

---

# Optimal Variance Control of the Score Function Gradient Estimator for Importance Weighted Bounds

---

Valentin Liévin<sup>1</sup>    Andrea Dittadi<sup>1</sup>    Anders Christensen<sup>1</sup>    Ole Winther<sup>1,2,3</sup>

<sup>1</sup> Section for Cognitive Systems, Technical University of Denmark

<sup>2</sup> Bioinformatics Centre, Department of Biology, University of Copenhagen

<sup>3</sup> Centre for Genomic Medicine, Rigshospitalet, Copenhagen University Hospital  
{valv,adit}@dtu.dk, anders.christensen321@gmail.com, olwi@dtu.dk

## Abstract

This paper introduces novel results for the score function gradient estimator of the importance weighted variational bound (IWAE). We prove that in the limit of large  $K$  (number of importance samples) one can choose the control variate such that the Signal-to-Noise ratio (SNR) of the estimator grows as  $\sqrt{K}$ . This is in contrast to the standard pathwise gradient estimator where the SNR decreases as  $1/\sqrt{K}$ . Based on our theoretical findings we develop a novel control variate that extends on VIMCO. Empirically, for the training of both continuous and discrete generative models, the proposed method yields superior variance reduction, resulting in an SNR for IWAE that increases with  $K$  without relying on the reparameterization trick. The novel estimator is competitive with state-of-the-art reparameterization-free gradient estimators such as Reweighted Wake-Sleep (RWS) and the thermodynamic variational objective (TVO) when training generative models.

## 1 Introduction

Gradient-based learning is now widespread in the field of machine learning, in which recent advances have mostly relied on the backpropagation algorithm, the workhorse of modern deep learning. In many instances, for example in the context of unsupervised learning, it is desirable to make models more expressive by introducing stochastic latent variables. Backpropagation thus has to be augmented with methodologies for marginalization over latent variables.

Variational inference using an inference model (amortized inference) has emerged as a key method for training and inference in latent variable models [1-7]. The pathwise gradient estimator, based on the reparameterization trick [2, 3], often gives low-variance estimates of the gradient for continuous distributions. However, since discrete distributions cannot be reparameterized, these methods are not applicable to inference in complex simulators with discrete variables, such as reinforcement learning or advanced generative processes [8-11]. While the score function (or Reinforce) estimator [12] is more generally applicable, it is well known to suffer from large variance. Consequently, most of the recent developments focus on reducing the variance using control variates [13-18] and using alternative variational objectives [9, 19-21].

Recently, variational objectives tighter than the traditional evidence lower bound (ELBO) have been proposed [21, 22]. In importance weighted autoencoders (IWAE) [22] the tighter bound comes with the price of a  $K$ -fold increase in the required number of samples from the inference network. Despite yielding a tighter bound, using more samples can be detrimental to the learning of the inference model [23]. In fact, the Signal-to-Noise ratio (the ratio of the expected gradient to its standard deviation) of the pathwise estimator has been shown to decrease at a rate  $\mathcal{O}(K^{-1/2})$  [23]. Although this can be improved to  $\mathcal{O}(K^{1/2})$  by exploiting properties of the gradient to cancel high-variance

terms [24], the variational distributions are still required to be reparameterizable. In this work we introduce OVIS (*Optimal Variance – Importance Sampling*), a novel score function-based estimator for importance weighted objectives with improved SNR.

The main contributions of this paper are: 1) A proof that, with an appropriate choice of control variate, the score function estimator for the IWAE objective can achieve a Signal-to-Noise Ratio  $\text{SNR} = \mathcal{O}(K^{1/2})$  as the number of importance samples  $K \rightarrow \infty$ . 2) A derivation of OVIS, a class of practical low-variance score function estimators following the principles of our theoretical analysis. 3) State-of-the-art results on a number of non-trivial benchmarks for both discrete and continuous stochastic variables, with comparison to a range of recently proposed score function methods.

## 2 Optimizing the Importance Weighted Bound

**Importance weighted bound (IWAE)** Amortized variational inference allows fitting a latent variable model  $p_\theta(\mathbf{x}, \mathbf{z})$  to the data using an approximate posterior  $q_\phi(\mathbf{z}|\mathbf{x})$  [2]. By using multiple importance weighted samples, we can derive a lower bound to the log marginal likelihood that is uniformly tighter as the number of samples,  $K$ , increases [22]. The *importance weighted bound* (IWAE) for one data point  $\mathbf{x}$  is:

$$\mathcal{L}_K(\mathbf{x}) := \mathbb{E} \left[ \log \hat{Z} \right] \quad \hat{Z} := \frac{1}{K} \sum_{k=1}^K w_k \quad w_k := \frac{p_\theta(\mathbf{x}, \mathbf{z}_k)}{q_\phi(\mathbf{z}_k|\mathbf{x})}, \quad (1)$$

where  $\mathbb{E}$  denotes an expectation over the  $K$ -copy variational posterior  $q_\phi(\mathbf{z}_{1:K}|\mathbf{x}) := \prod_{k=1}^K q_\phi(\mathbf{z}_k|\mathbf{x})$ . This bound coincides with the traditional evidence lower bound (ELBO) for  $K = 1$ . The log likelihood lower bound for the entire data set is  $\mathcal{L}_K(\mathbf{x}_{1:n}) = \sum_{i=1}^n \mathcal{L}_K(\mathbf{x}_i)$ . In the following we will derive results for one term  $\mathcal{L}_K = \mathcal{L}_K(\mathbf{x})$ .

**Score function estimator** Without making assumptions about the variational distribution, the gradient of the importance weighted bound (1) with respect to the parameters of the approximate posterior factorizes as (see Appendix A):

$$\nabla_\phi \mathcal{L}_K = \mathbb{E} \left[ \sum_k d_k \mathbf{h}_k \right] \quad d_k := \log \hat{Z} - v_k \quad v_k := \frac{w_k}{\sum_{l=1}^K w_l}, \quad (2)$$

where  $\mathbf{h}_k := \nabla_\phi \log q_\phi(\mathbf{z}_k|\mathbf{x})$  is the score function. A Monte Carlo estimate of the expectation in (2) yields the *score function* (or *Reinforce*) estimator.

**Control variates** The vanilla score function estimator of (2) is often not useful in practice due to its large sample-to-sample variance. By introducing control variates that aim to cancel out zero expectation terms, this variance can be reduced while keeping the estimator unbiased.

Given posterior samples  $\mathbf{z}_1, \dots, \mathbf{z}_K \sim q_\phi(\mathbf{z}_{1:K}|\mathbf{x})$ , let  $\mathbf{z}_{-k}$  denote  $[\mathbf{z}_1, \dots, \mathbf{z}_{k-1}, \mathbf{z}_{k+1}, \dots, \mathbf{z}_K]$ , let  $\mathbb{E}_k[\dots]$  and  $\mathbb{E}_{-k}[\dots]$  be the expectations over the variational distributions of  $\mathbf{z}_k$  and  $\mathbf{z}_{-k}$ , respectively, and let  $\{c_k\}_{k=1}^K$  be scalar control variates, with each  $c_k = c_k(\mathbf{z}_{-k})$  independent of  $\mathbf{z}_k$ . Using the independence of  $c_k$  and  $\mathbf{h}_k$  for each  $k$ , and the fact that the score function has zero expectation, we have  $\mathbb{E}[c_k \mathbf{h}_k] = \mathbb{E}_{-k}[c_k] \mathbb{E}_k[\mathbf{h}_k] = 0$ . Thus, we can define an unbiased estimator of (2) as:

$$\mathbf{g} := \sum_k (d_k - c_k) \mathbf{h}_k \quad (3)$$

$$\mathbb{E}[\mathbf{g}] = \mathbb{E} \left[ \sum_k (d_k - c_k) \mathbf{h}_k \right] = \mathbb{E} \left[ \sum_k d_k \mathbf{h}_k \right] = \nabla_\phi \mathcal{L}_K. \quad (4)$$

In the remainder of this paper, we will use the decomposition  $d_k = f_k + f_{-k}$ , where  $f_k = f_k(\mathbf{z}_k, \mathbf{z}_{-k})$  and  $f_{-k} = f_{-k}(\mathbf{z}_{-k})$  denote terms that depend and do not depend on  $\mathbf{z}_k$ , respectively. This will allow us to exploit the mutual independence of  $\{\mathbf{z}_k\}_{k=1}^K$  to derive optimal control variates.

**Signal-to-Noise Ratio (SNR)** We will compare the different estimators on the basis of their Signal-to-noise ratio. Following [23], we define the SNR for each component of the gradient vector as

$$\text{SNR}_i := \frac{|\mathbb{E}[g_i]|}{\sqrt{\text{Var}[g_i]}}, \quad (5)$$

where  $g_i$  denotes the  $i$ th component of the gradient vector.

In Section 3 we derive the theoretical SNR for the optimal choice of control variates in the limit  $K \rightarrow \infty$ . In Section 4 we derive the optimal scalar control variates  $\{c_k\}_{k=1}^K$  by optimizing the trace of the covariance of the gradient estimator  $\mathbf{g}$ , and in Section 6 we experimentally compare our approach with state-of-the-art gradient estimators in terms of SNR.

### 3 Asymptotic Analysis of the Signal-to-Noise Ratio

Assuming the importance weights have finite variance, i.e.  $\text{Var}[w_k] < \infty$ , we can derive the asymptotic behavior of the SNR as  $K \rightarrow \infty$  by expanding  $\log \hat{Z}$  as a Taylor series around  $Z := p_\theta(\mathbf{x}) = \int p_\theta(\mathbf{x}, \mathbf{z}) d\mathbf{z}$  [23]. A direct application of the pathwise gradient estimator (reparameterization trick) to the importance weighted bound results in an SNR that scales as  $\mathcal{O}(K^{-1/2})$  [23], which can be improved to  $\mathcal{O}(K^{1/2})$  by exploiting properties of the gradient [24]. In the following we will show that, for a specific choice of control variate, the SNR of the score function estimator scales as  $\mathcal{O}(K^{1/2})$ . Thus, a score function estimator exists for which *increasing the number of importance samples benefits the gradient estimate of the parameters of the variational distribution*.

For the asymptotic analysis we rewrite the estimator as  $\mathbf{g} = \sum_k \left( -\frac{\partial \log \hat{Z}}{\partial w_k} w_k + \log \hat{Z} - c_k \right) \mathbf{h}_k$  and apply a second-order Taylor expansion to  $\log \hat{Z}$ . The resulting expression  $\mathbf{g} = \sum_k (f_k + f_{-k} - c_k) \mathbf{h}_k$  separates terms  $f_k$  that contribute to the expected gradient from terms  $f_{-k}$  that have zero expectation and thus only contribute to the variance (cf. Appendix B):

$$f_k \approx \frac{w_k^2}{2K^2 Z^2} \quad (6)$$

$$f_{-k} \approx \log Z - \frac{3}{2} + \frac{2}{KZ} \sum_{l \neq k} w_l - \frac{1}{2K^2 Z^2} \left( \sum_{l \neq k} w_l \right)^2. \quad (7)$$

Since  $f_{-k}$  and  $c_k$  are independent of  $\mathbf{h}_k$ , the expected gradient is (cf. Appendix C.1):

$$\mathbb{E}[\mathbf{g}] = \sum_k \mathbb{E}[f_k \mathbf{h}_k] \approx \frac{1}{2Z^2 K} \mathbb{E}_1 [w_1^2 \mathbf{h}_1] = \mathcal{O}(K^{-1}), \quad (8)$$

where  $\mathbb{E}_1$  denotes an expectation over the first latent distribution  $q_\phi(\mathbf{z}_1 | \mathbf{x})$ . Since the choice of control variates  $c_k = c_k(\mathbf{z}_{-k})$  is free, we can choose  $c_k = f_{-k}$  to cancel out all zero expectation terms. The resulting covariance, derived in Appendix C.2 is:

$$\text{Cov}[\mathbf{g}] = \text{Cov} \left[ \sum_k f_k \mathbf{h}_k \right] \approx \frac{1}{4K^3 Z^4} \text{Cov}_1 [w_1^2 \mathbf{h}_1] = \mathcal{O}(K^{-3}) \quad (9)$$

with  $\text{Cov}_1$  indicating the covariance over  $q_\phi(\mathbf{z}_1 | \mathbf{x})$ . Although as we discuss in Section 4 this is not the minimal variance choice of control variates, it is sufficient to achieve an SNR of  $\mathcal{O}(K^{1/2})$ .

### 4 Optimal Control Variate

The analysis above shows that in theory it is possible to attain a good SNR with the score function estimator. In this section we derive the optimal (in terms of variance of the resulting estimator) control variates  $\{c_k\}_{k=1}^K$  by decomposing  $\mathbf{g} = \sum_k (f_k + f_{-k} - c_k) \mathbf{h}_k$  as above, and minimizing the trace of the covariance matrix, i.e.  $\mathbb{E}[\|\mathbf{g}\|^2] - \|\mathbb{E}[\mathbf{g}]\|^2$ . Since  $\mathbb{E}[f_{-k} \mathbf{h}_k]$  and  $\mathbb{E}[c_k \mathbf{h}_k]$  are both zero,  $\mathbb{E}[\mathbf{g}] = \nabla_\phi \mathcal{L}_K$  does not depend on  $c_k$ . Thus, the minimization only involves the first term:

$$\begin{aligned} \frac{1}{2} \frac{\partial}{\partial c_k} \mathbb{E} [\|\mathbf{g}\|^2] &= \mathbb{E} \left[ \mathbf{h}_k^T \sum_l (f_l + f_{-l} - c_l) \mathbf{h}_l \right] \\ &= \mathbb{E}_{-k} \left[ \sum_l \mathbb{E}_k [f_l \mathbf{h}_k^T \mathbf{h}_l] + (f_{-k} - c_k) \mathbb{E}_k [\|\mathbf{h}_k\|^2] \right]. \end{aligned}$$

where  $\mathbb{E}_k$  and  $\mathbb{E}_{-k}$  indicate expectations over  $q_\phi(\mathbf{z}_k | \mathbf{x})$  and  $q_\phi(\mathbf{z}_{-k} | \mathbf{x})$ , respectively. Setting the argument of  $\mathbb{E}_{-k}$  to zero, we get the optimal control variates  $c_k = c_k(\mathbf{z}_{-k})$  and gradient estimator  $\mathbf{g}$ :

$$c_k = f_{-k} + \sum_l \frac{\mathbb{E}_k [f_l \mathbf{h}_k^T \mathbf{h}_l]}{\mathbb{E}_k [\|\mathbf{h}_k\|^2]} \quad (10)$$

$$\mathbf{g} = \sum_k \left( f_k - \sum_l \frac{\mathbb{E}_k [f_l \mathbf{h}_k^T \mathbf{h}_l]}{\mathbb{E}_k [\|\mathbf{h}_k\|^2]} \right) \mathbf{h}_k. \quad (11)$$

Applying (11) in practice requires marginalizing over one latent variable and decoupling terms that do not depend on  $\mathbf{z}_k$  from those that do. In the remainder of this section we will 1) make a series of approximations to keep computation tractable, and 2) consider two limiting cases for the *effective sample size* (ESS) [25] in which we can decouple terms.

**Simplifying approximations to Equation (11)** First, we consider a term with  $l \neq k$ , define  $\Delta f_l := f_l - \mathbb{E}_k[f_l]$ , and subtract and add  $\mathbb{E}_k[f_l]$  from inside the expectation:

$$\mathbb{E}_k [f_l \mathbf{h}_k^T] \mathbf{h}_l = \mathbb{E}_k [\Delta f_l \mathbf{h}_k^T] \mathbf{h}_l + \mathbb{E}_k[f_l] \mathbb{E}_k [\mathbf{h}_k^T] \mathbf{h}_l = \mathbb{E}_k [\Delta f_l \mathbf{h}_k^T] \mathbf{h}_l$$

where we used the fact that  $\mathbb{E}_k[\mathbf{h}_k] = 0$ . The  $l \neq k$  terms thus only contribute to fluctuations relative to a mean value, and we assume they can be neglected.

Second, we assume that  $|\phi|$ , the number of parameters of  $q_\phi$ , is large, and the terms of the sum  $\|\mathbf{h}_k\|^2 = \sum_{i=1}^{|\phi|} h_{ki}^2$  are approximately independent with finite variances  $\sigma_i^2$ . By the Central Limit Theorem we approximate the distribution of  $\Delta \|\mathbf{h}_k\|^2 := \|\mathbf{h}_k\|^2 - \mathbb{E}_k[\|\mathbf{h}_k\|^2]$  with a zero-mean Gaussian with standard deviation  $(\sum_{i=1}^{|\phi|} \sigma_i^2)^{1/2}$ . Seeing that  $\mathbb{E}_k[\|\mathbf{h}_k\|^2]$  is  $\mathcal{O}(|\phi|)$ , we have

$$\frac{\mathbb{E}_k [f_k \|\mathbf{h}_k\|^2]}{\mathbb{E}_k [\|\mathbf{h}_k\|^2]} = \mathbb{E}_k [f_k] + \frac{\mathbb{E}_k [f_k \Delta \|\mathbf{h}_k\|^2]}{\mathbb{E}_k [\|\mathbf{h}_k\|^2]} = \mathbb{E}_k [f_k] + \mathcal{O}(|\phi|^{-1/2}),$$

where we used that the argument in the numerator scales as  $(\sum_{i=1}^{|\phi|} \sigma_i^2)^{1/2} = \mathcal{O}(|\phi|^{1/2})$ .

Finally, the expectation can be approximated with a sample average. Writing  $f_k = f_k(\mathbf{z}_k, \mathbf{z}_{-k})$  and drawing  $S$  new samples  $\mathbf{z}^{(1)}, \dots, \mathbf{z}^{(S)} \sim q_\phi(\mathbf{z}|\mathbf{x})$ :

$$\mathbb{E}_k [f_k] \approx \frac{1}{S} \sum_{s=1}^S f_k(\mathbf{z}^{(s)}, \mathbf{z}_{-k}).$$

This will introduce additional fluctuations with scale  $S^{-1/2}$ .

Putting these three approximations together and using  $d_k(\mathbf{z}_k, \mathbf{z}_{-k}) = f_k(\mathbf{z}_k, \mathbf{z}_{-k}) + f_{-k}(\mathbf{z}_{-k})$ , we obtain the sample-based expression of the OVIS estimator, called OVIS<sub>MC</sub> in the following:

$$\text{OVIS}_{\text{MC}} : \quad \mathbf{g} \approx \sum_k \left( d_k(\mathbf{z}_k, \mathbf{z}_{-k}) - \frac{1}{S} \sum_{s=1}^S d_k(\mathbf{z}^{(s)}, \mathbf{z}_{-k}) \right) \mathbf{h}_k. \quad (12)$$

Naively, this will produce a large computational overhead because we now have in total  $KS$  terms. However, we can reduce this to  $\mathcal{O}(K+S)$  because the bulk of the computation comes from evaluating the importance weights and because the  $S$  auxiliary samples can be reused for all  $K$  terms.

**Effective sample size (ESS)** The ESS [25] is a commonly used yardstick of the efficiency of an importance sampling estimate, defined as

$$\text{ESS} := \frac{(\sum_k w_k)^2}{\sum_k w_k^2} = \frac{1}{\sum_k v_k^2} \in [1, K]. \quad (13)$$

A low ESS occurs when only a few weights dominate, which indicates that the proposal distribution  $q$  poorly matches  $p$ . In the opposite limit, the variance of importance weights is finite and the ESS will scale with  $K$ . Therefore the limit  $\text{ESS} \gg 1$  corresponds to the asymptotic limit studied in Section 3.

**Optimal control for ESS limits and unified interpolation** In the following, we consider the two extreme limits  $\text{ESS} \gg 1$  and  $\text{ESS} \approx 1$  to derive sample-free approximations to the optimal control. We can thus in these limits avoid the sample fluctuations and excess computation of OVIS<sub>MC</sub>.

We first consider  $\text{ESS} \gg 1$  and for each  $k$  we introduce the unnormalized leave- $w_k$ -out approximation to  $\hat{Z}$ :

$$\tilde{Z}_{[-k]} := \frac{1}{K} \sum_{l \neq k} w_l \quad \text{such that} \quad \hat{Z} - \tilde{Z}_{[-k]} = \frac{w_k}{K}. \quad (14)$$

Assuming  $\text{Var}[w_k] < \infty$ , this difference is  $\mathcal{O}(K^{-1})$  as  $K \rightarrow \infty$ , thus we can expand  $\log \hat{Z}$  around  $\hat{Z} = \tilde{Z}_{[-k]}$ . In this limit, the optimal control variate simplifies to (cf. Appendix [D.1](#)):

$$\text{ESS} \gg 1 : \quad c_k \approx \log \frac{1}{K-1} \sum_{l \neq k} w_l + \log \left(1 - \frac{1}{K}\right). \quad (15)$$

When  $\text{ESS} \approx 1$ , one weight is much larger than the others and the assumption above is no longer valid. To analyze this frequently occurring scenario, assume that  $k' = \text{argmax}_l w_l$  and  $w_{k'} \gg \sum_{l \neq k'} w_l$ . In this limit  $\log \hat{Z} \approx \log w_{k'}/K$  and  $v_k \approx \delta_{k,k'}$  and thus  $d_k = \log w_{k'}/K - \delta_{k,k'}$ . In Appendix [D.2](#) we show we can approximate Equation [\(10\)](#) with

$$\text{ESS} \approx 1 : \quad c_k \approx \log \frac{1}{K-1} \sum_{l \neq k} w_l - v_k. \quad (16)$$

We introduce  $\text{OVIS}_\gamma$  to interpolate between the two limits (Appendix [D.3](#)):

$$c_k^\gamma := \log \frac{1}{K-1} \sum_{l \neq k} w_l - \gamma v_k + (1-\gamma) \log \left(1 - \frac{1}{K}\right) \quad \gamma \in [0, 1]. \quad (17)$$

In this paper we will only conduct experiments for the two limiting cases  $\gamma = 0$ , corresponding to Equation [\(15\)](#), and  $\gamma = 1$  approximating Equation [\(16\)](#). Tuning the parameter  $\gamma$  in the range  $[0, 1]$  will be left for future work. We discuss the implementation in the appendix [K](#).

**Higher ESS with looser lower bound** Empirically we observe that training may be impaired by a low ESS and by *posterior collapse* [\[4, 26-29\]](#). This motivates trading the tight IWAE objective for a gradient estimator with higher ESS. To that end, we use the importance weighted Rényi (IWR) bound:

$$\mathcal{L}_K^\alpha(\mathbf{x}) := \frac{1}{1-\alpha} \mathbb{E} \left[ \log \hat{Z}(\alpha) \right] \quad \hat{Z}(\alpha) := \frac{1}{K} \sum_k w_k^{1-\alpha} \quad (18)$$

which for  $\alpha \in [0, 1]$  is a lower bound on the Rényi objective  $\log \mathbb{E}_1 [w_1^{1-\alpha}] / (1-\alpha)$  [\[30\]](#). The Rényi objective in itself coincides with  $\log p(\mathbf{x})$  for  $\alpha = 0$  and is monotonically non-increasing in  $\alpha$ , i.e. is an evidence lower bound [\[30\]](#). So we have a looser bound but higher  $\text{ESS}(\alpha) = 1 / \sum_k v_k^2(\alpha) \geq \text{ESS}(0)$  for  $\alpha \in [0, 1]$  with  $v_k(\alpha) = w_k^{1-\alpha} / \sum_l w_l^{1-\alpha}$ . Furthermore, for  $\alpha = 1$  the bound corresponds to the ELBO and the divergence  $\mathcal{D}_{\text{KL}}(q_\phi(\mathbf{z}|\mathbf{x}) || p_\theta(\mathbf{z}|\mathbf{x}))$  is guaranteed to be minimized. In Appendix [E](#) we derive the score function estimator and control variate expressions for  $\mathcal{L}_K^\alpha$ . The objective can either be used in a warm-up scheme by gradually decreasing  $\alpha \rightarrow 0$  throughout iterations or can be run with a constant  $0 < \alpha < 1$ .

## 5 Related Work

The score function estimator with control variates can be used with all the commonly used variational families. By contrast, the reparameterization trick is only applicable under specific conditions. We now give a brief overview of the existing alternatives and refer the reader to [\[31\]](#) for a more extensive review. The importance of handling discrete distributions without relaxations is discussed in [\[9\]](#).

NVIL [\[13\]](#), DARN [\[17\]](#), and MuProp [\[18\]](#) demonstrate that score function estimators with carefully crafted control variates allow to train deep generative models. VIMCO [\[14\]](#) extends this to multi-sample objectives, and recycles the Monte Carlo samples  $\mathbf{z}_{-k}$  to define a control variate  $c_k = c_k(\mathbf{z}_{-k})$ . Unlike OVIS, VIMCO only controls the variance of the term  $\log \hat{Z}$  in  $d_k = \log \hat{Z} - v_k$ , leaving  $v_k$  uncontrolled, and causing the SNR to decrease with the number of particles  $K$  as we empirically observe in Section [6.1](#). We provide a detailed review of VIMCO in Appendix [F](#).

The Reweighted Wake-Sleep (RWS) algorithm [\[20\]](#) is an extension of the original Wake-Sleep algorithm (ws) [\[19\]](#) that alternates between two distinct learning phases for optimizing importance weighted objectives. A detailed review of RWS and ws is available in Appendix [F](#).

The Thermodynamic Variational Objective (TVO) [\[21\]](#) is a lower bound to  $\log p_\theta(\mathbf{x})$  that stems from a Riemannian approximation of the Thermodynamic Variational Identity (TVI), and unifies the objectives of Variational Inference and Wake-Sleep. Evaluating the gradient involves differentiating

through an expectation over a distribution with an intractable normalizing constant. To accommodate this, the authors propose an estimator that generalizes the score function estimator based on a tractable covariance term. We review the TVO in more detail in Appendix F.

Given a deterministic *sampling path*  $g(\epsilon; \theta)$  such that  $\mathbf{z} \sim p_\theta(\mathbf{z})$  and  $\mathbf{z} = g(\epsilon; \theta), \epsilon \sim p(\epsilon)$  are equivalent, one can derive a *pathwise gradient estimator* of the form  $\nabla_\theta \mathbb{E}_{p_\theta(\mathbf{z})} [f_\theta(\mathbf{z})] = \mathbb{E}_{p(\epsilon)} [\nabla_\theta f_\theta(g(\epsilon; \theta))]$ . This estimator – introduced in machine learning as the *reparameterization trick* or *stochastic backpropagation* [2, 3] – exhibits low variance thanks to the structural information provided by the sampling path. Notably, a zero expectation term can be removed from the estimator [32]. Extending on this, [24] derives an alternative gradient estimator for IWAE that exhibits  $\text{SNR} \sim K^{1/2}$ , as opposed to  $\text{SNR} \sim K^{-1/2}$  for the *standard* IWAE objective [23].

Continuous relaxations of discrete distributions yield a biased low-variance gradient estimate thanks to the reparameterization trick [16, 33]. Discrete samples can be obtained using the Straight-Through estimator [5, 34]. The resulting gradient estimate remains biased, but can be used as a control variate for the score function objective, resulting in an unbiased low-variance estimate of the gradient [15, 35].

## 6 Experimental Results

We conduct a number of experiments<sup>1</sup> on benchmarks that have previously been used to test score function based estimators. All models are trained via stochastic gradient ascent using the Adam optimizer [36] with default parameters. We use regular gradients on the training objective for the generative model parameters  $\theta$ . The SNR for  $\theta$  scales as  $\mathcal{O}(K^{1/2})$  [23].

### 6.1 Asymptotic Variance

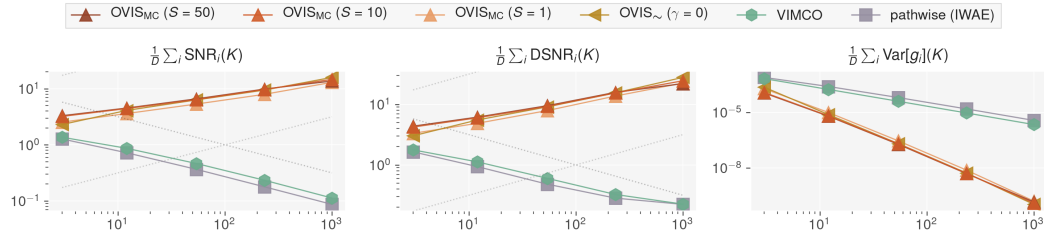


Figure 1: Gaussian model. Parameter-wise average of the asymptotic SNR, DSNR and variance of the gradients of the parameter  $b$  for different number of particles  $K \in [3, 1000]$  using  $10^4$  MC samples. The dotted lines stand for  $y = 10^{\pm 1} K^{\pm 0.5}$ .

Following [23], we empirically corroborate the asymptotic properties of the OVIS gradient estimator by means of the following simple model:

$$\mathbf{z} \sim \mathcal{N}(\mathbf{z}; \boldsymbol{\mu}, \mathbf{I}), \quad \mathbf{x}|\mathbf{z} \sim \mathcal{N}(\mathbf{x}; \mathbf{z}, \mathbf{I}), \quad q_\phi(\mathbf{z}|\mathbf{x}) = \mathcal{N}(\mathbf{z}; \mathbf{A}\mathbf{x} + \mathbf{b}, \frac{2}{3}\mathbf{I}).$$

where  $\mathbf{x}$  and  $\mathbf{z}$  are real vectors of size  $D = 20$ . We sample  $N = 1024$  points  $\{\mathbf{x}^{(n)}\}_{n=1}^N$  from the *true* model where  $\boldsymbol{\mu}^* \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ . The optimal parameters are  $\mathbf{A}^* = \mathbf{I}/2$ ,  $\mathbf{b}^* = \boldsymbol{\mu}^*/2$ , and  $\boldsymbol{\mu}^* = \frac{1}{N} \sum_{n=1}^N \mathbf{x}^{(n)}$ . The model parameters are obtained by adding Gaussian noise of scale  $\epsilon = 10^{-3}$ . We measure the variance and the SNR of the gradients with  $10^4$  MC samples. We also measured the *directional* SNR (DSNR [23]) to probe if our results hold in the multidimensional case.

In Figure 1 we report the gradient statistics for  $\mathbf{b}$ . We observe that using more samples in the standard IWAE leads to a decrease in SNR as  $\mathcal{O}(K^{-1/2})$  for both VIMCO and the pathwise-IWAE [23]. The tighter variance control provided by OVIS leads the variance to decrease almost at a rate  $\mathcal{O}(K^{-3})$ , resulting in a measured SNR not far from  $\mathcal{O}(K^{1/2})$  both for  $\text{OVIS}_{\text{MC}}$  and  $\text{OVIS}_{\sim}$ . This shows that, despite the approximations, the proposed gradient estimators  $\text{OVIS}_{\text{MC}}$  and  $\text{OVIS}_{\sim}$  are capable of achieving the theoretical SNR of  $\mathcal{O}(K^{1/2})$  derived in the asymptotic analysis in Section 3.

<sup>1</sup>The full experimental framework is available at [github.com/vlievin/ovis](https://github.com/vlievin/ovis)



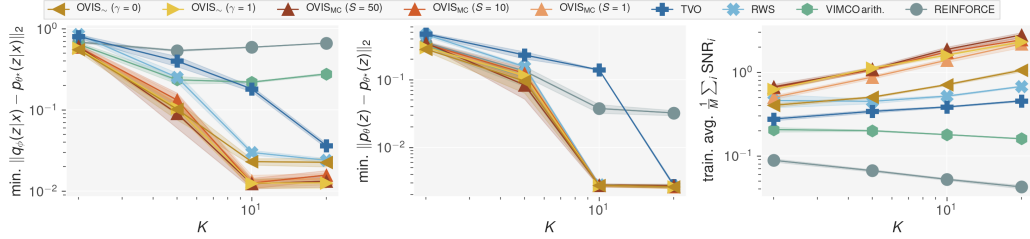


Figure 2: Training of the Gaussian mixture model. Minimum test-diagnostics recorded during training and training average of the SNR of the gradients of  $\phi$  with  $M = \text{card}(\phi)$ . In contrast to VIMCO,  $\text{OVIS}_{\sim}$  and  $\text{OVIS}_{\text{MC}}$  all benefit from the increase of the particles budget,  $\text{OVIS}_{\text{MC}}$  yields the most accurate posterior among the compared methods.

In Appendix C we learn the parameters of the Gaussian model using OVIS, RWS, VIMCO and the TVO. We find that optimal variance reduction translates into a more accurate estimation of the optimal parameters of the inference network when compared to RWS, VIMCO and the TVO.

## 6.2 Gaussian Mixture Model

We evaluate OVIS on a Gaussian Mixture Model and show that, unlike VIMCO [9], our method yields better inference networks as the number of particles  $K$  increases. Following [9], we define:

$$p_{\theta}(z) = \text{Cat}(z | \text{softmax}(\theta)) \quad p(x|z) = \mathcal{N}(x | \mu_z, \sigma_z^2) \quad q_{\phi}(z|x) = \text{Cat}(z | \text{softmax}(\eta_{\phi}(x)))$$

where  $z \in \{0, \dots, C-1\}$ ,  $\mu_z = 10z$ ,  $\sigma_z = 5$ , and  $C = 20$  is the number of clusters. The inference network  $\eta_{\phi}$  is parameterized by a multilayer perceptron with architecture 1-16- $C$  and tanh activations. The true generative model is set to  $p_{\theta^*}(z=c) = (c+5) / \sum_{i=1}^C (i+5)$ .

All models are trained for 100k steps with 5 random seeds. We compare OVIS with VIMCO, RWS with wake- $\phi$  update, Reinforce, and the TVO. For the latter we chose to use 5 partitions and  $\beta_1 = 10^{-2}$ , after a hyperparameter search over  $\beta_1 \in \{10^{-1}, 10^{-1.5}, 10^{-2}, 10^{-2.5}, 10^{-3}\}$  and  $\{2, 5\}$  partitions.

Each model is evaluated on a held-out test set of size  $M = 100$ . We measure the accuracy of the learned posterior  $q_{\phi}(z|x)$  by its average  $L_2$  distance from the true posterior, i.e.  $\frac{1}{M} \sum_{m=1}^M \|q_{\phi}(z|x^{(m)}) - p_{\theta^*}(z|x^{(m)})\|_2$ . As a sanity check, we assess the quality of the generative model using  $\|\text{softmax}(\theta) - \text{softmax}(\theta^*)\|_2$ . The SNR of the gradients for the parameters  $\phi$  is evaluated on one mini-batch of data using 500 MC samples.

We report our main results in Figure 2, and training curves in Appendix H. In contrast to VIMCO, the accuracy of the posteriors learned using  $\text{OVIS}_{\text{MC}}$  and  $\text{OVIS}_{\sim}$  all improve monotonically with  $K$  and outperform the baseline estimators, independently of the choice of the number of auxiliary particles  $S$ . All OVIS methods outperform the state-of-the-art estimators RWS and the TVO, as measured by the  $L_2$  distance between the approximate and the true posterior.

## 6.3 Deep Generative Models

We utilize the OVIS estimators to learn the parameters of both discrete and continuous deep generative models using stochastic gradient ascent. The base learning rate is fixed to  $3 \cdot 10^{-4}$ , we use mini-batches of size 24 and train all models for  $4 \cdot 10^6$  steps. We use the statically binarized MNIST dataset [37] with the original training/validation/test splits of size 50k/10k/10k. We follow the experimental protocol as detailed in [21], including the  $\beta$  partition for the TVO and the exact architecture of the models. We use a three-layer Sigmoid Belief Network [38] as an archetype of discrete generative model [13, 14, 21] and a Gaussian Variational Autoencoder [2] with 200 latent variables. All models are trained with three initial random seeds and for  $K \in \{5, 10, 50\}$  particles.

We assess the performance based on the marginal log-likelihood estimate  $\log \hat{p}_{\theta}(\mathbf{x}) = \mathcal{L}_{5000}(\mathbf{x})$ , that we evaluate on 10k training data points, such as to disentangle the training dynamics from the regularisation effect that is specific to each method. We measure the quality of the inference network solution using the divergence  $\mathcal{D}_{\text{KL}}(q_{\phi}(z|\mathbf{x}) || p_{\theta}(z|\mathbf{x})) \approx \log \hat{p}_{\theta}(\mathbf{x}) - \mathcal{L}_1(\mathbf{x})$ . The full training curves

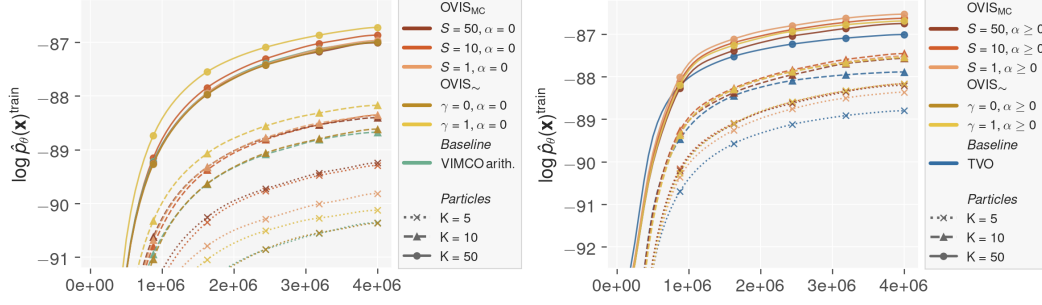


Figure 3: Training a Sigmoid Belief Network on Binarized MNIST. (Left) Optimizing for the importance weighted bound  $\mathcal{L}_K$  using OVIS. (Right) Optimizing for the Rényi importance lower bound  $\mathcal{L}_K^\alpha$  using OVIS with  $\alpha$  annealing  $0.99 \rightarrow 0$ . The curves are averaged over three seeds and smoothed for clarity.

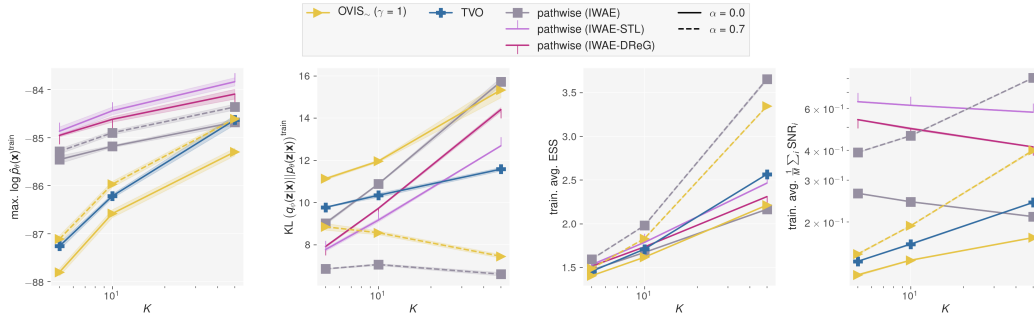


Figure 4: Training a one layer Gaussian VAE. Maximum recorded training  $\log \hat{p}_\theta(\mathbf{x})$ , final estimate of the bound  $\mathcal{D}_{\text{KL}}(q_\phi(\mathbf{z}|\mathbf{x})||p_\theta(\mathbf{z}|\mathbf{x}))$  and training average of the ESS and of the SNR. OVIS yields similar likelihood performances as the TVO but benefits from a tighter bound thanks to optimizing for the IWR bound.

– including the test log likelihood and divergences – are available in Appendix [J](#). We will show that OVIS improves over VIMCO, on which it extends, and we show that combining OVIS $_\sim$  with the Variational Rényi bound (IWR) as described in Section [4](#) outperforms the TVO.

### 6.3.1 Sigmoid Belief Network (SBN)

**A. Comparison with VIMCO** We learn the parameters of the SBN using the OVIS estimators for the IWAE bound and use VIMCO as a baseline. We report  $\log \hat{p}_\theta(\mathbf{x})$  in the left plot of Figure [3](#). All OVIS methods outperform VIMCO, ergo supporting the advantage of optimal variance reduction. When using a small number of particles  $K = 5$ , learning can be greatly improved by using an accurate MC estimate of the optimal control variate, as suggested by OVIS $_{\text{MC}}(S = 50)$  which allows gaining +1.0 nats over VIMCO. While OVIS( $\gamma = 0$ ), designed for large ESS barely improved over VIMCO, the biased OVIS $_\sim(\gamma = 1)$  for low ESS performed significantly better than other methods for  $K \geq 10$ , which coincides with the ESS measured in the range  $[1.0, 3.5]$  for all methods. We attribute the relative decrease of performances observed for OVIS $_{\text{MC}}$  for  $K = 50$  to *posterior collapse*.

**B. Training using IWR bounds** In Figure [3](#) (right) we train the SBN using OVIS and the TVO. OVIS is coupled with the objective  $\mathcal{L}_K^\alpha$  for which we anneal the parameter  $\alpha$  from 0.99 ( $\mathcal{L}_K^{0.99} \approx \mathcal{L}_1$ ) to 0 ( $\mathcal{L}_K^0 = \mathcal{L}_K$ ) during  $1e6$  steps using geometric interpolation. For all  $K$  values, OVIS outperform the TVO and OVIS $_\sim(\gamma = 1)$  performs comparably with OVIS $_{\text{MC}}$ .



### 6.3.2 Gaussian Variational Autoencoder (VAE)

In Figure 4 we train the Gaussian VAE using the standard pathwise IWAE, Sticking the Landing (STL) [32], DReG [24], the TVO and  $\text{OVIS}_{\sim}(\gamma = 1)$ .

OVIS is applied to the IWR bound with  $\alpha = 0.7$ . As measured by the training likelihood,  $\text{OVIS}_{\sim}(\gamma = 1)$  coupled with the IWR bound performs on par with the TVO, which bridges the gap to the standard pathwise IWAE for  $K = 50$ , although different objectives are at play. The advanced pathwise estimators (STL and DReG) outperform all other methods. Measuring the quality of the learned proposals  $q_{\phi}(\mathbf{z}|\mathbf{x})$  using the KL divergence allows disentangling the TVO and  $\text{OVIS}_{\sim}$  methods, as  $\text{OVIS}(\gamma = 1)$  applied to the IWR bound outputs higher-quality approximate posteriors for all considered number of particles.

### 6.4 A final Note on $\text{OVIS}_{\sim}(\gamma = 1)$

$\text{OVIS}_{\sim}(\gamma = 1)$  generates training dynamics that are superior to the baseline TVO and to  $\text{OVIS}_{\text{MC}}$  given a comparable particle budget (appendix I). We interpret this result as a consequence of the ESS-specific design, which also appeared to be robust to the choice of  $\alpha$  in the IWR objective. This also corroborates the results of [32], that suppressing the term  $-\sum_k v_k \mathbf{h}_k$  from the gradient estimate improves learning. We therefore recommend the practitioner to first experiment with  $\text{OVIS}_{\sim}(\gamma = 1)$  since it delivers competitive results at a reasonable computational cost.

## 7 Conclusion

We proposed OVIS, a gradient estimator that is generally applicable to deep models with stochastic variables, and is empirically shown to have optimal variance control. This property is achieved by identifying and canceling terms in the estimator that solely contribute to the variance. We expect that in practice it will often be a good trade-off to use a looser bound with a higher effective sample size, e.g. by utilizing the OVIS estimator with the importance weighted Rényi bound, allowing control of this trade-off via an additional scalar smoothing parameter. This sentiment is supported by our method demonstrating better performance than the current state-of-the-art.

## 8 Financial Disclosure

The PhD program supporting Valentin Liévin is partially funded by Google. This research was supported by the NVIDIA Corporation with the donation of GPUs.

## 9 Broader Impact

This work proposes OVIS, an improvement to the score function gradient estimator in the form of optimal control variates for variance reduction. As briefly touched upon in the introduction, OVIS has potential practical use cases across several branches of machine learning. As such, the potential impact of this research is broad, and we will therefore limit the scope of this section to a few clear applications.

Improved inference over discrete spaces such as action spaces encountered within e.g. model-based reinforcement learning has the potential of reducing training time and result in more optimal behavior of the learning agent. This advancement has the capability to increase efficiency of e.g. autonomous robots used within manufacturing. Such progress is often coveted due to cost optimization, increased safety, and reduced manual labor for humans. However, as argued in [39], this development can also lead to immediate disadvantages such as worker displacement, potentially in terms of both tasks and geographic location.

Another probable avenue of impact of this research is within machine comprehension. A topic within this field is reading, with practical applications such as chatbots. This use of machine learning has seen rapid growth and commercial interest over recent years [40]. Apart from the clear consumer benefits of these bots, focus has also broadened to other cases of use for social benefits [41]. However, as with most other machine learning inventions, chatbots can be exploited for malicious purposes such as automated spread of misinformation, e.g. during elections [42].

As with other theoretical advances such as those presented in this paper, consequences are not immediate and depend on the applications in which the research is utilized. It is our hope that this research will ultimately be of practical use with a tangible positive impact.

## References

- [1] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.
- [2] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.
- [3] Danilo Jimenez Rezende, Shakir Mohamed, and Daan Wierstra. Stochastic backpropagation and approximate inference in deep generative models. *arXiv preprint arXiv:1401.4082*, 2014.
- [4] Samuel R Bowman, Luke Vilnis, Oriol Vinyals, Andrew M Dai, Rafal Jozefowicz, and Samy Bengio. Generating sentences from a continuous space. *arXiv preprint arXiv:1511.06349*, 2015.
- [5] Aaron van den Oord, Oriol Vinyals, et al. Neural discrete representation learning. In *Advances in Neural Information Processing Systems*, pages 6306–6315, 2017.
- [6] Irina Higgins, Loic Matthey, Arka Pal, Christopher Burgess, Xavier Glorot, Matthew Botvinick, Shakir Mohamed, and Alexander Lerchner. beta-vae: Learning basic visual concepts with a constrained variational framework. *Iclr*, 2(5):6, 2017.
- [7] Lars Maaløe, Marco Fraccaro, Valentin Liévin, and Ole Winther. BIVA: A very deep hierarchy of latent variables for generative modeling. In *Advances in neural information processing systems*, pages 6548–6558, 2019.
- [8] Richard S Sutton, David A McAllester, Satinder P Singh, and Yishay Mansour. Policy gradient methods for reinforcement learning with function approximation. In *Advances in neural information processing systems*, pages 1057–1063, 2000.
- [9] Tuan Anh Le, Adam R Kosiosek, N Siddharth, Yee Whye Teh, and Frank Wood. Revisiting reweighted wake-sleep. *arXiv preprint arXiv:1805.10469*, 2018.
- [10] SM Ali Eslami, Nicolas Heess, Theophane Weber, Yuval Tassa, David Szepesvari, Geoffrey E Hinton, et al. Attend, infer, repeat: Fast scene understanding with generative models. In *Advances in Neural Information Processing Systems*, pages 3225–3233, 2016.
- [11] Yishu Miao and Phil Blunsom. Language as a latent variable: Discrete generative models for sentence compression. *arXiv preprint arXiv:1609.07317*, 2016.
- [12] Ronald J Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning*, 8(3-4):229–256, 1992.
- [13] Andriy Mnih and Karol Gregor. Neural variational inference and learning in belief networks. *arXiv preprint arXiv:1402.0030*, 2014.
- [14] Andriy Mnih and Danilo J Rezende. Variational inference for monte carlo objectives. *arXiv preprint arXiv:1602.06725*, 2016.
- [15] George Tucker, Andriy Mnih, Chris J Maddison, John Lawson, and Jascha Sohl-Dickstein. Rebar: Low-variance, unbiased gradient estimates for discrete latent variable models. In *Advances in Neural Information Processing Systems*, pages 2627–2636, 2017.
- [16] Chris J Maddison, Andriy Mnih, and Yee Whye Teh. The concrete distribution: A continuous relaxation of discrete random variables. *arXiv preprint arXiv:1611.00712*, 2016.
- [17] Karol Gregor, Ivo Danihelka, Andriy Mnih, Charles Blundell, and Daan Wierstra. Deep autoregressive networks. *arXiv preprint arXiv:1310.8499*, 2013.

- [18] Shixiang Gu, Sergey Levine, Ilya Sutskever, and Andriy Mnih. Muprop: Unbiased backpropagation for stochastic neural networks. *arXiv preprint arXiv:1511.05176*, 2015.
- [19] Geoffrey E Hinton, Peter Dayan, Brendan J Frey, and Radford M Neal. The "wake-sleep" algorithm for unsupervised neural networks. *Science*, 268(5214):1158–1161, 1995.
- [20] Jörg Bornschein and Yoshua Bengio. Reweighted wake-sleep. *arXiv preprint arXiv:1406.2751*, 2014.
- [21] Vaden Masrani, Tuan Anh Le, and Frank Wood. The thermodynamic variational objective. In *Advances in Neural Information Processing Systems*, pages 11521–11530, 2019.
- [22] Yuri Burda, Roger Grosse, and Ruslan Salakhutdinov. Importance weighted autoencoders. *arXiv preprint arXiv:1509.00519*, 2015.
- [23] Tom Rainforth, Adam R Kosiorek, Tuan Anh Le, Chris J Maddison, Maximilian Igl, Frank Wood, and Yee Whye Teh. Tighter variational bounds are not necessarily better. *arXiv preprint arXiv:1802.04537*, 2018.
- [24] George Tucker, Dieterich Lawson, Shixiang Gu, and Chris J Maddison. Doubly reparameterized gradient estimators for monte carlo objectives. *arXiv preprint arXiv:1810.04152*, 2018.
- [25] Augustine Kong. A note on importance sampling using standardized weights. *University of Chicago, Dept. of Statistics, Tech. Rep*, 348, 1992.
- [26] Xi Chen, Diederik P Kingma, Tim Salimans, Yan Duan, Prafulla Dhariwal, John Schulman, Ilya Sutskever, and Pieter Abbeel. Variational lossy autoencoder. *arXiv preprint arXiv:1611.02731*, 2016.
- [27] Casper Kaae Sønderby, Tapani Raiko, Lars Maaløe, Søren Kaae Sønderby, and Ole Winther. Ladder variational autoencoders. In *Advances in neural information processing systems*, pages 3738–3746, 2016.
- [28] Durk P Kingma, Tim Salimans, Rafal Jozefowicz, Xi Chen, Ilya Sutskever, and Max Welling. Improved variational inference with inverse autoregressive flow. In *Advances in neural information processing systems*, pages 4743–4751, 2016.
- [29] Adji B Dieng, Yoon Kim, Alexander M Rush, and David M Blei. Avoiding latent variable collapse with generative skip models. *arXiv preprint arXiv:1807.04863*, 2018.
- [30] Yingzhen Li and Richard E Turner. Rényi divergence variational inference. In *Advances in Neural Information Processing Systems*, pages 1073–1081, 2016.
- [31] Shakir Mohamed, Mihaela Rosca, Michael Figurnov, and Andriy Mnih. Monte carlo gradient estimation in machine learning. *arXiv preprint arXiv:1906.10652*, 2019.
- [32] Geoffrey Roeder, Yuhuai Wu, and David K Duvenaud. Sticking the landing: Simple, lower-variance gradient estimators for variational inference. In *Advances in Neural Information Processing Systems*, pages 6925–6934, 2017.
- [33] Eric Jang, Shixiang Gu, and Ben Poole. Categorical reparameterization with gumbel-softmax. *arXiv preprint arXiv:1611.01144*, 2016.
- [34] Yoshua Bengio, Nicholas Léonard, and Aaron Courville. Estimating or propagating gradients through stochastic neurons for conditional computation. *arXiv preprint arXiv:1308.3432*, 2013.
- [35] Will Grathwohl, Dami Choi, Yuhuai Wu, Geoffrey Roeder, and David Duvenaud. Backpropagation through the void: Optimizing control variates for black-box gradient estimation. *arXiv preprint arXiv:1711.00123*, 2017.
- [36] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [37] Ruslan Salakhutdinov and Iain Murray. On the quantitative analysis of deep belief networks. In *Proceedings of the 25th international conference on Machine learning*, pages 872–879, 2008.

- [38] Radford M Neal. Connectionist learning of belief networks. *Artificial intelligence*, 56(1): 71–113, 1992.
- [39] Mikell P. Groover. Encyclopædia britannica, May 2019. URL <https://www.britannica.com/technology/automation>.
- [40] Mai-Hanh Nguyen. The latest market research, trends, and landscape in the growing ai chatbot industry, Jan 2020. URL <https://www.businessinsider.com/chatbot-market-stats-trends?r=US&IR=T>.
- [41] Asbjørn Følstad, Petter Bae Brandtzaeg, Tom Feltwell, Effie L-C. Law, Manfred Tscheligi, and Ewa A. Luger. Sig: Chatbots for social good. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI EA '18, page 1–4, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450356213. doi: 10.1145/3170427.3185372. URL <https://doi.org/10.1145/3170427.3185372>.
- [42] Kayla Matthews. The dangers of weaponized chatbots, Aug 2019. URL <https://chatbotslife.com/the-dangers-of-weaponized-chatbots-900a0cefa08f>.