below). The most salient difference between (31a) and (31b) is the fact that $\boldsymbol{X}_{t_i}$ is replaced by $\boldsymbol{X}_{t_{i+1}}$ in (34b).

Remark 11 (Convergence of the limits in (34))**.** If we only assume that $\boldsymbol{X}, \boldsymbol{Z} \in C([t_0, t_1]; \mathbb{R}^d)$, possibly pathwise, that is, deterministically, then the limits in (34) might not exist, or when they do, their values might depends on the specific sequence of mesh refinements. The following approaches are available to make the definitions (34) rigorous:

1. Itô calculus (see, for example, Revuz and Yor (2013, Chapter 9)) uses adaptedness and semimartingale properties for the forward integral in (34a), but note that the definition is not pathwise (that is, the limit (34a) is defined up to a set of measure zero). For the backward integrals in (34b) and, importantly for us, in (57), it can then be shown that the relevant processes are (continuous) reverse-time martingales (see Kunita (2019) for a discussion of the corresponding filtrations). The latter property is guaranteed under Assumption A.1, see the discussion around Theorem 2.3 in Russo and Vallois (1996).

2. Föllmer's 'Itô calculus without probabilities' (Föllmer, 2006b) is convenient, since it allows to us to perform calculations using (31) and Proposition 2.2 without introducing filtrations and related stochastic machinery. The caveat is that the results may in principle depend on the sequence of mesh refinements, but under Assumption A.1, those differences only appear on a set of measure zero, see Russo and Vallois (1995); Föllmer and Protter (2000).

3. Similarly, the integrals in (34) can be defined in a pathwise fashion using rough path techniques, see Friz and Hairer (2020, Section 5.4).

For the current paper, the following conversion formulas are crucial,

$$\int_0^t \boldsymbol{X}_s \cdot \overleftarrow{\mathrm{d}}\, \boldsymbol{Z}_s - \int_0^t \boldsymbol{X}_s \cdot \overrightarrow{\mathrm{d}}\, \boldsymbol{Z}_s = \langle \boldsymbol{X}, \boldsymbol{Z} \rangle_t, \tag{35a}$$

$$\int_0^t \boldsymbol{X}_s \cdot \overleftarrow{\mathrm{d}}\, \boldsymbol{Z}_s + \int_0^t \boldsymbol{X}_s \cdot \overrightarrow{\mathrm{d}}\, \boldsymbol{Z}_s = 2 \int_0^t \boldsymbol{X}_s \circ \boldsymbol{Z}_s, \tag{35b}$$

where $\langle \boldsymbol{X}, \boldsymbol{Z} \rangle$ is the quadratic variation process (if defined, see Russo and Vallois (1995)), and $\circ$ denotes Stratonovich integration. For solutions to (31), we obtain (16) from (35a). In particular, we can often trade backward integrals for divergence terms (see Appendix C.1), using the (backward) martingale properties

$$\mathbb{E}\left[\int_0^t f_t(\boldsymbol{Y}_t) \cdot \overrightarrow{\mathrm{d}}\, \boldsymbol{W}_t\right] = 0, \quad \text{if } (\boldsymbol{Y}_t)_{0\le t\le T} \text{ solves (31}a\text{)}, \tag{36a}$$

$$\mathbb{E}\left[\int_t^T f_t(\boldsymbol{Y}_t) \cdot \overleftarrow{\mathrm{d}}\, \boldsymbol{W}_t\right] = 0, \quad \text{if } (\boldsymbol{Y}_t)_{0\le t\le T} \text{ solves (31}b\text{)}. \tag{36b}$$

B. Variational inference and divergences

Various concepts well-known in the variational inference community have direct counterparts in the diffusion setting. In this appendix we review a few that are directly relevant to this paper.

Maximum likelihood. The traditional approach (Blei et al., 2017; Kingma et al., 2019) towards the *evidence lower bound* (ELBO) in (7) is via maximum likelihood in latent variable models. Using the notation and set-up from the introduction, one can show using Jensen's inequality (or dual representations of the KL divergence), that

$$\ln\left(\int p_\theta(\boldsymbol{x}, \boldsymbol{z})\, \mathrm{d}\boldsymbol{z}\right) = \ln p_\theta(\boldsymbol{x}) \ge \mathrm{ELBO}_{\boldsymbol{x}}(\phi, \theta), \tag{37}$$

with equality if and only if $q_\phi(\boldsymbol{z}|\boldsymbol{x}) = p_\theta(\boldsymbol{z}|\boldsymbol{x})$. The bound in (37) motivates maximising the (tractable) right-hand side, performing model selection (according to Bayesian evidence) and posterior approximation (in terms of the variational family

$q_\phi(\boldsymbol{z}|\boldsymbol{x})$) at the same time. The calculation in (7) shows that this objective can equivalently be derived from Framework 1 and connected to the KL divergence between the joint distributions $q_\phi(\boldsymbol{x}, \boldsymbol{z})$ and $p_\theta(\boldsymbol{x}, \boldsymbol{z})$.

Reparameterisation trick (Kingma and Welling, 2014; Rezende et al., 2014). For optimising $\mathrm{ELBO}_{\boldsymbol{x}}(\phi, \theta)$, it is crucial to select efficient low-variance gradient estimators. In this context, it has been observed that reparameterising $\boldsymbol{z} \sim q_\phi(\boldsymbol{z}|\boldsymbol{x})$ in the form $\boldsymbol{z} = g(\epsilon, \phi, \boldsymbol{x})$, see Kingma et al. (2019, Section 2.4.1), substantially stabilises the training procedure. Here, ϵ is an auxiliary random variable with tractable 'base distribution' that is independent of ϕ and $\boldsymbol{x}$, and g is a deterministic function (transforming ϵ into $\boldsymbol{z}$), parameterised by ϕ and $\boldsymbol{x}$. We would like to point out that many (although not all, see below) objectives in diffusion modelling are already reparameterised, since the SDEs (13) transform the 'auxiliary' variables $(\boldsymbol{W}_t)_{0\leq t\leq T}$ into $(\boldsymbol{Y}_t)_{0\leq t\leq T}$. With this viewpoint, the vector field a_t corresponds to the parameter ϕ, $(\boldsymbol{W}_t)_{0\leq t\leq T}$ corresponds to ϵ, and g corresponds to the solution map associated to the SDE (13a), sometimes referred to as the Itô map. In this sense, the objectives (29), (44) and (45) are reparameterised, but (19) is not if the gradients are detached according to Remark 5. We mention in passing that *sticking the landing* (Roeder et al., 2017) offers a further variance reduction close to optimality, and that the same method can be employed for diffusion objectives, see Vargas et al. (2023b); Xu et al. (2021).

Reinforce gradient estimators. As an alternative to the KL-divergence, Nüsken and Richter (2021) investigated the family of *log-variance divergences*

$$D^u_{\mathrm{Var}}(q||p) = \mathrm{Var}_{\boldsymbol{x}\sim u}\left(\ln \frac{\mathrm{d}q}{\mathrm{d}p}(\boldsymbol{x})\right), \tag{38}$$

parameterised by an auxiliary distribution u, in order to connect variational inference to backward stochastic differential equations (Zhang, 2017). The fact that gradients of (38) do not have to be taken with respect to $\boldsymbol{x}$ (see Remark 5) reduces the computational cost and provides additional flexibility in the choice of u, but the gradient estimates potentially suffer from higher variance since the reparameterisation trick is not available. The latter drawback is alleviated somewhat by the fact that particular choices of u can be linked to control variate enhanced reinforce gradient estimators (Richter et al., 2020) that are particularly useful when reparameterisation is not available (such as in discrete models). We note that the same divergence has also been used as a variational inference objective in El Moselhy and Marzouk (2012).

Importance weighted autoencoders (IWAE). Burda et al. (2015) have developed a multi-sample version of $\mathrm{ELBO}_{\boldsymbol{x}}(\phi, \theta)$ that achieves a tighter lower bound on the marginal log-likelihood in (37). To develop similar objectives in a diffusion setting, we observe that for each $K \geq 1$,

$$D^{(K)}_{\mathrm{KL}}(q||p) = \mathbb{E}_{x_1,\ldots,x_K \overset{iid}{\sim} q}\left[\ln\left(\frac{1}{K}\sum_{i=1}^{K}\frac{\mathrm{d}q}{\mathrm{d}p}(x_i)\right)\right] \tag{39}$$

defines a generalised KL divergence[8] that reproduces the IWAE lower bound as per Framework 1, in the sense of equation (7). To the best of our knowledge, the precise formulation in (39) is new, but similar to the previous works Hernandez-Lobato et al. (2016); Li and Turner (2016); Daudel et al. (2022). We exhibit an example of (39) applied in a diffusion context in Section C.3, see Remark 12.

C. Connections to previous work

C.1. Discussion of equivalent expressions for $D_{\mathrm{KL}}(\overrightarrow{\mathbb{P}}^{\mu,a}||\overleftarrow{\mathbb{P}}^{\nu,b})$

Notice that we can realise samples from $\overleftarrow{\mathbb{P}}^{\nu,b}$ both via the reverse-time SDE in (13b) or via its time reversal given by the following forward SDE (Nelson, 1967; Anderson, 1982; Haussmann and Pardoux, 1985):

$$\mathrm{d}\widehat{\boldsymbol{Y}}_t = \left(b_{T-t}(\widehat{\boldsymbol{Y}}_t) - \sigma^2\nabla \ln \overleftarrow{\rho}^{\nu,b}_{T-t}(\widehat{\boldsymbol{Y}}_t)\right)\mathrm{d}t + \sigma\,\overrightarrow{\mathrm{d}}\boldsymbol{W}_t, \qquad \widehat{\boldsymbol{Y}}_0 \sim \nu, \tag{40}$$

using $\widehat{\boldsymbol{Y}}_t := \widehat{\boldsymbol{Y}}_{T-t}$. This allows us to obtain an expression for $D_{\mathrm{KL}}(\overrightarrow{\mathbb{P}}^{\mu,a}|\overleftarrow{\mathbb{P}}^{\nu,b})$ via Girsanov's theorem:

$$D_{\mathrm{KL}}(\overrightarrow{\mathbb{P}}^{\mu,a}||\overleftarrow{\mathbb{P}}^{\nu,b}) = D_{\mathrm{KL}}(\overleftarrow{\rho}^{\nu,b}_0||\nu) + \mathbb{E}\left[\tfrac{1}{2\sigma^2}\int_0^T \left|\left|a_t(\boldsymbol{Y}_t) - \left(b_t(\boldsymbol{Y}_t) - \sigma^2\nabla \ln \overleftarrow{\rho}^{\nu,b}_t(\boldsymbol{Y}_t)\right)\right|\right|^2 \mathrm{d}t\right]. \tag{41}$$

[8] Indeed, by Jensen's inequality, we have that $D^{(K+1)}_{\mathrm{KL}}(q||p) \geq D^{(K)}_{\mathrm{KL}}(q||p)$, so that in particular $q \neq p$ implies $D^{(K)}_{\mathrm{KL}}(q||p) > 0$.

However there are several terms here that we cannot estimate or realise in a tractable manner, one being the score $\nabla \ln \overleftarrow{\rho}_t^{\nu,b}$ and the other being sampling from the distribution $\overleftarrow{\rho}_0^{\nu,b}$. [9]

In order to circumvent the score term, the authors Vargas et al. (2021b); Chen et al. (2022) use the Fokker-Plank (FPK) equation and integration by parts, respectively, trading of the score with a divergence term, whilst Huang et al. (2021a) use a variant of the Feynman Kac formula to arrive at an equivalent solution. From Proposition 2.2, we can avoid the divergence entirely and replace it by a backwards integral (making use of the conversion formula (16) and the fact that the ensuing forward integral is zero in expectation). As hinted at in Remark 3, this replacement might have favourable variance-reducing properties, but numerical evidence would be necessary.

C.2. Score-based generative modeling

Generative modeling is concerned with the scenario where $\mu(\boldsymbol{x})$ can be sampled from (but its density is unknown), and the goal is to learn a backward diffusion as in (13b) that allows us to generate further samples from $\mu(\boldsymbol{x})$, see Song et al. (2021). We may fix a reference forward drift a_t, and, motivated by Proposition 2.1, parameterise the backward drift as $b_t = a_t - s_t$, so that in the case when $\overrightarrow{\mathbb{P}}^{\mu,a} = \overleftarrow{\mathbb{P}}^{\nu,b}$, the variable drift component s_t will represent the score $\sigma^2 \nabla \ln \rho_t^{\mu,a}$. When the diffusion associated to a_t is ergodic and T is large, $\overrightarrow{\mathbb{P}}^{\mu,a} = \overleftarrow{\mathbb{P}}^{\nu,b}$ requires that $\nu(\boldsymbol{z})$ is close to the corresponding invariant measure. Choosing $\gamma_t^- = a_t$, and, for simplicity $\sigma = 1$, direct calculations using Proposition 2.2 show that

$$\mathcal{L}_{\mathrm{ISM}}(s) := D_{\mathrm{KL}}(\overrightarrow{\mathbb{P}}^{\mu,a} || \overleftarrow{\mathbb{P}}^{\nu,a-s}) = \mathbb{E}_{\boldsymbol{Y} \sim \overrightarrow{\mathbb{P}}^{\mu,a}} \left[\tfrac{1}{2} \int_0^T s_t^2(\boldsymbol{Y}_t) \, \mathrm{d}t + \int_0^T (\nabla \cdot s_t)(\boldsymbol{Y}_t) \, \mathrm{d}t \right] + \text{const.} \tag{42}$$

recovers the implicit score matching objective (Hyvärinen and Dayan, 2005), up to a constant that does not depend on s_t.

Proof. We start by noticing that the contributions in (15a) and (15b) do not depend on s_t, and can therefore be absorbed in the constant in (42) Notice that the precise forms of Γ_0, Γ_T and γ^+ are left unspecified or unknown, but this does not affect the argument. We find

$$\begin{aligned} D_{\mathrm{KL}}(\overrightarrow{\mathbb{P}}^{\mu,a} || \overleftarrow{\mathbb{P}}^{\nu,a-s}) &= \mathbb{E}_{\boldsymbol{Y} \sim \overrightarrow{\mathbb{P}}^{\mu,a}} \left[\int_0^T s_t(\boldsymbol{Y}_t) \cdot \left(\overleftarrow{\mathrm{d}} \boldsymbol{Y}_t - \tfrac{1}{2} \left(2a_t - s_t \right) (\boldsymbol{Y}_t) \, \mathrm{d}t \right) \right] + \text{const.} \\ &= \mathbb{E} \left[\int_0^T s_t(\boldsymbol{Y}_t) \cdot \left(\sigma \overleftarrow{\mathrm{d}} \boldsymbol{W}_t + \tfrac{1}{2} s_t(\boldsymbol{Y}_t) \, \mathrm{d}t \right) \right] + \text{const.} = \mathbb{E} \left[\tfrac{1}{2} \int_0^T s_t^2(\boldsymbol{Y}_t) \, \mathrm{d}t + \int_0^T (\nabla \cdot s_t)(\boldsymbol{Y}_t) \, \mathrm{d}t \right] + \text{const.}, \end{aligned}$$

where in the first line we use Proposition 2.2 together with $b_t = a_t - s_t$ and $\gamma_t^- = a_t$, and to proceed to the second line we substitute $\overleftarrow{\mathrm{d}} \boldsymbol{Y}_t$ using the SDE in (13a). The last equality follows from the conversion formula between forward and backward Itô integrals, see (16), and the fact that forward integrals with respect to Brownian motion have zero (forward) expectation, see (36a). □

Notice that the nonuniqueness in Framework 1′ has been circumvented by fixing the forward drift a_t; indeed $\mathcal{L}_{\mathrm{ISM}}$ is convex in s, confirming Note that using integration by parts, $\mathcal{L}_{\mathrm{ISM}}$ is equivalent to denoising score matching (Song et al., 2020; 2021):

$$D_{\mathrm{KL}}(\overrightarrow{\mathbb{P}}^{\mu,a} || \overleftarrow{\mathbb{P}}^{\nu,a-s}) = \mathbb{E}_{\boldsymbol{Y} \sim \overrightarrow{\mathbb{P}}^{\mu,a}} \left[\tfrac{1}{2\sigma^2} \int_0^T \left\| s_t(\boldsymbol{Y}_t) - \nabla \ln \rho_{t|0}^{\mu,a}(\boldsymbol{Y}_t | \boldsymbol{Y}_0) \right\|^2 \mathrm{d}t \right] + \text{const.}. \tag{43}$$

Framework 1′ accommodates modifications of (42); in particular the divergence term in (42) can be replaced by a backward integral, see Appendix C.1 and Remark 3. Note that the settings discussed in this section are also akin to the formulations in Kingma et al. (2021); Huang et al. (2021a).

Finally, it is worth highlighting that this setting is not limited to ergodic models and can in fact accommodate finite time models in the exact same fashion as the Föllmer drift is used for sampling (Section C.3) by using a Doob's transform (Rogers and Williams, 2000) based SDE for $\overrightarrow{\mathbb{P}}^{\mu,a}$ as opposed to the classical VP-SDE see Example 2.4 in Ye et al. (2022).

[9] When $\widehat{\boldsymbol{Y}}_t$ is an OU process and μ is Gaussian we are in the traditional DDPM setting (Song et al., 2021) and these two quantities admit the classical tractable score matching approximations

C.3. Score-based sampling

Consider the setting when $\nu(\boldsymbol{z})$ is a target distribution that can be evaluated pointwise up to a normalisation constant. In order to construct a diffusion process that transports an appropriate auxiliary distribution $\mu(\boldsymbol{x})$ to $\nu(\boldsymbol{z})$, one approach is to fix a drift b_t in the backward diffusion (13b), and then learn the corresponding forward diffusion (13a) by minimising $a \mapsto D(\overrightarrow{\mathbb{P}}^{\mu,a} | \overleftarrow{\mathbb{P}}^{\nu,b})$. Tractability of this objective requires that $\mu := \overleftarrow{\mathbb{P}}_0^{\nu,b}$ be known explicitly, at least approximately. In the following we review possible choices.

The Föllmer drift. Choosing $b_t(x) = x/t$, one can show using Doob's transform (Rogers and Williams, 2000, Theorem 40.3(iii)), that $\overleftarrow{\mathbb{P}}_0^{\nu,b}(\boldsymbol{x}) = \delta(\boldsymbol{x})$, for any terminal distribution $\nu(\boldsymbol{z})$. Hence, minimising $a \mapsto D_{\mathrm{KL}}(\overrightarrow{\mathbb{P}}^{\delta_0,a} | \overleftarrow{\mathbb{P}}^{\nu,b})$ leads to a tractable objective. In particular consider the choice $\Gamma_0 = \delta_0$, $\gamma^+ = 0$, corresponding to a standard Brownian motion, then it follows that $\gamma^- = \frac{x}{t}$, $\Gamma_T = \mathcal{N}(0, T\sigma^2)$ and thus via Proposition 2.2:

$$D_{\mathrm{KL}}(\overrightarrow{\mathbb{P}}^{\delta_0,a} | \overleftarrow{\mathbb{P}}^{\nu,b}) = \mathbb{E}_{\boldsymbol{Y} \sim \overrightarrow{\mathbb{P}}^{\mu,a}} \left[\frac{1}{\sigma^2} \int_0^T a^2(\boldsymbol{Y}_t) \, \mathrm{d}t + \ln \left(\frac{\mathrm{d}\mathcal{N}(0, T\sigma^2)}{\mathrm{d}\nu} \right) (\boldsymbol{Y}_T) \right] + \text{const.}, \tag{44}$$

in accordance with (Dai Pra, 1991; Vargas et al., 2023b; Zhang and Chen, 2022). For further details, see Follmer (1984); Vargas et al. (2023b); Zhang and Chen (2022); Huang et al. (2021b). As hinted at in Appendix B, replacing D_{KL} in (44) by the log-variance divergence (38) leads to an objective that directly links to BSDEs, see (Nüsken and Richter, 2021, Section 3.2).

Ergodic diffusions. Vargas et al. (2023a); Berner et al. (2022) fix a backward drift b_t that induces an ergodic backward diffusion, so that for large T, the marginal at initial time $\overleftarrow{\mathbb{P}}_{t=0}^{\nu,b}$ is close to the corresponding invariant distribution, and in particular (almost) independent of $\nu(\boldsymbol{z})$.[10] Defining $\mu := \overleftarrow{\mathbb{P}}_{t=0}^{\nu,b}$, Vargas et al. (2023a); Berner et al. (2022) set out to minimise the denoising diffusion sampler loss $\mathcal{L}_{\mathrm{DDS}}(f) := D_{\mathrm{KL}}(\overrightarrow{\mathbb{P}}^{\mu,b+\sigma^2 f} | \overleftarrow{\mathbb{P}}^{\nu,b})$. Choosing the reference process to be $\Gamma_{0,T} = \mu$, $\gamma^\pm = b$ (that is, the reference process is at stationarity, with invariant measure $\mu(\boldsymbol{z})$), direct calculation based on (15) shows that

$$\mathcal{L}_{\mathrm{DDS}}(f) = \mathbb{E}_{\boldsymbol{Y} \sim \overrightarrow{\mathbb{P}}^{\mu,b+\sigma^2 f}} \left[\sigma^2 \int_0^T f^2(\boldsymbol{Y}_t) \, \mathrm{d}t + \ln \left(\frac{\mathrm{d}\Gamma_T}{\mathrm{d}\nu} \right) (\boldsymbol{Y}_T) \right], \tag{45}$$

Remark 12 (IWAE-objective)**.** In line with (39), we may also consider the multi-sample objective

$$\begin{aligned} \mathcal{L}_{\mathrm{DDS}}^{(K)}(f) &:= D_{\mathrm{KL}}^{(K)}(\overrightarrow{\mathbb{P}}^{\mu,b+\sigma^2 f} | \overleftarrow{\mathbb{P}}^{\nu,b}) \\ &= \mathbb{E}_{\boldsymbol{Y}^1, \dots, \boldsymbol{Y}^K \overset{iid}{\sim} \overrightarrow{\mathbb{P}}^{\mu,b+\sigma^2 f}} \left[\ln \left(\frac{1}{K} \sum_{i=1}^K \exp \left(\sigma^2 \int_0^T f^2(\boldsymbol{Y}_t^i) \, \mathrm{d}t + \ln \left(\frac{\mathrm{d}\Gamma_T}{\mathrm{d}\nu} \right) (\boldsymbol{Y}_T^i) \right) \right) \right] \end{aligned}$$

Proof. We start by noticing that the choice $\gamma_t^- = b_t$ cancels the terms in (15c), and the choice $\Gamma_0 = \mu$ cancels the first term in (15a). Using $a_t = b_t + \sigma^2 f_t$, we therefore obtain

$$\begin{aligned} \mathcal{L}_{\mathrm{DDS}}(f) = D_{\mathrm{KL}}(\overrightarrow{\mathbb{P}}^{\mu,b+\sigma^2 f} | \overleftarrow{\mathbb{P}}^{\nu,b}) &= \mathbb{E} \left[\sigma^2 \int_0^T f_t(\boldsymbol{Y}_t) \cdot \big((b_t + f_t)(\boldsymbol{Y}_t) \, \mathrm{d}t - \tfrac{1}{2}(2b_t + f_t)(\boldsymbol{Y}_t) \, \mathrm{d}t \big) + \ln \left(\frac{\mathrm{d}\Gamma_T}{\mathrm{d}\nu} \right) (\boldsymbol{Y}_T) \right] \\ &= \mathbb{E} \left[\sigma^2 \int_0^T f_t^2(\boldsymbol{Y}_t) \, \mathrm{d}t + \ln \left(\frac{\mathrm{d}\Gamma_T}{\mathrm{d}\nu} \right) (\boldsymbol{Y}_T) \right]. \end{aligned} \tag{46a}$$

As is implicit in Berner et al. (2022), it is also possible to choose $\gamma^\pm = 0$ for the reference process, with $\Gamma_0 = \Gamma_T = \mathrm{Leb}$, the Lebesgue measure on $\mathbb{R}^d$. We notice in passing that although the Lebesgue measure is not normalisable, it is invariant under Brownian motion (the forward and backward drifts are both zero), and the arguments can be made rigorous by a

[10] Vargas et al. (2023a) chose a (time-inhomoegenous) backward Ornstein-Uhlenbeck process, so that $\overleftarrow{\mathbb{P}}_{t=0}^{\nu,b}$ is close to a Gaussian, but generalisations are straightforward.

limiting argument (take Gaussians with diverging variances), or by using the techniques in Léonard (2014a, Appendix A.1). By similar calculations as above, we obtain

$$\mathcal{L}_{\mathrm{DDS}}(f) = \mathbb{E}\left[\sigma^2 \int_0^T f_t^2(\boldsymbol{Y}_t)\,\mathrm{d}t - \tfrac{1}{\sigma}\int_0^T b_t(\boldsymbol{Y}_t)\cdot \overleftarrow{\mathrm{d}}\boldsymbol{W}_t + \ln\mu(\boldsymbol{Y}_0) - \ln\nu(\boldsymbol{Y}_T)\right] \tag{47a}$$

$$= \mathbb{E}\left[\sigma^2 \int_0^T f_t^2(\boldsymbol{Y}_t)\,\mathrm{d}t - \int_0^T (\nabla\cdot b_t)(\boldsymbol{Y}_t)\,\mathrm{d}t - \ln\nu(\boldsymbol{Y}_T)\right] + \text{const.}, \tag{47b}$$

where we overload notation and denote the Lebesgue densities of μ and ν with the same letters. In the second line we have used the conversion formula in (16), together with the fact that the forward Itô integrals are forward martingales (Kunita, 2019), and therefore have zero expectation. Comparing (45) and (47b), we notice the additional divergence term, due to the fact that the choice $\gamma^- = 0$ does not cancel the terms in (57). See also the discussion in Appendix C.1. □

Finally we note that whilst the work in Berner et al. (2022) focuses on exploring a VP-SDE-based approach which is ergodic, their overarching framework generalises beyond ergodic settings, notice this objective is akin to the KL expressions in Vargas (2021, Proposition 1) and Liu et al. (b, Proposition 9).

C.4. Action matching (Neklyudov et al.)

Similar to our approach in Section 3.1, Neklyudov et al. fix a curve of distributions $(\pi_t)_{t\in[0,T]}$. In contrast to us, they assume that samples from π_t are available, for each $t \in [0,T]$ (but scores and unnormalised densities are not). Still, we can use Framework 1′ to rederive their objective:

Akin to the proof of Proposition 3.1, under mild conditions on $(\pi_t)_{t\in[0,T]}$, there exists a unique vector field $\nabla\phi_t^*$ that satisfies the Fokker-Planck equation

$$\partial_t \pi_t + \nabla\cdot(\pi_t \nabla\phi_t^*) = \tfrac{\sigma^2}{2}\Delta\pi_t. \tag{48}$$

We can now use the reference process $\overrightarrow{\mathbb{P}}^{\pi_0,\nabla\phi^*}$ (that is, $\Gamma_0 = \pi_0$, $\gamma_t^+ = \nabla\phi_t^*$, $\Gamma_T = \pi_T$, $\gamma_t^- = \nabla\phi_t^* - \sigma^2\nabla\ln\pi_t$) to compute the objective

$$\psi \mapsto D_{\mathrm{KL}}(\overrightarrow{\mathbb{P}}^{\pi_0,\nabla\psi} || \overleftarrow{\mathbb{P}}^{\pi_T,\nabla\psi-\sigma^2\nabla\ln\pi}),$$

relying on the same calculational techniques as in Sections C.2 and C.3 (the particular choice of reference process cancels the terms in (15a)). Notice that the parameterisation in this objective constrains the target diffusion to have time-marginals π_t, just as in Section 3.1. By direct calculation, we obtain (up to a factor of $2/\sigma^2$) the action-gap in equation (5) in Neklyudov et al.. Indeed, we see that

$$D_{\mathrm{KL}}(\overrightarrow{\mathbb{P}}^{\pi_0,\nabla\psi} || \overleftarrow{\mathbb{P}}^{\pi_T,\nabla\psi-\sigma^2\nabla\ln\pi}) = \mathbb{E}_{\overrightarrow{\mathbb{P}}^{\pi_0,\nabla\psi}}\left[\ln\left(\frac{\mathrm{d}\overrightarrow{\mathbb{P}}^{\pi_0,\nabla\psi}}{\mathrm{d}\overleftarrow{\mathbb{P}}^{\nabla\psi-\sigma^2\nabla\ln\pi}}\right)\right]$$

$$= \mathbb{E}\left[\tfrac{1}{\sigma^2}\int_0^T (\nabla\psi_t - \nabla\phi_t^*)^2(\boldsymbol{Y}_t)\,\mathrm{d}t - \tfrac{1}{\sigma}\int_0^T (\nabla\psi_t - \nabla\phi_t^*)(\boldsymbol{Y}_t)\cdot\overleftarrow{\mathrm{d}}\boldsymbol{W}_t - \int_0^T \nabla\ln\pi_t(\boldsymbol{Y}_t)\cdot(\nabla\psi_t - \nabla\phi_t^*)(\boldsymbol{Y}_t)\,\mathrm{d}t\right]$$

$$= \mathbb{E}\left[\tfrac{1}{\sigma^2}\int_0^T (\nabla\psi_t - \nabla\phi_t^*)^2(\boldsymbol{Y}_t)\,\mathrm{d}t\right],$$

where in the last line we have used the conversion formula (16) together with (36a) to compute

$$\mathbb{E}\left[\tfrac{1}{\sigma}\int_0^T (\nabla\psi_t - \nabla\phi_t^*)(\boldsymbol{Y}_t)\cdot\overleftarrow{\mathrm{d}}\boldsymbol{W}_t\right] = \mathbb{E}\left[\int_0^T (\nabla\cdot(\nabla\psi_t - \nabla\phi_t^*))(\boldsymbol{Y}_t)\,\mathrm{d}t\right]$$

$$= \int_0^T\int_{\mathbb{R}^d} (\nabla\cdot(\nabla\psi_t - \nabla\phi_t^*))(\boldsymbol{x})\pi_t(\mathrm{d}\boldsymbol{x})\,\mathrm{d}t = -\int_0^T\int_{\mathbb{R}^d} (\nabla\psi_t - \nabla\phi_t^*)(\boldsymbol{x})\cdot\nabla\ln\pi_t(\boldsymbol{x})\pi_t(\mathrm{d}\boldsymbol{x})\,\mathrm{d}t$$

$$= \mathbb{E}\left[\int_0^T \nabla\ln\pi_t(\boldsymbol{Y}_t)\cdot(\nabla\psi_t - \nabla\phi_t^*)(\boldsymbol{Y}_t)\,\mathrm{d}t\right]$$

and cancel the two last terms in the penultimate line.

D. Score-based annealing (Section 3.1)

D.1. Derivation of the RND in equation (19)

In this section, we first verify the expression in (19), using Proposition 2.2, and choosing $\Gamma_0 = \Gamma_T$ to be the Lebesgue measure, $\gamma^+ = \gamma^- = 0$. We recall that although the Lebesgue measure in not normalisable, the arguments can be made rigorous using the techniques in Léonard (2014a, Appendix A).

The Radon-Nikodym derivative (RND) along (17) reads

$$\begin{aligned}
&\left(\ln \frac{\mathrm{d}\overrightarrow{\mathbb{P}}_{\pi_0,\sigma^2\nabla\ln\pi+\nabla\phi}}{\mathrm{d}\overleftarrow{\mathbb{P}}_{\pi_T,-\sigma^2\nabla\ln\pi+\nabla\phi}}\right)(\boldsymbol{Y}) = (\ln\pi_0)(\boldsymbol{Y}_0) - (\ln\pi_T)(\boldsymbol{Y}_T)\\
&+ \tfrac{1}{2\sigma^2}\int_0^T (\sigma^2\nabla\ln\pi_t + \nabla\phi_t)(\boldsymbol{Y}_t)\left((\sigma^2\nabla\ln\pi_t + \nabla\phi_t)(\boldsymbol{Y}_t)\,\mathrm{d}t + \sqrt{2}\sigma\,\overrightarrow{d}\,\boldsymbol{W}_t - \tfrac{1}{2}(\sigma^2\nabla\ln\pi_t + \nabla\phi_t)(\boldsymbol{Y}_t)\,\mathrm{d}t\right)\\
&- \tfrac{1}{2\sigma^2}\int_0^T (-\sigma^2\nabla\ln\pi_t + \nabla\phi_t)(\boldsymbol{Y}_t)\left((\sigma^2\nabla\ln\pi_t + \nabla\phi_t)(\boldsymbol{Y}_t)\,\mathrm{d}t + \sqrt{2}\sigma\,\overleftarrow{d}\,\boldsymbol{W}_t - \tfrac{1}{2}(-\sigma^2\nabla\ln\pi_t + \nabla\phi_t)(\boldsymbol{Y}_t)\,\mathrm{d}t\right)\\
&= (\ln\pi_0)(\boldsymbol{Y}_0) - (\ln\pi_T)(\boldsymbol{Y}_T) + \sigma^2\int_0^T |\nabla\ln\pi_t(\boldsymbol{Y}_t)|^2\,\mathrm{d}t\\
&+ \tfrac{\sigma}{\sqrt{2}}\left(\int_0^T \nabla\ln\pi_t(\boldsymbol{Y}_t)\cdot\overrightarrow{\mathrm{d}}\,\boldsymbol{W}_t + \int_0^T \nabla\ln\pi_t(\boldsymbol{Y}_t)\cdot\overleftarrow{\mathrm{d}}\,\boldsymbol{W}_t\right) + \tfrac{1}{\sigma\sqrt{2}}\left(\int_0^T \nabla\phi(\boldsymbol{Y}_t)\cdot\overrightarrow{\mathrm{d}}\,\boldsymbol{W}_t - \int_0^T \nabla\phi(\boldsymbol{Y}_t)\cdot\overleftarrow{\mathrm{d}}\,\boldsymbol{W}_t\right).
\end{aligned}$$

Using (35b), we obtain

$$\tfrac{\sigma}{\sqrt{2}}\left(\int_0^T \nabla\ln\pi_t(\boldsymbol{Y}_t)\cdot\overrightarrow{\mathrm{d}}\,\boldsymbol{W}_t + \int_0^T \nabla\ln\pi_t(\boldsymbol{Y}_t)\cdot\overleftarrow{\mathrm{d}}\,\boldsymbol{W}_t\right) = \sqrt{2}\sigma\int_0^T \nabla\ln\pi_t(\boldsymbol{Y}_t)\circ\mathrm{d}\boldsymbol{W}_t.$$

Furthermore, from (16) we see that

$$\tfrac{1}{\sigma\sqrt{2}}\left(\int_0^T \nabla\phi(\boldsymbol{Y}_t)\cdot\overrightarrow{\mathrm{d}}\,\boldsymbol{W}_t - \int_0^T \nabla\phi(\boldsymbol{Y}_t)\cdot\overleftarrow{\mathrm{d}}\,\boldsymbol{W}_t\right) = -\int_0^T \Delta\phi_t(\boldsymbol{Y}_t)\,\mathrm{d}t, \tag{52}$$

from which the claim follows.

Remark 13 (Estimating (19) without second derivatives)**.** Using (52), we can equivalently write the RND as

$$\begin{aligned}
&\left(\ln \frac{\mathrm{d}\overrightarrow{\mathbb{P}}_{\pi_0,\sigma^2\nabla\ln\pi+\nabla\phi}}{\mathrm{d}\overleftarrow{\mathbb{P}}_{\pi_T,-\sigma^2\nabla\ln\pi+\nabla\phi}}\right)(\boldsymbol{Y}) = \ln\pi_T(\boldsymbol{Y}_T) - \ln\pi_0(\boldsymbol{Y}_0)\\
&- \tfrac{1}{\sigma\sqrt{2}}\left(\int_0^T \nabla\phi(\boldsymbol{Y}_t)\cdot\overrightarrow{\mathrm{d}}\,\boldsymbol{W}_t - \int_0^T \nabla\phi(\boldsymbol{Y}_t)\cdot\overleftarrow{\mathrm{d}}\,\boldsymbol{W}_t\right) - \sigma\sqrt{2}\int_0^T \nabla\ln\pi_t(\boldsymbol{Y}_t)\circ\mathrm{d}\boldsymbol{W}_t - \sigma^2\int_0^T |\nabla\ln\pi_t(\boldsymbol{Y}_t)|^2\,\mathrm{d}t,
\end{aligned}$$

so that the loss in (19) can be estimated without the need to evaluate $\Delta\phi$. Note that the identity (52) is similar to a finite difference approximation of $\Delta\phi$ along the process $\boldsymbol{Y}_t$.

D.2. Existence and uniqueness of the drift

Before proceeding to the proof of Propostion 3.1, we state the following assumption on the curve of distributions $(\pi_t)_{t\in[0,T]}$:

Assumption D.1. Assume that $\pi \in C^\infty([0,T]\times\mathbb{R}^d;\mathbb{R})$, and that for all $t\in[0,T]$

1. the time derivative $\partial_t\pi_t$ is square-integrable, that is, $\partial_t\pi_t(t,\cdot) \in L^2(\mathbb{R}^d)$,
2. π_t satisfies a Poincaré inequality, that is, there exists a constant $C_t > 0$ such that

$$\mathrm{Var}_{\pi_t}(f) \le C_t\int_{\mathbb{R}^d} |\nabla f|^2\mathrm{d}\pi_t, \tag{54}$$

for all $f \in C_b^1(\mathbb{R}^d)$.

Note that at the boundary $\partial[0,T] = \{0,T\}$, we agree to denote by $\partial_t \pi_t$ the 'inward-pointing derivative' and interpret $C^\infty([0,T]\times\mathbb{R}^d;\mathbb{R})$ in that way. We remark that the Poincaré inequality (54) is satisfied under relatively mild conditions on the tails of π_t (for instance, Gaussian tails) and control of its derivatives, see, e.g., Bakry et al. (2014, Chapter 4). Under Assumption D.1, we can prove Proposition 3.1 as follows:

Proof of Proposition 3.1. The Fokker-Planck equation associated to (17) is given by

$$\partial_t \pi_t + \nabla \cdot (\pi_t \nabla \phi_t) = 0. \tag{55}$$

The operator $\phi \mapsto -\nabla\cdot(\pi_t\nabla\phi)$ is essentially self-adjoint in $L^2(\mathbb{R}^d)$, and, by (54) coercive on $L^2_0(\mathbb{R}^d) := \{f \in L^2(\mathbb{R}^d) : \int f \mathrm{d}x = 0\}$. Therefore, there exists a unique solution $\phi_t^* \in L^2_0(\mathbb{R}^d)$ to (55), for any $t \in [0,T]$. This solution is smooth by elliptic regularity. By Proposition 2.1 and our general framework, any minimiser $\widetilde{\phi}$ of (18) necessarily satisfies (55) as well. We then obtain

$$\nabla \cdot (\pi_t \nabla(\phi_t - \widetilde{\phi}_t)) = 0.$$

Multiplying this equation by $\phi_t - \widetilde{\phi}_t$, integrating, and integrating by parts shows that $\int \|\nabla(\phi - \widetilde{\phi})\|^2 \,\mathrm{d}\pi_t = 0$, proving the claim. □

Remark 14 (Relation to previous work)**.** Note we can carry out a change of variables to equation (55),

$$\partial_t \ln \pi_t = -\pi_t^{-1}(\nabla\pi_t \cdot \nabla\phi_t + \pi_t \Delta\phi) = -\nabla \ln \pi_t \cdot \nabla\phi - \Delta\phi,$$

yielding the PDE

$$\partial_t \ln \pi_t + \nabla \ln \pi_t \cdot \nabla\phi + \Delta\phi = 0,$$

which when considered in terms of the unnormalised flow $\hat{\pi}_t = Z_t \pi_t$ coincides with PDE in Vaikuntanathan and Jarzynski (2008); Arbel et al. (2021):

$$\partial_t \ln \hat{\pi}_t + \nabla \ln \hat{\pi}_t \cdot \nabla\phi + \Delta\phi - \mathbb{E}_{\pi_t}[\partial_t \ln \hat{\pi}_t] = 0.$$

In particular, we note that the Markov chain proposed in Arbel et al. (2021) converges to our proposed parameterisation in equation (17) (see equation (12) in (Arbel et al., 2021)).

E. Proofs

E.1. Proof of Proposition 2.2 (forward-backward Radon-Nikodym derivatives)

Proof. We begin with the forward Radon-Nikodym derivative

$$\ln\left(\frac{\mathrm{d}\overrightarrow{\mathbb{P}}^{\mu,a}}{\mathrm{d}\overrightarrow{\mathbb{P}}^{\nu,b}}\right)(\boldsymbol{Y}) = \ln\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\right)(\boldsymbol{Y}_0) + \tfrac{1}{\sigma^2}\int_0^T (a_t - b_t)(\boldsymbol{Y}_t)\cdot \overrightarrow{\mathrm{d}}\boldsymbol{Y}_t + \tfrac{1}{2\sigma^2}\int_0^T \left(b_t^2 - a_t^2\right)(\boldsymbol{Y}_t)\,\mathrm{d}t, \tag{56}$$

following from Girsanov's theorem (see, for instance, Nüsken and Richter (2021, Lemma A.1) and substitute $\sigma u = a - b$). To compute the backward Radon-Nikodym derivative, we temporarily introduce the time-reversal operator $\mathcal{R}$, acting as $(\mathcal{R}\boldsymbol{Y})_t := \boldsymbol{Y}_{T-t}$ on paths[11], and as $(\mathcal{R}a)_t(\boldsymbol{y}) := a_{T-t}(\boldsymbol{y})$ on vector fields. We then observe that

$$\ln\left(\frac{\mathrm{d}\overleftarrow{\mathbb{P}}^{\mu,\mathcal{R}a}}{\mathrm{d}\overleftarrow{\mathbb{P}}^{\nu,\mathcal{R}b}}\right)(\mathcal{R}\boldsymbol{Y}) = \ln\left(\frac{\mathrm{d}\overrightarrow{\mathbb{P}}^{\mu,a}}{\mathrm{d}\overrightarrow{\mathbb{P}}^{\nu,b}}\right)(\boldsymbol{Y}),$$

for instance by comparing the discrete-time processes in (10a) and (10b). Equivalently,

$$\ln\left(\frac{\mathrm{d}\overleftarrow{\mathbb{P}}^{\mu,a}}{\mathrm{d}\overleftarrow{\mathbb{P}}^{\nu,b}}\right)(\boldsymbol{Y}) = \ln\left(\frac{\mathrm{d}\overrightarrow{\mathbb{P}}^{\mu,\mathcal{R}a}}{\mathrm{d}\overrightarrow{\mathbb{P}}^{\nu,\mathcal{R}b}}\right)(\mathcal{R}\boldsymbol{Y}),$$

[11] Although pathwise definitions should be treated with care (because Itô integrals are defined only up to a set of measure zero), the arguments can be made rigorous using the machinery referred to in Appendix A.

since $\mathcal{R}^2$ is the identity. Building on (56), the backward Radon-Nikodym derivative therefore reads

$$\begin{aligned}\ln\left(\frac{\mathrm{d}\overleftarrow{\mathbb{P}}^{\mu,a}}{\mathrm{d}\overleftarrow{\mathbb{P}}^{\nu,b}}\right)(\boldsymbol{Y}) &= \ln\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\right)((\mathcal{R}\boldsymbol{Y})_0) + \tfrac{1}{\sigma^2}\int_0^T ((\mathcal{R}a)_t - (\mathcal{R}b)_t)(\mathcal{R}\boldsymbol{Y}_t)\cdot\overrightarrow{\mathrm{d}}(\mathcal{R}\boldsymbol{Y})_t \\ &\quad + \tfrac{1}{2\sigma^2}\int_0^T \left((\mathcal{R}b)_t^2 - (\mathcal{R}a)_t^2\right)((\mathcal{R}\boldsymbol{Y})_t)\,\mathrm{d}t, \\ &= \ln\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\right)(\boldsymbol{Y}_T) + \tfrac{1}{\sigma^2}\int_0^T (a_t - b_t)(\boldsymbol{Y}_t)\cdot\overleftarrow{\mathrm{d}}\boldsymbol{Y}_t + \tfrac{1}{2\sigma^2}\int_0^T \left(b_t^2 - a_t^2\right)(\boldsymbol{Y}_t)\,\mathrm{d}t, \end{aligned} \tag{57}$$

where the integrals have been transformed using the substitution $t \mapsto T - t$. The result in (15) now follows by writing

$$\ln\left(\frac{\mathrm{d}\overrightarrow{\mathbb{P}}^{\mu,a}}{\mathrm{d}\overleftarrow{\mathbb{P}}^{\nu,b}}\right)(\boldsymbol{Y}) = \ln\left(\frac{\mathrm{d}\overrightarrow{\mathbb{P}}^{\mu,a}}{\mathrm{d}\overrightarrow{\mathbb{P}}^{\Gamma_0,\gamma^+}}\right)(\boldsymbol{Y}) + \ln\left(\frac{\mathrm{d}\overleftarrow{\mathbb{P}}^{\Gamma_T,\gamma^-}}{\mathrm{d}\overleftarrow{\mathbb{P}}^{\nu,b}}\right)(\boldsymbol{Y}),$$

using the assumption $\overrightarrow{\mathbb{P}}^{\Gamma_0,\gamma^+} = \overleftarrow{\mathbb{P}}^{\Gamma_T,\gamma^-}$, and inserting (56) as well as (57). □

E.2. Proof of Proposition 3.2 (Crooks' fluctuation theorem and the Jarzinky equality)

Proof. For simplicity (and in order to use Proposition 2.2), we parameterise the SDE (17) in the same way as (13),

$$\mathrm{d}\boldsymbol{Y}_t = \frac{\sigma^2}{2}\nabla\ln\pi_t(\boldsymbol{Y}_t)\,\mathrm{d}t + \sigma\mathrm{d}\boldsymbol{W}_t,$$

which is simply relabeling $\sigma\sqrt{2} \mapsto \sigma$.

Using the divergence-based conversion formula for the backwards integral (see (16)) and letting $a = -b = \frac{\sigma^2}{2}\nabla\ln\pi_t$, we obtain

$$\begin{aligned}\ln\left(\frac{\mathrm{d}\overrightarrow{\mathbb{P}}^{\mu,\frac{\sigma^2}{2}\nabla\ln\pi}}{\mathrm{d}\overleftarrow{\mathbb{P}}^{\nu,-\frac{\sigma^2}{2}\nabla\ln\pi}}\right)(\boldsymbol{Y}) &= \ln\mu(\boldsymbol{Y}_0) - \ln\nu(\boldsymbol{Y}_T) + \tfrac{\sigma^2}{2}\int_0^T ||\nabla\ln\pi_t||^2(\boldsymbol{Y}_t)\,\mathrm{d}t \\ &\quad + \sigma\int_0^T \nabla\ln\pi_t(\boldsymbol{Y}_t)\cdot\overrightarrow{\mathrm{d}}\boldsymbol{W}_t + \frac{\sigma^2}{2}\int_0^T \Delta\ln\pi_t(\boldsymbol{Y}_t)\,\mathrm{d}t.\end{aligned}$$

Then via Itô's lemma applied to the unnormalised annealed log target $\ln\hat{\pi}_t = \ln\pi_t - \ln\mathcal{Z}_t$ we have:

$$\begin{aligned}&\ln\hat{\pi}_T(\boldsymbol{Y}_T) - \ln\hat{\pi}_0(\boldsymbol{Y}_0) - \int_0^T \partial_t\ln\hat{\pi}_t(\boldsymbol{Y}_t)\,\mathrm{d}t \\ &\quad = \frac{\sigma^2}{2}\int_0^T ||\nabla\ln\pi_t||^2(\boldsymbol{Y}_t)\,\mathrm{d}t + \frac{\sigma^2}{2}\int_0^T \Delta\ln\pi(\boldsymbol{Y}_t)\,\mathrm{d}t + \sigma\int_0^T \nabla\ln\pi_t(\boldsymbol{Y}_t)\cdot\overrightarrow{\mathrm{d}}\boldsymbol{W}_t,\end{aligned}$$

thus we arrive at Crooks' generalised fluctuation theorem (Crooks, 1999),

$$\ln\left(\frac{\mathrm{d}\overrightarrow{\mathbb{P}}^{\mu,\frac{\sigma^2}{2}\nabla\ln\pi}}{\mathrm{d}\overleftarrow{\mathbb{P}}^{\nu,-\frac{\sigma^2}{2}\nabla\ln\pi}}\right)(\boldsymbol{Y}) = \ln\mu(\boldsymbol{Y}_0) - \ln\nu(\boldsymbol{Y}_T) + \ln\hat{\pi}_T(\boldsymbol{Y}_T) - \ln\hat{\pi}_0(\boldsymbol{Y}_0) - \int_0^T \partial_t\ln\hat{\pi}_t(\boldsymbol{Y}_t)\,\mathrm{d}t, \tag{59}$$

for arbitrary initial initial and final densities μ and ν. Now notice that:

$$\begin{aligned}1 &= \mathbb{E}_{\overrightarrow{\mathbb{P}}^{\mu,\frac{\sigma^2}{2}\nabla\ln\pi}}\left[\left(\frac{\mathrm{d}\overrightarrow{\mathbb{P}}^{\mu,\frac{\sigma^2}{2}\nabla\ln\pi}}{\mathrm{d}\overleftarrow{\mathbb{P}}^{\nu,-\frac{\sigma^2}{2}\nabla\ln\pi}}\right)^{-1}\right] \\ &= \mathbb{E}_{\overrightarrow{\mathbb{P}}^{\mu,\frac{\sigma^2}{2}\nabla\ln\pi}}\left[\exp\left(-\ln\mu(\boldsymbol{Y}_0) + \ln\nu(\boldsymbol{Y}_T) - \ln\hat{\pi}_T(\boldsymbol{Y}_T) + \ln\hat{\pi}_0(\boldsymbol{Y}_0) + \int_0^T \partial_t\ln\hat{\pi}_t(\boldsymbol{Y}_t)\,\mathrm{d}t\right)\right],\end{aligned}$$

which implies the Jarzynski equality when considering the boundaries $\mu = \pi_0$ and $\nu = \pi_T$, resulting in:

$$\mathbb{E}_{\overrightarrow{\mathbb{P}}^{\pi_0, \frac{\sigma^2}{2}\nabla \ln \pi}}\left[\exp\left(\int_0^T \partial_t \ln \hat{\pi}_t(\boldsymbol{Y}_t)\,\mathrm{d}t\right)\right] = e^{-(\ln \mathcal{Z}_0 - \ln \mathcal{Z}_T)}.$$

Similarly considering the same boundaries in (59) yields Crooks' identity:

$$\left(\frac{\mathrm{d}\overrightarrow{\mathbb{P}}^{\pi_0, \sigma^2\nabla \ln \pi}}{\mathrm{d}\overleftarrow{\mathbb{P}}^{\pi_T, -\sigma^2\nabla \ln \pi}}\right)(\boldsymbol{Y}) = e^{-\frac{1}{\sigma^2}(\mathcal{F}_T - \mathcal{F}_0)} e^{\frac{1}{\sigma^2}\mathcal{W}_T}.$$

□

E.3. Proof of Proposition 4.1: EM $\iff$ IPF

Proof. The proof proceeds by induction.

To begin with, the update formula in (24a) implies that

$$\pi^1(\boldsymbol{x}, \boldsymbol{z}) = \underset{\pi(\boldsymbol{x}, \boldsymbol{z})}{\arg\min} \left\{D_{\mathrm{KL}}(\pi(\boldsymbol{x}, \boldsymbol{z}) || r(\boldsymbol{x}, \boldsymbol{z})) : \ \pi_{\boldsymbol{x}}(\boldsymbol{x}) = \mu(\boldsymbol{x})\right\},$$

recalling the initialisation $\pi(\boldsymbol{x}, \boldsymbol{z}) = r(\boldsymbol{x}, \boldsymbol{z})$. To take account of the marginal constraint, we may write $\pi(\boldsymbol{x}, \boldsymbol{z}) = \mu(\boldsymbol{x})\pi(\boldsymbol{z}|\boldsymbol{x})$ and vary over the conditionals $\pi(\boldsymbol{z}|\boldsymbol{x})$. By the chain rule for D_{KL}, we see that

$$D_{\mathrm{KL}}(\mu(\boldsymbol{x})\pi(\boldsymbol{z}|\boldsymbol{x}) || r(\boldsymbol{x}, \boldsymbol{z})) = D_{\mathrm{KL}}(\mu(\boldsymbol{x}) || r(\boldsymbol{x})) + \mathbb{E}_{\boldsymbol{x}\sim\mu(\boldsymbol{x})}[D_{\mathrm{KL}}(\pi(\boldsymbol{z}|\boldsymbol{x}) || r(\boldsymbol{z}|\boldsymbol{x}))], \tag{60}$$

which is minimised at $\pi(\boldsymbol{z}|\boldsymbol{x}) = r(\boldsymbol{z}|\boldsymbol{x})$. From this, it follows that $\pi^1(\boldsymbol{x}, \boldsymbol{z}) = \mu(\boldsymbol{x})r(\boldsymbol{z}|\boldsymbol{x})$ for the first IPF iterate. By assumption, the EM iteration is initialised in such a way that $q^{\phi_0}(\boldsymbol{z}|\boldsymbol{x}) = r(\boldsymbol{z}|\boldsymbol{x})$, so that indeed $\pi^1(\boldsymbol{x}, \boldsymbol{z}) = q^{\phi_0}(\boldsymbol{z}|\boldsymbol{x})\mu(\boldsymbol{x})$.

The induction step is split (depending on whether n is odd or even):

1.) First assume that the first line of (26) holds for a fixed odd $n \geq 1$. Our aim is to show that this implies that

$$\pi^{n+1}(\boldsymbol{x}, \boldsymbol{z}) = p^{\theta_{(n+1)/2}}(\boldsymbol{x}|\boldsymbol{z})\nu(\boldsymbol{z}), \tag{61}$$

that is, the second line of (26) with n replaced by $n + 1$. From (24b), we see that

$$\pi^{n+1}(\boldsymbol{x}, \boldsymbol{z}) = \underset{\pi(\boldsymbol{x}, \boldsymbol{z})}{\arg\min} \left\{D_{\mathrm{KL}}(\pi(\boldsymbol{x}, \boldsymbol{z}) || \pi^n(\boldsymbol{x}, \boldsymbol{z})) : \ \pi_{\boldsymbol{z}}(\boldsymbol{z}) = \nu(\boldsymbol{z})\right\}.$$

Again, we enforce the marginal constraint by setting $\pi(\boldsymbol{x}, \boldsymbol{z}) = \pi(\boldsymbol{z}|\boldsymbol{x})\nu(\boldsymbol{z})$ and proceed as in (60) to obtain $\pi^{n+1}(\boldsymbol{x}, \boldsymbol{z}) = \pi^n(\boldsymbol{x}|\boldsymbol{z})\nu(\boldsymbol{z})$. The statement in (61) is therefore equivalent to $\pi^n(\boldsymbol{x}|\boldsymbol{z}) = p^{\theta_{(n+1)/2}}(\boldsymbol{x}|\boldsymbol{z})$. To show this, we observe from the EM-scheme in (25) that

$$\theta_{(n+1)/2} = \underset{\theta}{\arg\min}\, \mathcal{L}_{D_{\mathrm{KL}}}(\phi_{(n-1)/2}, \theta).$$

In combination with the second line of (26) and the definition of $\mathcal{L}_D(\phi, \theta)$ in (5), we obtain

$$\theta_{(n+1)/2} = \underset{\theta}{\arg\min}\, D_{\mathrm{KL}}(\pi^n(\boldsymbol{x}, \boldsymbol{z}) || p^\theta(\boldsymbol{x}|\boldsymbol{z})\nu(\boldsymbol{z})) = \underset{\theta}{\arg\min}\, \mathbb{E}_{\boldsymbol{z}\sim\pi^n_{\boldsymbol{z}}(\boldsymbol{z})}\left[D_{\mathrm{KL}}(\pi^n(\boldsymbol{x}|\boldsymbol{z}) || p^\theta(\boldsymbol{x}|\boldsymbol{z}))\right],$$

where the second equality follows from the chain rule for D_{KL} as in (60). Since by assumption the parameterisation of $p^\theta(\boldsymbol{x}|\boldsymbol{z})$ is flexible, we indeed conclude that $\pi^n(\boldsymbol{x}|\boldsymbol{z}) = p^{\theta_{(n+1)/2}}(\boldsymbol{x}|\boldsymbol{z})$.

2.) Assume now that the second line of (26) holds for a fixed even $n \geq 2$. We need to show that the first line holds with n replaced by $n + 1$, that is,

$$\pi^{n+1}(\boldsymbol{x}, \boldsymbol{z}) = q^{\phi_{n/2}}(\boldsymbol{z}|\boldsymbol{x})\mu(\boldsymbol{x}).$$

Using similar arguments as before, we see that $\pi^{n+1}(\boldsymbol{x}, \boldsymbol{z}) = \pi^n(\boldsymbol{x}|\boldsymbol{z})\mu(\boldsymbol{x})$, so that it is left to show that $\pi^n(\boldsymbol{x}|\boldsymbol{z}) = q^{\phi_{n/2}}(\boldsymbol{z}|\boldsymbol{x})$. Along the same lines as in 1.), we obtain

$$\phi_{n/2} = \underset{\phi}{\arg\min}\, \mathcal{L}_{D_{\mathrm{KL}}}(\phi, \theta_{n/2}) = \underset{\phi}{\arg\min}\, D_{\mathrm{KL}}(q^\phi(\boldsymbol{z}|\boldsymbol{x})\mu(\boldsymbol{x}) || \pi^n(\boldsymbol{x}, \boldsymbol{z}))$$

$$= \arg\min_{\phi} \mathbb{E}_{\boldsymbol{x}\sim\mu(\boldsymbol{x})} \left[q^{\phi}(\boldsymbol{z}|\boldsymbol{x}) || \pi^{n}(\boldsymbol{z}|\boldsymbol{x}) \right].$$

Again, this allows us to conclude, since the parameterisation in $q^{\phi}(\boldsymbol{z}|\boldsymbol{x})$ is assumed to be flexible enough to allow for $q^{\phi_{n/2}}(\boldsymbol{z}|\boldsymbol{x}) = \pi^{n}(\boldsymbol{x}|\boldsymbol{z})$.

The proof for the path space IPF scheme is verbatim the same after adjusting the notation. For completeness, we consider a drift-wise version below. □

E.4. Drift based EM

As remarked in the previous subsection, the proof of the equivalence between IPF and EM in path space follows the exact same lines, replacing the chain rule of D_{KL} with the (slightly more general) disintegration theorem (Léonard, 2014b). In this section, we provide a direct extension to the control setting, yielding yet another IPF-type algorithm and motivating certain design choices for the family of methods we study.

Corollary E.1. *For the intialisation $\phi_0 = 0$, the alternating scheme*

$$\theta_{n+1} = \arg\min_{\theta} D_{\mathrm{KL}}(\overrightarrow{\mathbb{P}}^{\mu, f+\sigma^2\nabla\phi_n}, \overleftarrow{\mathbb{P}}^{\nu, f+\sigma^2\nabla\theta}), \quad \phi_{n+1} = \arg\min_{\phi} D_{\mathrm{KL}}(\overrightarrow{\mathbb{P}}^{\mu, f+\sigma^2\nabla\phi}, \overleftarrow{\mathbb{P}}^{\nu, f+\sigma^2\nabla\theta_{n+1}})$$

agrees with the path space IPF iterations in Bernton et al. (2019); Vargas et al. (2021a); De Bortoli et al. (2021).

Proof. For brevity let $\mathcal{L}_{\mathrm{FB}}(\phi, \theta) := D_{\mathrm{KL}}(\overrightarrow{\mathbb{P}}^{\mu, f+\sigma^2\nabla\phi}, \overleftarrow{\mathbb{P}}^{\nu, f+\sigma^2\nabla\theta})$. Additionally, we parameterise the forwards and backwards SDEs with respective path distributions $\overrightarrow{\mathbb{P}}^{\mu, f+\sigma^2\nabla\phi}$, $\overleftarrow{\mathbb{P}}^{\nu, f+\sigma^2\nabla\theta}$ as:

$$\begin{aligned} \mathrm{d}\boldsymbol{Y}_t &= f_t(\boldsymbol{Y}_t)\,\mathrm{d}t + \sigma^2\nabla\phi_t(\boldsymbol{Y}_t)\,\mathrm{d}t + \sigma\,\overrightarrow{\mathrm{d}}\boldsymbol{W}_t, \quad \boldsymbol{Y}_0 \sim \mu, \\ \mathrm{d}\boldsymbol{Y}_t &= f_t(\boldsymbol{Y}_t)\,\mathrm{d}t + \sigma^2\nabla\theta_t(\boldsymbol{Y}_t)\,\mathrm{d}t + \sigma\,\overleftarrow{\mathrm{d}}\boldsymbol{W}_t, \quad \boldsymbol{Y}_T \sim \nu. \end{aligned}$$

The proof will proceed quite similarly, so instead we will consider just the inductive step for the odd half bridge:

$$\theta_n = \arg\min_{\theta} \mathcal{L}_{\mathrm{FB}}(\phi_{n-1}, \theta).$$

We can show via the D_{KL} chain rule and the disintegration theorem (Léonard, 2014b) that the above is minimised when θ satisfies $\overleftarrow{\mathbb{P}}^{\nu, f+\sigma^2\nabla\theta} = \overrightarrow{\mathbb{P}}^{\mu, f+\sigma^2\nabla\phi_{n-1}} \frac{\mathrm{d}\nu}{\mathrm{d}\rho_T^{\mu, f+\sigma^2\nabla\phi_{n-1}}}$ which corresponds to $\nabla\theta_n = \sigma^2\nabla\phi_{n-1} - \sigma^2\nabla\ln\rho_t^{\mu, f+\sigma^2\nabla\phi_{n-1}}$ following Observation 1 in Vargas et al. (2021a). Similarly as per Proposition 4.1 the results will follow for the even half bridges.

□

EM initialisation: The above corollary provides us with convergence guarantees when performing coordinate descent on $D_{\mathrm{KL}}(\overrightarrow{\mathbb{P}}^{\mu, f+\sigma^2\nabla\phi}, \overleftarrow{\mathbb{P}}^{\nu, f+\sigma^2\nabla\theta})$ subject to initialising $\phi_0 = 0$. n practice, this indicates that the way of initialising ϕ has a major impact on which bridge we converge to.

Thus as a rule of thumb we propose initialising $\phi_0 = 0$ such that we initialise at the Schrödinger prior: then one may carry out joint updates as an alternate heuristic, we call this approach DNF (EM Init), as it is effectively a clever initialisation of DNF inspired by the relationship between IPF and EM.

E.5. Proof of Proposition 4.3 (HJB-regularisers)

This result can be found in Chen et al. (2021, Proposition 5.1), for instance, but since it is relevant to the connections pointed out in Remark 15 below, we present an independent proof:

Proof. We denote the path measures associated to the SDE

$$\mathrm{d}\boldsymbol{Y}_t = f_t(\boldsymbol{Y}_t)\,\mathrm{d}t + \sigma\,\mathrm{d}\boldsymbol{W}_t \tag{63}$$

by $\mathbb{P}$ and the SDE (27) by $\mathbb{P}^\phi$, respectively. According to Girsanov's theorem, the Radon-Nikodym derivative satisfies

$$\frac{\mathrm{d}\mathbb{P}^\phi}{\mathrm{d}\mathbb{P}} = \exp\left(\sigma \int_0^T \nabla\phi_t(\boldsymbol{Y}_t)\cdot \mathrm{d}\boldsymbol{W}_t - \tfrac{\sigma^2}{2}\int_0^T |\nabla\phi_t|^2(\boldsymbol{Y}_t)\,\mathrm{d}t\right), \tag{64}$$

provided that the marginals agree at initial time, $\mathbb{P}_0 = \mathbb{P}_0^\phi$. Along solutions of (63), we have by Itô's formula

$$\begin{aligned}\phi_T(\boldsymbol{Y}_T) - \phi_0(\boldsymbol{Y}_0) &= \int_0^T \partial_t\phi_t(\boldsymbol{Y}_t)\,\mathrm{d}t + \int_0^T (f_t\cdot\nabla\phi_t)(\boldsymbol{Y}_t)\,\mathrm{d}t + \tfrac{\sigma^2}{2}\int_0^T \Delta\phi_t(\boldsymbol{Y}_t)\,\mathrm{d}t + \sigma\int_0^T \nabla\phi_t(\boldsymbol{Y}_t)\cdot\mathrm{d}\boldsymbol{W}_t \\ &= -\tfrac{\sigma^2}{2}\int_0^T |\nabla\phi_t|^2(\boldsymbol{Y}_t)\,\mathrm{d}t + \sigma\int_0^T \nabla\phi_t(\boldsymbol{Y}_t)\cdot\mathrm{d}\boldsymbol{W}_t = \ln\left(\frac{\mathrm{d}\mathbb{P}^\phi}{\mathrm{d}\mathbb{P}}\right),\end{aligned} \tag{65}$$

where we have used the HJB-equation (28) in the second line. Combining this with (64), we see that

$$\frac{\mathrm{d}\mathbb{P}^\phi}{\mathrm{d}\mathbb{P}}(\boldsymbol{Y}) = \exp\left(-\phi_0(\boldsymbol{Y}_0)\right)\exp\left(\phi_T(\boldsymbol{Y}_T)\right). \tag{66}$$

The claim now follows, since the unique solution to the Schrödinger problem is characterised by the product-form expression in (66, see Léonard (2014a, Section 2), together with the marginal constraints $\mathbb{P}_0^\phi = \mu$ and $\mathbb{P}_T^\phi = \nu$, which are satisfied by assumption. □

Remark 15 (Connection to *reciprocal classes* (Shi et al., 2023; Peluchetti, 2023) and *TD learning* (Liu et al., b))**.** The calculation in equation (65) makes the relationship between the HJB equation (28) and reciprocal classes manifest (since reciprocal classes can essentially be defined through the relationship (66), see Léonard et al. (2014); Rœlly (2013)). Moreover, equation (65) showcases the relationship between TD learning (Sutton and Barto, 2018, Chapter 6) as suggested in Liu et al. (b) and HJB regularisation. Indeed,

$$\mathrm{Reg}_{\mathrm{BSDE}}(\phi) := \mathrm{Var}\left(\phi_T(\boldsymbol{Y}_T) - \phi_0(\boldsymbol{Y}_0) + \tfrac{\sigma^2}{2}\int_0^T |\nabla\phi_t|^2(\boldsymbol{Y}_t)\,\mathrm{d}t - \sigma\int_0^T \nabla\phi_t(\boldsymbol{Y}_t)\cdot\mathrm{d}\boldsymbol{W}_t\right), \tag{67}$$

where the variance is taken with respect to the path measure induced by (63), is a valid HJB-regulariser in the sense of Corollary 4.4. The equivalence between $\mathrm{Reg}_{\mathrm{BSDE}}(\phi) = 0$ and the HJB equation (28) follows from the theory of backward stochastic differential equations (BSDEs)[12], see, for example, the proof of Proposition 3.4 in Nüsken and Richter (2023) and the discussion in Nüsken and Richter (2021, Section 3.2).

In the following, we present an analogue of Proposition 4.3 involving the backward drift (Chen et al., 2019):

Proposition E.2. *Assume that $\theta \in C^{1,2}([0,T]\times\mathbb{R}^d;\mathbb{R})$ satisfies the following two conditions:*

1. *The backward SDE*
$$\mathrm{d}\boldsymbol{Y}_t = f_t(\boldsymbol{Y}_t)\,\mathrm{d}t + \sigma^2\nabla\theta_t(\boldsymbol{Y}_t)\,\mathrm{d}t + \sigma\,\overleftarrow{\mathrm{d}}\,\boldsymbol{W}_t, \qquad \boldsymbol{Y}_T \sim \nu \tag{68}$$
admits a unique strong solution on $[0,T]$, satisfying moreover the initial constraint $\boldsymbol{Y}_0 \sim \mu$.

2. *The Hamilton-Jacobi-Bellmann (HJB) equation*
$$\partial_t\theta + f\cdot\nabla\theta - \frac{\sigma^2}{2}\Delta\theta + \tfrac{\sigma^2}{2}|\nabla\theta|^2 - \nabla\cdot f = 0 \tag{69}$$
holds for all $(t,x)\in[0,T]\times\mathbb{R}^d$.

Assuming furthermore that the solution to (68) admits a smooth positive density ρ, we have that $a_t = \nabla\theta_t + \sigma^2\nabla\ln\rho_t$ provides the unique solution to the Schrödinger problem as posed in (23).

Remark 16. As opposed to Chen et al. (2016, equation (41)), the HJB-equation (69) does not involve the time reversal of the Schrödinger prior; the form of the HJB equations is not uniquely determined. On the other hand, (69) contains the divergence term $\nabla\cdot f$, which discourages us from enforcing this constraint in the same way as (28). An akin result can be found in (Liu et al., b) stated in terms of BSDEs.

[12] ... not to be confused with reverse-time SDEs as in (13).

Proof of Corollary E.2. Using the forward-backward Radon-Nikodym derivative in (15), we compute

$$\ln\left(\frac{\mathrm{d}\overrightarrow{\mathbb{P}}^{\mu,f}}{\mathrm{d}\overleftarrow{\mathbb{P}}^{\nu,f+\sigma^2\nabla\psi}}\right)(\boldsymbol{Y}) = \ln\left(\frac{\mathrm{d}\mu}{\mathrm{dLeb}}\right) - \ln\left(\frac{\mathrm{d}\nu}{\mathrm{dLeb}}\right) + \sigma\int_0^T f_t(\boldsymbol{Y}_t)\cdot\mathrm{d}\boldsymbol{W}_t - \sigma\int_0^T f_t(\boldsymbol{Y}_t)\cdot\overleftarrow{\mathrm{d}}\boldsymbol{W}_t$$
$$-\sigma\int_0^T \theta_t(\boldsymbol{Y}_t)\cdot\overleftarrow{\mathrm{d}}\boldsymbol{W}_t + \tfrac{\sigma^2}{2}\int_0^T |\nabla\theta_t|^2(\boldsymbol{Y}_t)\,\mathrm{d}t$$
$$= \ln\left(\frac{\mathrm{d}\mu}{\mathrm{dLeb}}\right) - \ln\left(\frac{\mathrm{d}\nu}{\mathrm{dLeb}}\right) - \sigma\int_0^T (\nabla\cdot f_t)(\boldsymbol{Y}_t)\,\mathrm{d}t - \sigma\int_0^T \theta_t(\boldsymbol{Y}_t)\cdot\overleftarrow{\mathrm{d}}\boldsymbol{W}_t + \tfrac{\sigma^2}{2}\int_0^T |\nabla\theta_t|^2(\boldsymbol{Y}_t)\,\mathrm{d}t.$$

Here we have chosen $\overrightarrow{\gamma} = \overleftarrow{\gamma} = 0$, and $\Gamma_0 = \Gamma_T = \mathrm{Leb}$. The initial measure for the Schrödinger prior is μ, but the argument is unaffected by this choice (as the solution is independent of this). We now use the (backward) Itô formula along the Schrödinger prior,

$$\theta_t(\boldsymbol{Y}_T) - \theta_0(\boldsymbol{Y}_0) = \int_0^T \partial_t\theta_t(\boldsymbol{Y}_t)\,\mathrm{d}t + \int_0^T \nabla\theta_t(\boldsymbol{W}_t)\cdot\overleftarrow{\mathrm{d}}\boldsymbol{W}_t + \int_0^T \nabla\theta_t(\boldsymbol{Y}_t)\cdot f_t(\boldsymbol{Y}_t)\,\mathrm{d}t - \tfrac{1}{2}\int_0^T \Delta\theta_t(\boldsymbol{Y}_t)\,\mathrm{d}t.$$

Using the HJB-equation (69), we see that

$$\ln\left(\frac{\mathrm{d}\overrightarrow{\mathbb{P}}^{\mu,f}}{\mathrm{d}\overleftarrow{\mathbb{P}}^{\nu,f+\sigma^2\nabla\theta}}\right)(\boldsymbol{Y}) = \ln\left(\frac{\mathrm{d}\mu}{\mathrm{dLeb}}\right) - \ln\left(\frac{\mathrm{d}\nu}{\mathrm{dLeb}}\right) - \theta_t(\boldsymbol{Y}_T) + \theta_0(\boldsymbol{Y}_0), \tag{70}$$

and we can conclude as in the proof of Proposition 4.3. □

F. Numerical implementation

F.1. Discretisation and connection to DNFs (diffusion normalising flows)

In this section we derive the main discretisation formula used in our implementations for the forward-backwards Radon-Nikodym derivative (RND).

Proposition F.1. *Letting $\Gamma_0 = \Gamma_T = \mathrm{Leb}$ and $\gamma^{\pm} = 0$, we have that the RND in (15) is given by*

$$\ln\left(\frac{\mathrm{d}\overrightarrow{\mathbb{P}}^{\mu,a}}{\mathrm{d}\overleftarrow{\mathbb{P}}^{\nu,b}}\right)(\boldsymbol{Y}) = \ln\mu(\boldsymbol{Y}_0) - \ln\nu(\boldsymbol{Y}_T) + \tfrac{1}{\sigma^2}\int_0^T a_t(\boldsymbol{Y}_t)\cdot\overrightarrow{\mathrm{d}}\boldsymbol{Y}_t - \tfrac{1}{2\sigma^2}\int_0^T ||a_t(\boldsymbol{Y}_t)||^2\,\mathrm{d}t$$
$$-\tfrac{1}{\sigma^2}\int_0^T b_t(\boldsymbol{Y}_t)\cdot\overleftarrow{\mathrm{d}}\boldsymbol{Y}_t + \tfrac{1}{2\sigma^2}\int_0^T ||b_t(\boldsymbol{Y}_t)||^2\,\mathrm{d}t, \qquad \overrightarrow{\mathbb{P}}^{\mu,a}\text{-almost surely,}$$

and admits the following discrete-time approximation up to constant terms in a_t and b_t (following Remark 3),

$$\ln\left(\widehat{\frac{\mathrm{d}\overrightarrow{\mathbb{P}}^{\mu,a}}{\mathrm{d}\overleftarrow{\mathbb{P}}^{\nu,b}}}\right)(\boldsymbol{Y}) = -\ln\nu(\boldsymbol{Y}_T) + \sum_{i=0}^{K-1}\tfrac{1}{2\sigma^2\sqrt{t_{i+1}-t_{i+1}}}||\boldsymbol{Y}_{t_i} - \boldsymbol{Y}_{t_{i+1}} + b_{t_{i+1}}(\boldsymbol{Y}_{t_{i+1}})(t_{i+1}-t_{i+1})||^2 + \text{const.},$$

when using the Euler-Maruyama discretisation:

$$\boldsymbol{Y}_{t_{i+1}} = \boldsymbol{Y}_{t_i} + a_{t_i}(\boldsymbol{Y}_{t_i})(t_{i+1}-t_i) + \sqrt{(t_{i+1}-t_i)}\sigma\xi, \qquad \xi\sim\mathcal{N}(0,I).$$

Proof. The first part follows by direct computation.

From here on, we will use the notation $f_{t_i} = f_{t_i}(\boldsymbol{Y}_{t_i})$ for brevity. Following Remark 3 we have that

$$\ln\left(\frac{\mathrm{d}\overrightarrow{\mathbb{P}}^{\mu,a}}{\mathrm{d}\overleftarrow{\mathbb{P}}^{\nu,b}}\right)(\boldsymbol{Y}) \approx \ln\mu(\boldsymbol{Y}_0) - \ln\nu(\boldsymbol{Y}_T)$$
$$+\tfrac{1}{\sigma^2}\sum_{i=0}^{K-1} a_{t_i}\cdot(\boldsymbol{Y}_{t_{i+1}} - \boldsymbol{Y}_{t_i}) - \tfrac{1}{2\sigma^2}\sum_{i=0}^{K-1}||a_{t_i}||^2(t_{i+1}-t_i)$$

$$- \frac{1}{\sigma^2} \sum_{i=0}^{K-1} b_{t_{i+1}} \cdot (\boldsymbol{Y}_{t_{i+1}} - \boldsymbol{Y}_{t_i}) + \frac{1}{2\sigma^2} \sum_{i=0}^{K-1} ||b_{t_{i+1}}||^2 (t_{i+1} - t_i).$$

Adding and subtracting $\boldsymbol{Y}_{t_{i+1}} - \boldsymbol{Y}_{t_i}$ allows us to complete the square in each sum, resulting in:

$$\ln \left(\frac{\mathrm{d} \overrightarrow{\mathbb{P}}^{\mu,a}}{\mathrm{d} \overleftarrow{\mathbb{P}}^{\nu,b}} \right) (\boldsymbol{Y}) \approx \ln \mu(\boldsymbol{Y}_0) - \ln \nu(\boldsymbol{Y}_T) + \sum_{i=0}^{K-1} \frac{1}{2\sigma^2 \sqrt{t_{i+1} - t_i}} ||\boldsymbol{Y}_{t_{i+1}} - \boldsymbol{Y}_{t_i} - a_{t_i}(t_{i+1} - t_i)||^2$$
$$+ \sum_{i=0}^{K-1} \frac{1}{2\sigma^2 \sqrt{t_{i+1} - t_i}} ||\boldsymbol{Y}_{t_i} - \boldsymbol{Y}_{t_{i+1}} + b_{t_{i+1}}(t_{i+1} - t_i)||^2. \tag{73}$$

Now notice that under the Euler-Maruyama discretisation $||\boldsymbol{Y}_{t_{i+1}} - \boldsymbol{Y}_{t_i} - a_{t_i}(t_{i+1} - t_i)||^2 = \sigma^2 ||\xi||^2$ where $\xi \sim \mathcal{N}(0, I)$ does not depend on a_t or b_t; in particular when using D_{KL} for the divergence we have that $\mathbb{E}_{\overrightarrow{\mathbb{P}}^{\mu,a}_{\mathrm{EM}}} ||\boldsymbol{Y}_{t_{i+1}} - \boldsymbol{Y}_{t_i} - a_{t_i}(t_{i+1} - t_i)||^2 = \sigma^2$ and thus:

$$\ln \left(\widehat{\frac{\mathrm{d} \overrightarrow{\mathbb{P}}^{\mu,a}}{\mathrm{d} \overleftarrow{\mathbb{P}}^{\nu,b}}} \right) (\boldsymbol{Y}) \propto \ln \mu(\boldsymbol{Y}_0) - \ln \nu(\boldsymbol{Y}_T) + \sum_{i=0}^{K-1} \frac{1}{2\sigma^2 \sqrt{t_{i+1} - t_i}} ||\boldsymbol{Y}_{t_i} - \boldsymbol{Y}_{t_{i+1}} + b_{t_{i+1}}(t_{i+1} - t_i)||^2. \tag{74}$$

□

Notice that in expectation (for computing D_{KL}), equation (74) matches equation (15) in Zhang and Chen (2021) and thus provides a theoretical backing to the objective used in Zhang and Chen (2021). Resolving the term $\mathbb{E}_{\overrightarrow{\mathbb{P}}^{\mu,a}_{\mathrm{EM}}} ||\boldsymbol{Y}_{t_{i+1}} - \boldsymbol{Y}_{t_i} - a_{t_i}(t_{i+1} - t_i)||^2$ analytically may offer a variance reduction similar to the analytic calculations in Sohl-Dickstein et al. (2015, Equation 14) and the Rao-Blackwelizations of D_{KL} in Ho et al. (2020).

Remark 17. The time discretised RND in equation (73) can be expressed as the ratio of the transition densities corresponding to two discrete-time Markov chains $\mu(\boldsymbol{y}_0) q^a(\boldsymbol{y}_{1:K}|\boldsymbol{y}_0) / p^b(\boldsymbol{y}_{0:K-1}|\boldsymbol{y}_K) \nu(\boldsymbol{y}_K)$ with $\boldsymbol{y}_{0:K} \sim q^a(\boldsymbol{y}_{1:K}|\boldsymbol{y}_0)\mu(\boldsymbol{y}_0)$. As a result considering $\nu(x) = \hat{\nu}(\boldsymbol{x})/Z$ and the IS estimator $\hat{Z} = p^b(\boldsymbol{y}_{0:K-1}|\boldsymbol{y}_K)\hat{\nu}(\boldsymbol{y}_K) / \mu(\boldsymbol{y}_0) q^a(\boldsymbol{y}_{1:K}|\boldsymbol{y}_0)$ it follows that $\mathbb{E}_{q^a(\boldsymbol{y}_{1:K}|\boldsymbol{y}_0)\mu(\boldsymbol{y}_0)}[\ln \hat{Z}]$ is an ELBO of $\hat{Z}$ (e.g. $\mathbb{E}_{q^a(\boldsymbol{y}_{1:K}|\boldsymbol{y}_0)\mu(\boldsymbol{y}_0)}[\ln \hat{Z}] \leq \ln Z$).

Whilst superficially simple, Remark 17 guarantees that normalizing constant estimators arising from our discretisation do not overestimate the true normalizing constant. This result is beneficial in practice as it allows us to compare estimators possessing this property by selecting the one with the largest value. As highlighted in (Vargas et al., 2023a) many SDE discretisations can result in estimators that do not yield an ELBO: for example, the estimators used in (Berner et al., 2022) can result in overestimating the normalising constant. Note similar remarks have been established in the context of free energy computation and the Jarzynski equality see Stoltz et al. (2010, Remark 4.5).

F.2. Neural network parameterisations

Following Zhang and Chen (2021) and the recent success in score generative modelling we choose the following parameterisations:

$$a_t(\boldsymbol{x}) = f_t(\boldsymbol{x}) + \sigma^2 \nabla \phi(t, \boldsymbol{x}), \tag{75a}$$
$$b_t(\boldsymbol{x}) = f_t(\boldsymbol{x}) + \sigma^2 \nabla \phi(t, \boldsymbol{x}) - \sigma^2 s_\theta(t, \boldsymbol{x}), \tag{75b}$$

where s_θ is a score network (Song et al., 2021; De Bortoli et al., 2021; Zhang and Chen, 2021) and $\phi(t, \boldsymbol{x})$ is a neural network potential. We adapt the architectures proposed in Onken et al. (2021); Koshizuka and Sato (2023) to general activation functions. Note that these architectures allow for fast computation of $\Delta\phi$ comparable to that of Hutchinson's trace estimator (Grathwohl et al., 2019; Hutchinson, 1989).

Finally, we remark that the parametrisation in (75b) allows us to learn the score of the learned SDE and thus seamlessly adapt our approach to using the probability flow ODE (Song et al., 2021) at inference time.

F.3. 2D toy targets - generative modelling

For these tasks, we mirror the experimental setup in Zhang and Chen (2021) where we start from $t_0 = 0.001$ rather than 0 and go up to $T = 0.05$ and an exponential schedule with base 0.9 for the discretisation grid as specified in

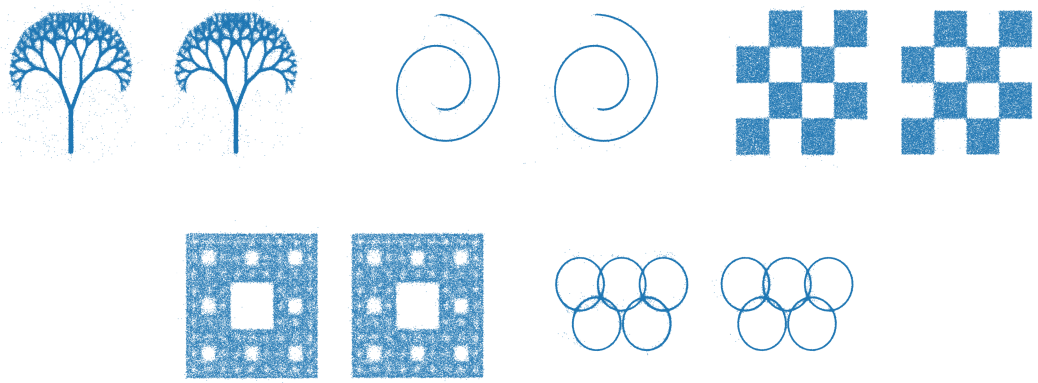


Figure 2. Generated samples trained by our approach ($\lambda = 0.5$) left and DNF ($\lambda = 0$) right. Qualitatively we can observe that both learned models have similarly matched marginals.

https://github.com/qsh-zh/DiffFlow/blob/62200f0e1089c97e16e7fc38cf3db2526bfeae11/utils/scalars.py#L58. Furthermore we use a diffusion coefficient of $\sigma = 0.2$.

F.4. Double well potential

We used the following potential (Vargas et al., 2021a):

$$U\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \frac{5}{2}(x^2 - 1)^2 + y^2 + \frac{1}{\delta}\exp\left(-\frac{x^2 + y^2}{\delta}\right), \tag{76}$$

with $\delta = 0.35$, furthermore, we used the boundary distributions:

$$\mu \sim \mathcal{N}\left(\begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.0125 & 0 \\ 0 & 0.15 \end{pmatrix}\right), \quad \nu \sim \mathcal{N}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.0125 & 0 \\ 0 & 0.15 \end{pmatrix}\right).$$

The Schrödinger prior is given by:

$$\mathrm{d}\boldsymbol{Y}_t = -\nabla_{\boldsymbol{Y}_t} U(\boldsymbol{Y}_t)\,\mathrm{d}t + \sigma\,\mathrm{d}\boldsymbol{W}_t, \tag{77}$$

with $\sigma = 0.4$. The terminal time is $T = 1$. Furthermore, we employ the same exponential discretisation scheme as in the generative modelling experiments.

F.5. PINN Loss

For the PINN loss across all tasks, we sample the trajectories from $\boldsymbol{Y}^{\phi}_{0:T} \sim \overrightarrow{\mathbb{P}}^{\mu,\nabla\phi}$ and thus employ the same discretisation as used in the KL loss. However, we detach the trajectories $\boldsymbol{Y}^{\mathrm{detach}(\phi)}_{0:T}$ before calculating the gradient updates in a similar fashion to Nüsken and Richter (2021).