

## Review



**Cite this article:** Doucet A, Moulines E, Thin A. 2023 Differentiable samplers for deep latent variable models. *Phil. Trans. R. Soc. A* **381**: 20220147.  
<https://doi.org/10.1098/rsta.2022.0147>

Received: 2 September 2022

Accepted: 15 February 2023

One contribution of 16 to a theme issue  
 ‘Bayesian inference: challenges, perspectives,  
 and prospects’.

**Subject Areas:**  
 statistics

**Keywords:**  
 Bayesian inference, importance sampling,  
 Monte Carlo methods, variational inference

**Authors for correspondence:**

Arnaud Doucet

e-mail: [doucet@stats.ox.ac.uk](mailto:doucet@stats.ox.ac.uk)

Eric Moulines

e-mail: [eric.moulines@polytechnique.edu](mailto:eric.moulines@polytechnique.edu)

# Differentiable samplers for deep latent variable models

Arnaud Doucet<sup>1</sup>, Eric Moulines<sup>2</sup> and Achille Thin<sup>2</sup>

<sup>1</sup>Department of Statistics, Oxford University, Oxford, UK

<sup>2</sup>Ecole Polytechnique, Centre de Mathématiques Appliquées, CNRS UMR 7641, Palaiseau, France

AD, 0000-0002-7662-419X

Latent variable models are a popular class of models in statistics. Combined with neural networks to improve their expressivity, the resulting deep latent variable models have also found numerous applications in machine learning. A drawback of these models is that their likelihood function is intractable so approximations have to be carried out to perform inference. A standard approach consists of maximizing instead an evidence lower bound (ELBO) obtained based on a variational approximation of the posterior distribution of the latent variables. The standard ELBO can, however, be a very loose bound if the variational family is not rich enough. A generic strategy to tighten such bounds is to rely on an unbiased low-variance Monte Carlo estimate of the evidence. We review here some recent importance sampling, Markov chain Monte Carlo and sequential Monte Carlo strategies that have been proposed to achieve this.

This article is part of the theme issue ‘Bayesian inference: challenges, perspectives, and prospects’.

## 1. Latent variable models

We consider here that observation  $x \in \mathcal{X} \subset \mathbb{R}^{d_x}$  is a sample from a probability density  $p^*(x)$ . We approximate  $p^*(x)$  by considering a parametric model  $p_\theta(x)$ , with parameters  $\theta \in \Theta \subset \mathbb{R}^d$ , defined by

$$p_\theta(x) = \int_{\mathcal{Z}} p_\theta(x, z) \, dz, \quad (1.1)$$

where  $z \in \mathcal{Z} \subset \mathbb{R}^{d_z}$  is a latent variable. Such an implicit distribution over  $x$  is very flexible. If  $z$  is discrete and  $p_\theta(x|z)$  is a Gaussian distribution, then  $p_\theta(x)$  is a Gaussian mixture. If  $z$  is continuous as considered

here,  $p_\theta(x)$  is an infinite mixture of the  $p_\theta(x|z)$ , thus more expressive than finite mixtures. We use the term **deep latent variable model (DLVM)** to refer to a latent variable model  $p_\theta(x, z)$  whose distribution is parameterized by neural networks. DLVMs include, in particular, the **deep latent Gaussian models (DLGMs)** [1]—which encompass the popular *variational auto-encoder* [2–4] described further. They have a wide range of applications, from generative modelling to semi-supervised learning and representation learning.

DLGMs rely on  $p_\theta(x, z) = p(z)p_\theta(x|z)$ , where  $p(z)$  is a standard multivariate Gaussian prior distribution over latent space and  $p_\theta(x|z)$  is the conditional distribution of the observation given the latent variables (also referred to as the *decoder*). An advantage of DLGMs and more generally DLVMs is that the marginal distribution  $p_\theta(x)$  can be very complex, even if each factor (prior and decoder) in the model is relatively simple. When considered as a function  $\theta \mapsto p_\theta(x)$  for a given  $x$ ,  $p_\theta(x)$  is called the *marginal likelihood* or *evidence*—we sometimes omit ‘marginal’, when there is no ambiguity. For a given value of  $x$ ,  $p_\theta(x)$  is the normalizing constant of the unnormalized distribution  $z \rightarrow p_\theta(x, z)$ .

Given data  $x$ , we would like to maximize the log marginal likelihood  $\ell(\theta, x) = \log p_\theta(x)$ . The main difficulty of maximum-likelihood learning in complex latent variable models such as DLVMs is that  $\ell(\theta, x)$  is intractable. This intractability is due to the integral (1.1) which must be computed to obtain the model evidence. Because of this intractability, we cannot easily differentiate and optimize these models according to their parameters. While Fisher’s identity shows that

$$\nabla_\theta \ell(\theta, x) = \int \nabla \log p_\theta(x, z) p_\theta(z|x) dz, \quad (1.2)$$

the posterior distribution  $p_\theta(z|x)$  is itself intractable and importance sampling or Markov chain Monte Carlo (MCMC) approximations of it are typically necessary.

The approach that has become prominent in the machine learning literature is to rely instead on **variational inference (VI)**. VI introduces a parametric family of distributions, the variational family  $\mathcal{Q} = \{q_\phi, \phi \in \Phi\}$ , which is parameterized by  $\phi \in \Phi \subset \mathbb{R}^d$ . The goal is to choose the parameter  $\phi$  such that  $q_\phi(z|x)$  approximates the true posterior  $p_\theta(z|x)$  for given  $x, \theta$ , typically intractable for DLVMs. The likelihood function is defined via a generally intractable integral. To circumvent marginalization, a common approach is to optimize a *variational lower bound* on the log marginal likelihood (often abusively referred to as the marginal log-likelihood); see [2,5,6].

$$\mathcal{L}(\theta, \phi) = \int_z \log \left( \frac{p_\theta(x, z)}{q_\phi(z|x)} \right) q_\phi(z|x) dz = \ell(\theta, x) - D_{\text{KL}}(q_\phi(\cdot|x) || p_\theta(\cdot|x)), \quad (1.3)$$

where  $D_{\text{KL}}$  denotes the Kullback–Leibler divergence. This objective is thus always lower than the marginal log-likelihood  $\ell(\theta, x)$  and is thus referred to as the evidence lower bound (ELBO). **The expressiveness of  $\mathcal{Q}$  is essential for good performance**; e.g. [7]. The maximization of the ELBO implicitly solves two optimization problems concurrently: (i) maximizing the log evidence  $\log p_\theta(x)$  and (ii) minimizing the divergence  $D_{\text{KL}}(q_\phi(\cdot|x) || p_\theta(\cdot|x))$  between the variational posterior and the true posterior. **Amortized variational inference** [2] proposes to learn one global mapping in  $x$  which maps each observation to a distribution in the latent space that approximates  $z \mapsto p_\theta(z|x)$ . The parameters  $\phi$  are thus shared among all the distributions  $\{q_\phi(\cdot|x_1), \dots, q_\phi(\cdot|x_N)\}$ . This is useful as practically we typically have a large dataset  $x_1, \dots, x_N$  and the resulting ELBO is a sum of  $N$  terms.

If the density of  $q_\phi(\cdot|x)$  is available in closed form, then the  $\mathcal{Q}$  is called explicit. Explicit models allow straightforward estimation of the VI target, but often result in reduced flexibility, which limits their overall performance. For example, the mean-field VI [8] requires restrictive independence assumptions among the latent variables of interest:  $\mathcal{Q} = \{q_\phi(\cdot|x) = N(\mu_\phi(x), \text{diag}(\sigma^2(x)))\}$ , where  $N(\mu, \Sigma)$  denotes the  $d$ -dimensional normal distribution with mean  $\mu$  and covariance  $\Sigma$ . In that case, the model is the very popular variational autoencoder (VAE) [2–4]. **Normalizing flows (NFs)** [1,4,9,11] provide an alternative family of explicit density models that have improved power compared with mean-field alternatives. These methods push samples

from a simple base distribution (typically Gaussian) through parameterized diffeomorphisms to produce complex density models. NFs have proven useful; e.g. [9,10,12,13] where flows have been shown to improve ELBO tightness. Although NFs can directly improve the expressiveness of mean-field VI schemes, their inherent bijectivity impose limitations, see [14–16].

We focus here on a complementary approach to obtain tight ELBOs that builds upon Monte Carlo methods and can be traced back to Mnih & Rezende [17]. This general framework is described in §2 where a basic importance sampling approach is introduced for illustration. We then explain how one can generalize this approach to state-of-the-art Monte Carlo estimates of the evidence including **annealed importance sampling (AIS)** and more general **sequential importance sampling (SIS)** schemes in §3. Finally, sequential Monte Carlo are discussed in §4.

## 2. Monte Carlo for variational inference

We first note that the standard ELBO (1.3) can be thought of as the expectation w.r.t. to  $q_\phi(z|x)$  of a one sample unbiased importance sampling estimate of  $p_\theta(x)$

$$\hat{p}_\theta(x, z) := w_{\theta, \phi}(z) \quad \text{and} \quad w_{\theta, \phi}(z) = \frac{p_\theta(x, z)}{q_\phi(z|x)}. \quad (2.1)$$

However, as pointed out by Mnih & Rezende [17], there is no need to limit ourselves to such an estimate. As long as we have access to a positive unbiased estimate  $\hat{p}_\theta(x, u)$  of  $p_\theta(x)$  where  $u$  represents all the random variables of density  $q(u)$ <sup>1</sup> necessary to compute this estimate then we can also define an ELBO through

$$\mathcal{L}(\theta, \phi) = \int \log \hat{p}_\theta(x, u) q(u) du \leq \log \left( \int \hat{p}_\theta(x, u) q(u) du \right) = \ell(\theta, x), \quad (2.2)$$

where we have used first Jensen's inequality then the unbiasedness of the evidence estimate. This approach is interesting as a Taylor expansion shows that

$$\mathcal{L}(\theta, \phi) \approx \ell(\theta, x) - \frac{1}{2} \text{var}_q \left[ \frac{\hat{p}_\theta(x, u)}{p_\theta(x)} \right]; \quad (2.3)$$

e.g. [16,18,19]. So if one can obtain an unbiased estimate of the evidence which has a small variance relative to  $p_\theta(x)^2$ , then the resulting ELBO will be tight.

As a first illustration of this principle, consider the case where we compute an estimate of the evidence using not one but  $K$  samples; i.e.  $u = (z_1, \dots, z_K)$ ,  $q(u) = \prod_{k=1}^K q_\phi(z_k|x)$  and

$$\hat{p}_\theta(x, u) := \frac{1}{K} \sum_{k=1}^K w_{\theta, \phi}(z_k). \quad (2.4)$$

The resulting ELBO (2.2) denoted  $\mathcal{L}_{\text{iw}}(\theta, \phi)$  is known as the **importance weighted ELBO (IW-ELBO)**. It was proposed by Burda *et al.* [20], who showed that it is monotonically increasing with  $K$  and converges asymptotically to  $\ell(\theta, x)$  under mild assumptions.

In this case, we have

$$\nabla_\theta \mathcal{L}_{\text{iw}}(\theta, \phi) = \mathbb{E}_q \left[ \sum_{k=1}^K \bar{w}_{\theta, \phi}^k \nabla \log p_\theta(x, z_k) \right] \quad \text{and} \quad \bar{w}_{\theta, \phi}^k = \frac{w_{\theta, \phi}(z_k)}{\sum_{l=1}^K w_{\theta, \phi}(z_l)}. \quad (2.5)$$

This shows that an unbiased gradient of this ELBO w.r.t.  $\theta$  can be obtained by building a normalized importance sampling approximation  $\hat{p}_\theta(z|x) = \sum_{k=1}^K \bar{w}_{\theta, \phi}^k \delta_{z_k}$  of  $p_\theta(z|x)$  and integrating  $\nabla_\theta \log p_\theta(x, z)$  w.r.t. it. The exact same gradient estimator would be obtained by replacing  $\hat{p}_\theta(z|x)$  in place of  $p_\theta(z|x)$  in Fisher's identity (1.2). This shows that the biased gradient approximation of the true log-likelihood  $\ell(\theta, x)$  obtained by importance sampling is nothing but an unbiased gradient of  $\mathcal{L}_{\text{iw}}(\theta, \phi)$ . The ELBO interpretation is more fruitful as it shows the exact objective one optimizes when using stochastic gradient ascent. Additionally, it also allows us

<sup>1</sup> $q(u)$  can depend on  $\theta, \phi$  and  $x$  but this is notationally omitted.

to optimize the algorithmic parameter  $\phi$ . However, standard gradient estimates of this ELBO for  $\phi$  have poor properties as  $K$  increases due to the fact that  $\mathcal{L}_{\text{iw}}(\theta, \phi)$  becomes then almost independent of  $\phi$  (because of the convergence of the ELBO to the true log marginal likelihood), and thus decreasingly strongly the signal-to-noise ratio of the estimator of  $\nabla_{\phi} \mathcal{L}_{\text{iw}}(\theta, \phi)$  [21]. One can remedy this behaviour by using a ‘sticking the landing’ variance reduction approach [22] or by doubly reparameterizing the gradient estimates of  $\nabla_{\phi} \mathcal{L}_{\text{iw}}(\theta, \phi)$  [23].

In this section, we have presented a generic way to obtain an ELBO from any unbiased positive estimate of the evidence. However, we have so far limited ourselves practically to a basic importance sampling estimate of the evidence which performs poorly in high dimension. In the next sections, we discuss more sophisticated Monte Carlo techniques.

### 3. Annealed and sequential importance sampling

The problem of standard importance sampling techniques is that it is difficult to build a ‘good’ proposal distribution in high-dimensional spaces. A popular approach addressing this problem is AIS, an approach pioneered by Crooks [24] and Neal [25] based on early work by Jarzynski [26]. Despite having been introduced over 20 years ago, this approach remains one of the ‘gold standard’ techniques to estimate the evidence unbiasedly and can thus be used to define an ELBO [27,28]. However, as detailed below, it is difficult to obtain low-variance gradient estimators of this ELBO so a generalized version of the AIS estimate based on SIS [29,30] has been instead favoured [14–16,31,33].

For ease of presentation, we avoid here measure-theoretic notation but some of the Markov kernels discussed here do not necessarily admit a density w.r.t. Lebesgue measure; e.g. Metropolis–Hastings kernels admit an atomic component. We refer the reader to Thin *et al.* [27] for a rigorous presentation.

#### (a) Evidence estimate using sequential importance sampling

The first ingredient of these techniques is to define a sequence of  $T+1$  unnormalized density functions  $\gamma_0, \dots, \gamma_T$  bridging smoothly an easy-to-sample distribution  $\gamma_0(z) = q_{\phi}(z | x)$  (the dependency in  $x$  is omitted to avoid overloading the notation) to the unnormalized target of interest  $\gamma_T(z) = p_{\theta}(x, z)$ . Typically one selects

$$\gamma_t(z) = q_{\phi}(z | x)^{1-\beta_t} p_{\theta}(x, z)^{\beta_t} \quad \text{and} \quad \pi_t(z) = \frac{\gamma_t(z)}{\int \gamma_t(z') dz'}, \quad (3.1)$$

for an annealing schedule  $0 = \beta_0 < \dots < \beta_T = 1$ . Typically, either the  $(\beta)_{t=0}^T$  are learned, or following Grosse *et al.* [34], we can select them as  $\beta_t = \tilde{\beta}_t - \tilde{\beta}_0 / (\tilde{\beta}_T - \tilde{\beta}_0)$ , with  $\tilde{\beta}_t = \sigma(\delta(2t - T))$ , where  $\delta$  can be fixed or optimized and  $\sigma$  is the sigmoid function. The second ingredient of these techniques is to define a sequence of forward Markov transition kernels  $m_t(z_{t-1}, z_t)$ , designed to “move” samples towards  $\pi_t$ . The most common choice is to select  $m_t$  as a MCMC kernel leaving  $\pi_t$  invariant. However, it is not necessary that  $m_t$  be exactly invariant for  $\pi_t$ : it is perfectly possible to use a kernel that leaves  $m_t$  ‘approximately’ invariant. As a result  $m_t$  depends on  $\theta, \phi$  and  $x$  but this is not emphasized notationally for simplicity. Finally, we present here for simplicity our results for distributions admitting densities w.r.t. the Lebesgue measure, but the reasoning in itself does not depend on this assumption, see [16,27] for a more general presentation.

Based on these two ingredients, we can build the following proposal distribution on the path space  $z_{0:T} \in \mathcal{Z}^{T+1}$

$$\tilde{q}(z_{0:T}) = q_{\phi}(z_0 | x) \prod_{t=1}^T m_t(z_{t-1}, z_t). \quad (3.2)$$

By construction, the  $T$ th marginal  $q_T(z_T) = \int \tilde{q}(z_{0:T}) dz_{0:T-1}$  is expected to be close to  $p_{\theta}(z | x)$  for  $T$  large enough and MCMC kernels mixing fast enough. If we were able to evaluate this marginal

density pointwise, then we could compute the following unbiased importance sampling estimate of the evidence

$$w_{\theta,\phi}^{\text{mar}}(z_T) = \frac{p_{\theta}(x, z_T)}{q_T(z_T)}. \quad (3.3)$$

However, the expression of this marginal density is generally not available. To bypass this issue, we introduce another unnormalized density on the path space

$$\bar{p}(x, z_{0:T}) = p_{\theta}(x, z_T) \prod_{t=0}^{T-1} l_t(z_{t+1}, z_t), \quad (3.4)$$

where  $l_t(z_{t+1}, z_t)$  is a sequence of auxiliary backward Markov transition kernels; i.e.  $\int l_t(z_{t+1}, z_t) dz_t = 1$ . These kernels will also depend typically on  $\theta, \phi$  and  $x$  but this is omitted notationally. By construction, we thus have  $\int \bar{p}(x, z_{0:T}) dz_{0:T-1} = p_{\theta}(x, z_T)$ .

Having now defined both a proposal (3.2) and an extended target (3.4) on path space, we can define the following SIS estimate of the evidence [29,30]

$$w_{\theta,\phi}^{\text{sis}}(z_{0:T}) = \frac{\bar{p}(x, z_{0:T})}{\bar{q}(z_{0:T})} = \prod_{t=1}^T \frac{\gamma_t(z_t) l_{t-1}(z_t, z_{t-1})}{\gamma_{t-1}(z_{t-1}) m_t(z_{t-1}, z_t)}. \quad (3.5)$$

This is a special case of the general framework introduced in §2 where  $u = z_{0:T}$ ,  $q(u) = \bar{q}(z_{0:T})$  and  $\hat{p}_{\theta}(x, u) := w_{\theta,\phi}^{\text{sis}}(z_{0:T})$ . As in §2, we can further average over  $N$  samples of  $\bar{q}(z_{0:T})$ .

## (b) Settings

In practice, we need to select appropriately the forward and backward kernels.

### (i) Annealed importance sampling

The standard AIS approach makes the following choice. First one selects  $m_t$  to be MCMC kernels of invariant distribution  $\pi_t$ ; e.g. Metropolis–Hastings. Then we select the corresponding backward kernel  $l_{t-1}$  as the reversal of  $m_t$ , i.e.

$$\pi_t(z_{t-1}) m_t(z_{t-1}, z_t) = \pi_t(z_t) l_{t-1}(z_t, z_{t-1}), \quad (3.6)$$

so in particular we have  $l_{t-1} = m_t$  if  $m_t$  is reversible w.r.t.  $\pi_t$ . Implicitly, the reversal of  $m_t$  means the kernel associated with the time-reversed process of that with transition dynamics given by  $m_t$ , that begins in stationarity. When (3.6) is satisfied, we can easily check that the evidence estimate equation (3.5) simplifies and becomes equal to

$$w_{\theta,\phi}^{\text{ais}}(z_{0:T}) = \prod_{t=1}^T \frac{\gamma_t(z_{t-1})}{\gamma_{t-1}(z_{t-1})}. \quad (3.7)$$

This estimate is elegant but it is unfortunately difficult to obtain a low-variance estimate of the gradient of this ELBO. This is due to the fact that the transition kernels  $m_t$  will typically involve accept/reject steps.

Consider for example that  $m_t$  is a Metropolis–Hastings type kernel. We follow here the derivation by Thin *et al.* [16]. If we use a reparameterized proposal then we can write

$$m_t(z_{t-1}, z_t) = \int n_t((z_{t-1}, u), z_t) g(u) du, \quad (3.8)$$

where  $g(u)$  is typically a standard multivariate normal and

$$n_t((z_{t-1}, u), z_t) = \alpha_t(z_{t-1}, u) \delta_{S_t(z_{t-1}, u)}(z_t) + \{1 - \alpha_t(z_{t-1}, u)\} \delta_{z_{t-1}}(z_t), \quad (3.9)$$

where  $\delta_z(z')$  denotes the Dirac delta function, and  $S_t(z_{t-1}, u)$  denotes the Metropolis–Hastings proposal from point  $z_{t-1}$  with noise  $u$  (sampled from the distribution  $g$ ). For example, if we consider  $m_t(z_{t-1}, z_t)$  a random walk Metropolis algorithm of proposal  $N(z'; z_{t-1}, \sigma_t^2 I)$  then we would have  $S_t(z_{t-1}, u) = z_{t-1} + \sigma_t u$  and  $\alpha_t(z_{t-1}, u) = \min(1, \gamma_t(S_t(z_{t-1}, u)) / \gamma_t(z_{t-1}))$ , with  $g$

the standard multivariate normal. With the convention  $\alpha_t(z_{t-1}, u) = \alpha_t^1(z_{t-1}, u)$ ,  $1 - \alpha_t(z_{t-1}, u) = \alpha_t^0(z_{t-1}, u)$ ,  $S_t^1(z_{t-1}, u) = S_t(z_{t-1}, u)$  and  $S_t^0(z_{t-1}, u) = z_{t-1}$ , we can thus rewrite

$$n_t((z_{t-1}, u), z_t) = \sum_{a_t=0}^1 \alpha_t^{a_t}(z_{t-1}, u) \delta_{S_t^{a_t}(z_{t-1}, u)}(z_t). \quad (3.10)$$

Piecing the  $T$  steps together, it thus follows that the proposal in equation (3.2) can be summarized as first draw the noise for the  $T$  steps from  $g(u_{1:T}) = \prod_{t=1}^T g(u_t)$ , and sequentially for  $1 \leq t \leq T$ , (i) conditionally upon  $u_{1:T}$ , draw the Bernoulli r.v.s  $a_t$  with  $\mathbb{P}(a_t = v \mid z_0, u_{1:t-1}, a_{1:t-1}) = \alpha_t^v(z_{t-1}, u_t)$  for  $v \in \{0, 1\}$  and, (ii) given  $z_0, u_{1:t}, a_{1:t}$ , compute the state  $z_t$  given by

$$z_t = \Phi_t(z_0, u_{1:t}, a_{1:t}) = S_t^{a_t}(z_{t-1}, u_t) = S_t^{a_t}(\Phi_{t-1}(z_0, u_{1:t-1}, a_{1:t-1}), u_t) \quad (3.11)$$

with  $\Phi_0(z_0, u_{1:0}, a_{1:0}) = z_0$ . Then, the joint distribution of the accept/reject binary random variables  $a_{1:T}$  is given by  $\beta(a_{1:T} \mid z_0, u_{1:T}) = \prod_{t=1}^T \alpha_t^{a_t}(\Phi_{t-1}(z_0, u_{1:t-1}, a_{1:t-1}), u_t)$ . The proposal  $\tilde{q}(z_{0:T})$  in equation (3.2) is thus equal to

$$q_\phi(z_0 \mid x) \int \cdots \int \sum_{a_{1:T} \in \{0,1\}^T} g(u_{1:T}) \beta(a_{1:T} \mid z_0, u_{1:T}) \\ \times \prod_{t=1}^T \delta_{S_t^{a_t}(\Phi_{t-1}(z_0, u_{1:t-1}, a_{1:t-1}), u_t)}(x_t) \, du_{1:T}, \quad (3.12)$$

and finally the ELBO corresponding to the AIS estimate in equation (3.7) can be rewritten as

$$\mathcal{L}_{\text{ais}} = \mathbb{E}_{\tilde{q}(z_{0:T})} [\log w_{\theta, \phi}^{\text{ais}}(z_{0:T})] = \mathbb{E}_{q_\phi(z_0 \mid x) g(u_{1:T}) \beta(a_{1:T} \mid z_0, u_{1:T})} [\log w_{\theta, \phi}^{\text{ais}}(z_{0:T})], \quad (3.13)$$

where the states  $z_{1:T}$  are fully deterministic functions of  $z_0, u_{1:T}, a_{1:T}$ , denoted  $z_{1:T} = \Phi(z_0, u_{1:T}, a_{1:T})$ . It follows that the gradient of this ELBO w.r.t.  $\phi, \theta$  is given by

$$\nabla \mathcal{L}_{\text{ais}} = \nabla \mathcal{L}_{\text{ais}}^{(1)} + \nabla \mathcal{L}_{\text{ais}}^{(2)}, \quad (3.14)$$

where

$$\left. \begin{aligned} \nabla \mathcal{L}_{\text{ais}}^{(1)} &= \mathbb{E}[\nabla \log w_{\theta, \phi}^{\text{ais}}(z_0, \Phi(z_0, u_{1:T}, a_{1:T}))], \\ \text{and} \quad \nabla \mathcal{L}_{\text{ais}}^{(2)} &= \mathbb{E}[\nabla \log \beta(a_{1:T} \mid z_0, u_{1:T}) \cdot \log w_{\theta, \phi}^{\text{ais}}(z_0, \Phi(z_0, u_{1:T}, a_{1:T}))]. \end{aligned} \right\} \quad (3.15)$$

The first term  $\nabla \mathcal{L}_{\text{ais}}^{(1)}$  can typically be estimated with low variance but the second term  $\nabla \mathcal{L}_{\text{ais}}^{(2)}$  has a higher variance estimate due to the gradient estimate for the accept/reject discrete random variables  $\nabla \log \beta(a_{1:T} \mid z_0, u_{1:T})$  [35], see [27] for a further discussion. In [28], the second term is neglected, which introduces a bias that is difficult to quantify. In [16], a control variate technique inspired by Mnih & Rezende [17, Section 2.5.3] is developed to reduce the variance of the second term but it remains high.

## (ii) Annealed importance sampling with unadjusted proposals

While the standard AIS estimate can be used to obtain an ELBO, we have seen that the resulting ELBO is difficult to optimize as AIS typically relies on MCMC kernels using accept/reject steps. To bypass this issue, it is possible instead to use unadjusted samplers such as the unadjusted Langevin algorithm (ULA) [16,28,36] and the unadjusted Hamiltonian algorithm (UHA) [28,31,33–37].

We start by considering the use of ULA transition kernels; i.e. we let

$$m_t(z_t \mid z_{t-1}) = N(z_t \mid z_{t-1} + \delta \nabla \log \pi_t(z_{t-1}), 2\delta \text{Id}). \quad (3.16)$$

The resulting proposal equation (3.2) can thus be thought of as an Euler–Maruyama discretization of the following time-inhomogeneous Langevin diffusion on the time interval  $[0, T_0]$  for  $\delta = T_0/T$

$$dz_s = \nabla \log \pi_s(z_s) \, ds + \sqrt{2} \, dB_s, \quad (3.17)$$



where  $(B_s)_{s \in [0, T_0]}$  is a multivariate Brownian motion. We slightly abuse notation here as  $\pi_t$  in (3.16) corresponds to  $\pi_{t\delta}$  in continuous time. Since the continuous-time homogeneous Langevin diffusion is known to be reversible, a reasonable choice for the backward kernel is to select  $l_{t-1} = m_t$ , as suggested in [27,28,36]. However, as  $m_t$  is not exactly  $\pi_t$ -reversible, we need to rely on the general expression given in (3.5) for the evidence estimate. As this proposal can be easily reparameterized as a function of the Gaussian random variables used to sample from  $m_t$ , we can apply the reparameterization trick to compute low-variance gradient estimates of the corresponding ELBO. It was shown in [16] that, compared with a standard AIS estimate using Metropolis-adjusted Langevin kernels, the ELBO computed using ULA kernels was not as tight but it was much easier to optimize.

Instead of having a proposal arising from the time-discretization of an overdamped Langevin diffusion (3.17), we can instead build a proposal by discretizing an underdamped Langevin diffusion initialized at  $z_0 \sim q_\phi(\cdot | x)$ ,  $p_0 \sim N(0, M)$

$$dx_s = M^{-1}p_s ds \quad \text{and} \quad dp_s = \nabla \log \pi_s(z_s) ds - \eta p_s ds + \sqrt{2\eta} M^{1/2} dB_s. \quad (3.18)$$

Here  $p_s \in \mathbb{R}^{d_z}$  is a momentum variable,  $M$  a positive definite mass matrix and  $\eta > 0$  a friction coefficient. If we had  $\pi_s = \pi$  then the invariant distribution of this diffusion would be given by  $\pi(z)N(p; 0, M)$ . This underdamped Langevin diffusion corresponds to a continuous-time version of Hamiltonian dynamics for which the momentum component is continuously refreshed using an Ornstein–Uhlenbeck process. In this case, we can use the following integrator to discretize the diffusion equation (3.18)

$$\tilde{p}_{t+1} \sim N(hp_t, (1 - h^2)M) \quad \text{and} \quad (z_{t+1}, p_{t+1}) = \text{LF}_t(z_t, \tilde{p}_{t+1}), \quad (3.19)$$

with  $h = \exp(-\eta\delta)$  and  $\text{LF}_t$  is the leapfrog integrator for  $\pi_{t\delta}$ . So we have forward transition kernels of the form

$$m_{t+1}((z_t, p_t), (\tilde{p}_{t+1}, z_{t+1}, p_{t+1})) = N(\tilde{p}_{t+1}; hp_t, (1 - h^2)M) \delta_{\text{LF}_t(z_t, \tilde{p}_{t+1})}(z_{t+1}, p_{t+1}). \quad (3.20)$$

We are also selecting backward transition kernels of the form

$$l_t((z_{t+1}, p_{t+1}), (\tilde{p}_{t+1}, z_t, p_t)) = \delta_{\text{LF}_t^{-1}(z_{t+1}, p_{t+1})}(z_{t+1}, \tilde{p}_{t+1}) N(p_t; h\tilde{p}_{t+1}, (1 - h^2)M). \quad (3.21)$$

Considering now the extended proposal and target

$$\tilde{q}(z_{0:T}, p_{0:T}, \tilde{p}_{1:T}) = q_\phi(z_0 | x) N(p_0; 0, M) \prod_{t=0}^{T-1} m_{t+1}((z_t, p_t), (\tilde{p}_{t+1}, z_{t+1}, p_{t+1})), \quad (3.22)$$

and

$$\bar{p}(x, z_{0:T}, p_{0:T}, \tilde{p}_{1:T}) = p_\theta(x, z_T) N(p_T; 0, M) \prod_{t=0}^{T-1} l_t((z_{t+1}, p_{t+1}), (\tilde{p}_{t+1}, z_t, p_t)), \quad (3.23)$$

the resulting SIS unbiased evidence estimate is given by

$$\begin{aligned} w_{\theta, \phi}^{\text{sis}}(z_{0:T}) &= \frac{\bar{p}(x, z_{0:T}, p_{0:T}, \tilde{p}_{1:T})}{\tilde{q}(z_{0:T}, p_{0:T}, \tilde{p}_{1:T})} \\ &= \prod_{t=0}^{T-1} \frac{\gamma_{t+1}(z_{t+1}) N(p_{t+1}; 0, M) l_t((z_{t+1}, p_{t+1}), (\tilde{p}_{t+1}, z_t, p_t))}{\gamma_t(z_t) N(p_t; 0, M) m_{t+1}((z_t, p_t), (\tilde{p}_{t+1}, z_{t+1}, p_{t+1}))} \\ &= \frac{p_\theta(x, z_T) N(p_T; 0, M)}{q_\phi(z_0 | x) N(p_0; 0, M)} \prod_{t=0}^{T-1} \frac{N(p_t; 0, M)}{N(\tilde{p}_{t+1}; 0, M)}, \end{aligned}$$

where we have used the fact that as  $\text{LF}_t$  is a symplectic integrator, it is volume preserving and thus  $|\text{J}_{\text{LF}_t}(z, p)| = 1$ , and that  $N(p_t; 0, M) N(\tilde{p}_{t+1}; hp_t, (1 - h^2)M) = N(\tilde{p}_{t+1}; 0, M) N(p_t; h\tilde{p}_{t+1}, (1 - h^2)M)$ . As for ULA, it is possible to easily reparameterize the proposal in terms of standard Gaussian random variables. This leads to low-variance estimates of the gradients of the ELBO. Doucet *et al.* [38]

provides an empirical comparison of UHA-type and ULA-type proposals in a limited Monte Carlo experiment. The performance is slightly better for UHA than for ULA, but the difference is rather small. This finding is somewhat consistent with the results of Bou-Rabee & Eberle [37], which show that the mixing time of UHA is better than that of ULA. However, we are far from having a clear theoretical guarantee of the relative advantage of using UHA compared with ULA in the context of AIS, which is very complex.

### (iii) Optimizing backward Markov kernels

In the schemes previously described, the backward Markov kernels are selected as the reversals or approximate reversals of the forward kernels. However, it was realized early on in the literature [29] that this choice is suboptimal and the optimal backward kernels were identified. Indeed from the law of total variance, we have

$$\begin{aligned} \text{var}_{\tilde{q}(z_{0:T})}[w_{\theta,\phi}^{\text{sis}}(z_{0:T})] &= \text{var}_{\tilde{q}_T(z_T)}[w_{\theta,\phi}^{\text{mar}}(z_T)] \\ &\quad + \mathbb{E}_{\tilde{q}_T(z_T)}[\text{var}_{\tilde{q}(z_{0:T-1}|z_T)}[w_{\theta,\phi}^{\text{sis}}(z_{0:T})]], \end{aligned} \quad (3.24)$$

where  $w_{\theta,\phi}^{\text{mar}}(z_T)$  is defined in (3.3) and the second term on the r.h.s. is null for  $\tilde{p}(z_{0:T-1} | z_T, x) = \tilde{q}(z_{0:T-1} | z_T)$ . So using the backward decomposition of  $\tilde{q}(z_{0:T})$  shows that the optimal backward kernels  $l_t^{\text{opt}}$  satisfy

$$l_t^{\text{opt}}(z_{t+1}, z_t) = \frac{q_t(z_t)m_t(z_t, z_{t+1})}{q_{t+1}(z_{t+1})}, \quad (3.25)$$

where  $q_t(z_t)$  is the marginal of  $z_t$  under  $\tilde{q}(z_{0:T})$ . While this expression is interesting, it is unclear how one could come up with an approximation of these optimal backward kernels as the marginals  $q_t(z_t)$  are intractable. However, inspired by recent developments in generative modelling [39], it was recently realized in [38] that it is possible to come up with sensible approximations to these optimal kernels for ULA and UHA proposals. We restrict ourselves here to ULA proposals and refer the reader to Doucet *et al.* [38] for UHA proposals. Recall that the ULA proposal we consider can be thought of as the time-discretization of the time-inhomogeneous Langevin diffusion equation (3.17). The **time-reversed process**  $(\tilde{z}_s)_{s \in [0, T_0]}$  is also a diffusion given by

$$d\tilde{z}_s = \{-\nabla \log \pi_{T_0-s}(\tilde{z}_s) + 2\nabla \log q_{T_0-s}(\tilde{z}_s)\} ds + \sqrt{2} d\tilde{B}_s \quad \text{and} \quad \tilde{z}_0 \sim q_{T_0}, \quad (3.26)$$

where  $(\tilde{B}_s)$  is another Brownian motion and  $q_s$  denotes the law of  $x_s$  under equation (3.17). The so-called **score terms**  $\nabla \log q_s$  are intractable but this expression suggests that one should consider parameterized backward Markov transition kernels of the form

$$l_t(z_{t+1}, z_t) = N(z_t; z_{t+1} - \delta \nabla \log \pi_{t+1}(z_{t+1}) + 2\delta s_\phi(t+1, z_{t+1}), 2\delta I), \quad (3.27)$$

where  $s_\phi$  is a **neural network approximating the scores**  $(\nabla q_t)_{t=1}^T$ . The parameters of this network can be obtained by maximizing the ELBO which coincides with a denoising score matching loss [40]. This approach can improve experimentally over the standard backward Markov kernels for both ULA and UHA proposals [38] for the estimation of the normalizing constant in these contexts. While it had been previously proposed to use backward Markov kernels of the form  $l_t(z_{t+1}, z_t) = N(z_t; \mu_\phi(z_{t+1}), \Sigma_\phi(z_{t+1}))$  [30,41], where the functions  $\mu_\phi, \Sigma_\phi$  are parameterized by neural networks, this parameterization does not exploit the structure of optimal backward kernels and performs poorly in simulations [16].

### (iv) Combining stochastic proposals and normalizing flows

It is well known that adding some deterministic moves to the stochastic transition kernels can further improve the performance of SIS/AIS type algorithms. In practice, we simply interleave stochastic transitions with deterministic invertible transitions of the form  $m_t(z_{t-1}, z_t) = \delta_{T_t(z_{t-1})}(z_t)$  with corresponding backward kernel  $l_{t-1}(z_t, z_{t-1}) = \delta_{T_t^{-1}(z_t)}(z_{t-1})$ . This was proposed early on by



Vaikuntanathan & Jarzynski [42]; see also [43]. In [28], these deterministic moves  $T_t$  are built using NFs whose parameters are learned by maximizing the ELBO.

#### (v) Limitations

The main limitation of the methods is that the unadjusted samplers considered here can be much more unstable than their Metropolis-adjusted counterparts. They also suffer from high memory consumption as they require storing the whole simulated Markov chain. As a result, one is limited to use short chains which limit their performance.

## 4. Further extensions

### (a) Sequential Monte Carlo

For SIS-type methods, degeneracy of importance weights typically occurs. SMC samplers, originally proposed by Del Moral *et al.* [29], introduce additional resampling, either at each step or adaptively using a non-degeneracy criterion for importance weights to mitigate this problem. The resampling has the effect of providing an unweighted cloud of particles distributed approximately according to  $\pi_t$ . Note, however, that using resampling steps for estimating normalizing constants is not always beneficial; see [29, Section 4.2.3.2]. As SMC samplers provide also an unbiased evidence of the evidence [44], they can be used to provide an ELBO. This was noted by Maddison *et al.* [19], Le *et al.* [45] and Naesseth *et al.* [46]. However, the resulting ELBO is difficult to maximize. Indeed, the resampling steps of SMC involve sampling discrete distributions. Hence, the variance of the gradient estimates is very large as we have to use REINFORCE gradient estimates, similarly to (3.14). Hence it was proposed in [19,45,46] to use biased gradient estimates which neglect the terms resorting to a REINFORCE gradient estimate. Unfortunately, as established by Corenflos *et al.* [47], this can introduce some very substantial bias. To address this problem, **differentiable resampling procedures** have been introduced; e.g. [47,48]. In these approaches, resampling is replaced by solving a regularized optimal transport problem. These methods perform well but are computationally quite expensive and have only been demonstrated for low-dimensional filtering problems where  $d_z$  is moderate. In the DLVM settings where  $d_z$  can be of the order of a hundred, it is not expected that they will fare well. One way to get around these issues is to abandon the ELBO for alternative loss functions.

In [49], the authors consider a SMC sampler combined with NFs. A NF is introduced after each resampling step to push particles towards the next target distribution. To learn these NFs, it is proposed to minimize

$$\sum_{t=1}^T D_{\text{KL}}[T_t^\# \pi_{t-1} || \pi_t], \quad (4.1)$$

where  $T_t^\# \pi_{t-1}$  denotes the push-forward of  $\pi_{t-1}$ ; see also [50]. This criterion and its gradient can be estimated using the particle approximation of  $\pi_{t-1}$ . A stop gradient operator is used to avoid differentiating through the SMC procedure, thus avoiding the increase in variance caused by the approximation of the differentiation through the resampling steps. Using a challenging example from lattice field theory and VAEs, the method has been shown to outperform alternatives such as the schemes proposed in [28,51]. The advantage of this method is that, unlike [19,45,46], it does not require differentiation of the SMC process.

Midgley *et al.* [52] propose an alternative approach to train a normalizing  $q_\phi$  flow to approximate a target  $\mu$ . Instead of using the mode seeking reverse KL to train the flow, one minimizes the variance of the importance weight  $p_\theta(z | x)/q_\phi(z)$ . This criterion and its gradient w.r.t.  $\phi$  is approximated using a standard Metropolis-adjusted AIS sampler targeting the optimal importance distribution  $\pi(z) \propto p_\theta(z | x)^2 / q_\phi(z)$ —an SMC sampler would also be applicable. A stop gradient operator is used to avoid differentiating through the AIS procedure.

## (b) Alternative unbiased estimates of the evidence

While most work on marginal likelihood estimation (to construct a valid ELBO) has focused on importance sampling methods and its variants, other approaches have recently been proposed. Building on the work of Rotskoff & Vanden-Eijnden [53], Thin *et al.* [54], for example, have proposed a novel unbiased estimate of evidence based on deterministic trajectories starting from points sampled by an importance distribution  $q_\phi(z|x)$ . Starting from a well-chosen transformation  $T$  (typically a conformal Hamiltonian dynamics), the estimator  $w_{\theta,\phi}^{\text{neo}}(z)$  is given by

$$w_{\theta,\phi}^{\text{neo}}(z) = \sum_{k=0}^K \varpi_k(z) w_{\theta,\phi}(T^k(z)), \quad (4.2)$$

where the weights  $\varpi_k(z) = q_\phi(T^k(z)|x)J_{T^k}(z)/\{\sum_{i=-k}^{K-k} q_\phi(T^{-i}(z)|x)J_{T^{-i}}(z)\}$  ensure unbiasedness of the estimator  $w_{\theta,\phi}^{\text{neo}}(z)$  when  $z \sim q_\phi(\cdot|x)$ . This estimator can be easily applied with the reparameterization trick [2] and leads to a differentiable ELBO. It has been successfully used by Thin *et al.* [54] to perform inference for a variety of latent variable models.

## (c) Other related works

We have mainly focused on variational SIS based approaches compared with other classical homogeneous MCMC methods which do not provide an estimate of the evidence. Hoffman [1] considers another type of DLGM learnt with Hamiltonian Monte Carlo (HMC), by decoupling the ELBO, optimizing it in  $\phi$  while approximating (1.2) using an HMC chain initialized at  $z_0 \sim q_\phi(\cdot|x)$ . This method, however, differs from the ones presented above, as the effective variational distribution built cannot be cast into a single objective maximization. Another related approach is the Hamiltonian VAE [55], which improves the variational family using a well-designed flow inspired by Hamiltonian dynamics. It is, however, more closely related to NFs than to MCMC samplers. Another relevant direction of work is unbiased and biased-reduced estimation of the gradient of the log-likelihood (1.2), as investigated in [56,57]. Both works extend the IWAE methodology by writing an iterated Sampling Importance Resampling scheme (i-SIR [58,59]) to re-interpret the IWAE ELBO on an extended space and use either coupled Markov chains methodology to obtain unbiased estimates of (1.2) or biased reduction techniques to improve on the IWAE's methodology, viewing (2.5) as a self normalized importance sampling estimate.

## 5. Discussion

The marginal likelihood/evidence for deep latent variable models is intractable but low-variance and unbiased Monte Carlo estimates of the evidence can be used to define a tight ELBO. However, because we are interested in maximizing such bounds with respect to both the model parameters  $\theta$  and the variational parameters  $\phi$ , it is also necessary to have access to low-variance gradient estimates of this ELBO. Unfortunately, most standard procedures used in the statistics literature, such as AIS and SMC, provide high variance gradient estimates because they are based on accept/reject steps and/or resampling steps. To address this problem, variants of AIS based on general SIS have been proposed based on unadjusted samplers such as ULA and UHA. These provide low-variance gradient estimates through the trick of reparameterization, but at the cost of a loss of numerical stability compared with their Metropolis-adjusted counterparts. Differentiable estimators of the evidence and the ELBO have also been proposed for SMC, but these are limited to low-to-medium dimension latent variables. For high dimensional scenarios, SMC estimators of ELBO remain difficult to optimize and it may be better to use alternative criteria.

**Data accessibility.** This article has no additional data.

**Authors' contributions.** A.D.: investigation, methodology, project administration, writing—original draft, writing—review and editing; E.M.: conceptualization, investigation, methodology, project administration, writing—original draft, writing—review and editing; A.T.: writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

**Conflict of interest declaration.** We declare we have no competing interests.

**Funding.** Part of this research has been supported by ANR-19-CHIA-SCAI-002 and has been carried out under the auspices of Lagrange research Center for Mathematics and Calculus. AD acknowledges support of EPSRC grants CoSines (EP/R034710/1) and Bayes4Health (EP/R018561/1).

## References

1. Hoffman MD. 2017 Learning deep latent Gaussian models with Markov chain Monte Carlo. In *Proceedings of the 34th International Conference on Machine Learning*, pp. 1510–1519, vol. 70. PMLR.
2. Kingma DP, Welling M. 2014 Auto-encoding variational Bayes. In *Int. Conf. on Learning Representations*.
3. Kingma DP, Welling M. 2019 An introduction to variational autoencoders. Preprint (<https://arxiv.org/abs/1906.02691>).
4. Rezende DJ, Mohamed S. 2015 Variational inference with normalizing flows. In *Proceedings of the 32nd International Conference on Machine Learning*, pp. 1530–1538, vol. 37. PMLR.
5. Jordan MI, Ghahramani Z, Jaakkola TS, Saul LK. 1999 An introduction to variational methods for graphical models. In *Learning in Graphical Models*, pp. 105–161. Cambridge, MA: MIT Press.
6. Rezende DJ, Mohamed S, Wierstra D. 2014 Stochastic backpropagation and approximate inference in deep generative models. In *Proceedings of the 31st International Conference on Machine Learning*, pp. 1278–1286, vol. 32. PMLR.
7. Yin M, Zhou M. 2018 Semi-implicit variational inference. In *Proceedings of the 35th International Conference on Machine Learning*, pp. 5660–5669, vol. 80. PMLR.
8. Blei D, Kucukelbir A, McAuliffe J. 2017 Variational inference: a review for statisticians. *J. Am. Stat. Assoc.* **112**, 859–877. (doi:10.1080/01621459.2017.1285773)
9. Papamakarios G, Nalisnick E, Rezende DJ, Mohamed S, Lakshminarayanan B. 2021 Normalizing flows for probabilistic modeling and inference. *J. Mach. Learn. Res.* **22**, 2617–2680.
10. Papamakarios G, Pavlakou T, Murray I. 2017 Masked autoregressive flow for density estimation. In *Advances in Neural Information Processing Systems* 30, pp. 2335–2344. Curran Associates, Inc.
11. Tabak EG, Vanden-Eijnden E. 2010 Density estimation by dual ascent of the log-likelihood. *Commun. Math. Sci.* **8**, 217–233. (doi:10.4310/CMS.2010.v8.n1.a11)
12. Ho J, Chen X, Srinivas A, Duan Y, Abbeel P. 2019 Flow++: improving flow-based generative models with variational dequantization and architecture design. Preprint (<https://arxiv.org/abs/1902.00275>).
13. Louizos C, Welling M. 2017 Multiplicative normalizing flows for variational Bayesian neural networks. In *Proceedings of the 34th International Conference on Machine Learning*, pp. 2218–2227, vol. 70. PMLR.
14. Caterini A, Cornish R, Sejdinovic D, Doucet A. 2021 Variational inference with continuously-indexed normalizing flows. In *Proceedings of the Thirty-Seventh Conference on Uncertainty in Artificial Intelligence*, pp. 44–53, vol. 61. PMLR.
15. Cornish R, Caterini A, Deligiannidis G, Doucet A. 2020 Relaxing bijectivity constraints with continuously indexed normalising flows. In *Proceedings of the 37th International Conference on Machine Learning*, pp. 2133–2143, vol. 119. PMLR.
16. Thin A, Kotelevskii N, Denain J-S, Grinsztajn L, Durmus A, Panov M, Moulines E. 2020 MetFlow: a new efficient method for bridging the gap between Markov chain Monte Carlo and variational inference. Preprint (<https://arxiv.org/abs/2002.12253>).
17. Mnih A, Rezende D. 2016 Variational inference for Monte Carlo objectives. In *Proceedings of The 33rd International Conference on Machine Learning*, pp. 2188–2196, vol. 48. PMLR.
18. Domke J, Sheldon DR. 2018 Importance weighting and variational inference. In *Advances in Neural Information Processing Systems* 31, pp. 4475–4484. Curran Associates, Inc.
19. Maddison CJ, Lawson D, Tucker G, Heess N, Norouzi M, Mnih A, Doucet A, Teh YW. 2017 Filtering variational objectives. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc.

20. Burda Y, Grosse R, Salakhutdinov R. 2016 Importance weighted autoencoders. In *Int. Conf. on Learning Representations*.
21. Rainforth T, Kosiorek A, Le TA, Maddison C, Igl M, Wood F, Teh YW. 2018 Tighter variational bounds are not necessarily better. In *Proceedings of the 35th International Conference on Machine Learning*, pp. 4277–4285, vol. 80. PMLR.
22. Roeder G, Wu Y, Duvenaud DK. 2017 Sticking the landing: simple, lower-variance gradient estimators for variational inference. In *Advances in Neural Information Processing Systems 31*, pp. 6928–6937. Curran Associates, Inc.
23. Tucker G, Lawson D, Gu S, Maddison CJ. 2019 Doubly reparameterized gradient estimators for Monte Carlo objectives. In *Int. Conf. on Learning Representations*.
24. Crooks GE. 1998 Nonequilibrium measurements of free energy differences for microscopically reversible Markovian systems. *J. Stat. Phys.* **90**, 1481–1487. (doi:10.1023/A:1023208217925)
25. Neal RM. 2001 Annealed importance sampling. *Stat. Comput.* **11**, 125–139. (doi:10.1023/A:1008923215028)
26. Jarzynski C. 1997 Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* **78**, 2690–2693. (doi:10.1103/PhysRevLett.78.2690)
27. Thin A, Kotelevskii N, Durmus A, Panov M, Moulines E, Doucet A. 2021 Monte Carlo variational auto-encoders. In *Proceedings of the 38th International Conference on Machine Learning*, pp. 10247–10257, vol. 139. PMLR.
28. Wu H, Köhler J, Noé F. 2020 Stochastic normalizing flows. In *Advances in Neural Information Processing Systems 34*, pp. 5933–5944. Curran Associates, Inc.
29. Del Moral P, Doucet A, Jasra A. 2006 Sequential Monte Carlo samplers. *J. R. Stat. Soc. B Stat. Methodol.* **68**, 411–436. (doi:10.1111/j.1467-9868.2006.00553.x)
30. Salimans T, Kingma D, Welling M. 2015 Markov chain Monte Carlo and variational inference: bridging the gap. In *Proceedings of the 32nd International Conference on Machine Learning*, pp. 1218–1226, vol. 37. PMLR.
31. Dai C, Heng J, Jacob PE, Whiteley N. 2022 An invitation to sequential Monte Carlo samplers. *J. Am. Stat. Assoc.* **117**, 1587–1600. (doi:10.1080/01621459.2022.2087659)
32. Geffner T, Domke J. 2021 MCMC variational inference via uncorrected Hamiltonian annealing. In *Advances in Neural Information Processing Systems 35*, pp. 639–651. Curran Associates, Inc.
33. Zhang G, Hsu K, Li J, Finn C, Grosse RB. 2021 Differentiable annealed importance sampling and the perils of gradient noise. In *Advances in Neural Information Processing Systems 35*, pp. 19398–19410. Curran Associates, Inc.
34. Grosse RB, Ghahramani Z, Adams RP. 2015 Sandwiching the marginal likelihood using bidirectional Monte Carlo. Preprint (<https://arxiv.org/abs/1511.02543>).
35. Williams RJ. 1992 Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Mach. Learn.* **8**, 229–256.
36. Heng J, Bishop AN, Deligiannidis G, Doucet A. 2020 Controlled sequential Monte Carlo. *Ann. Stat.* **48**, 2904–2929. (doi:10.1214/19-AOS1914)
37. Bou-Rabee N, Eberle A. 2023 Mixing time guarantees for unadjusted Hamiltonian Monte Carlo. *Bernoulli* **29**, 75–104. (doi:10.3150/21-BEJ1450)
38. Doucet A, Grathwohl WS, Matthews AGDG, Strathmann H. 2022 Score-based diffusion meets annealed importance sampling. In *Advances in Neural Information Processing Systems 36*. Curran Associates, Inc.
39. Song Y, Sohl-Dickstein J, Kingma DP, Kumar A, Ermon S, Poole B. 2021 Score-based generative modeling through stochastic differential equations. In *Int. Conf. on Learning Representations*.
40. Vincent P. 2011 A connection between score matching and denoising autoencoders. *Neural Comput.* **23**, 1661–1674. (doi:10.1162/NECO\_a\_00142)
41. Huang C-W, Tan S, Lacoste A, Courville AC. 2018 Improving explorability in variational inference with annealed variational objectives. In *Advances in Neural Information Processing Systems 32*, pp. 9724–9734. Curran Associates, Inc.
42. Vaikuntanathan S, Jarzynski C. 2011 Escorted free energy simulations. *J. Chem. Phys.* **134**, 054107. (doi:10.1063/1.3544679)
43. Heng J, Doucet A, Pokern Y. 2021 Gibbs flow for approximate transport with applications to Bayesian computation. *J. R. Stat. Soc. B Stat. Methodol.* **83**.

44. Del Moral P. 2004 *Feynman-Kac Formulae: Genealogical and Interacting Particle Approximations*. Berlin, Germany: Springer.
45. Le TA, Igl M, Rainforth T, Jin T, Wood F. 2018 Auto-encoding sequential Monte Carlo. In *Int. Conf. on Learning Representations*.
46. Naesseth CA, Linderman SW, Ranganath R, Blei DM. 2018 Variational sequential Monte Carlo. In *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics*, pp. 968–977, vol. 84. PMLR.
47. Corenflos A, Thornton J, Deligiannidis G, Doucet A. 2021 Differentiable particle filtering via entropy-regularized optimal transport. In *Proceedings of the 38th International Conference on Machine Learning*, pp. 2100–2111, vol. 139. PMLR.
48. Lai J, Domke J, Sheldon D. 2022 Variational marginal particle filters. In *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics*, pp. 875–895, vol. 151. PMLR.
49. Matthews AG, Arbel M, Rezende DJ, Doucet A. 2022 Continual repeated annealed flow transport Monte Carlo. In *Proceedings of the 39th International Conference on Machine Learning*, pp. 15196–15219, vol. 162. PMLR.
50. Zimmermann H, Wu H, Esmaili B, van de Meent J-W. 2021 Nested variational inference. In *Advances in Neural Information Processing Systems* 35, pp. 20423–20435. Curran Associates, Inc.
51. Arbel M, Matthews A, Doucet A. 2021 Annealed flow transport Monte Carlo. In *Proceedings of the 38th International Conference on Machine Learning*, pp. 318–330, vol. 139. PMLR.
52. Midgley LI, Stimper V, Simm GN, Schölkopf B, Hernández-Lobato JM. 2023 Flow annealed importance sampling bootstrap. In *Int. Conf. on Learning Representations*.
53. Rotskoff G, Vanden-Eijnden E. 2019 Dynamical computation of the density of states and Bayes factors using nonequilibrium importance sampling. *Phys. Rev. Lett.* **122**, 150602. (doi:10.1103/PhysRevLett.122.150602)
54. Thin A, Janati El Idrissi Y, Le Corff S, Ollion C, Moulines E, Doucet A, Durmus A, Robert CX. 2021 Neo: non equilibrium sampling on the orbits of a deterministic transform. In *Advances in Neural Information Processing Systems* 35, pp. 17060–17071. Curran Associates, Inc.
55. Caterini AL, Doucet A, Sejdinovic D. 2018 Hamiltonian variational auto-encode. In *Advances in Neural Information Processing Systems* 32, pp. 8178–8188. Curran Associates, Inc.
56. Cardoso G, Samsonov S, Thin A, Moulines E, Olsson J. 2022 Br-snis: bias reduced self-normalized importance sampling. In *Advances in Neural Information Processing Systems* 36. Curran Associates, Inc.
57. Ruiz FJ, Titsias MK, Cemgil T, Doucet A. 2021 Unbiased gradient estimation for variational auto-encoders using coupled Markov chains. In *Proceedings of the Thirty-Seventh Conference on Uncertainty in Artificial Intelligence*, pp. 707–717, vol. 161. PMLR.
58. Andrieu C, Doucet A, Holenstein R. 2010 Particle Markov chain Monte Carlo methods. *J. R. Stat. Soc. B* **72**, 269–342. (doi:10.1111/j.1467-9868.2009.00736.x)
59. Andrieu C, Lee A, Vihola M. 2018 Uniform ergodicity of the iterated conditional SMC and geometric ergodicity of particle Gibbs samplers. *Bernoulli* **24**, 842–872. (doi:10.3150/15-BEJ785)