

---

# The Thermodynamic Variational Objective

---

**Vaden Masrani**  
Department of Computer Science  
University of British Columbia  
Vancouver, B.C. V6T 1Z4  
vadmas@cs.ubc.ca

**Tuan Anh Le**  
Department of Engineering Science  
University of Oxford  
tuananh@robots.ox.ac.uk

**Frank Wood**  
Department of Computer Science  
University of British Columbia  
Vancouver, B.C. V6T 1Z4  
fwood@cs.ubc.ca

## Abstract

We introduce the thermodynamic variational objective (TVO) for learning in both continuous and discrete deep generative models. The TVO arises from a key connection between variational inference and thermodynamic integration that results in a tighter lower bound to the log marginal likelihood than the standard variational evidence lower bound (ELBO), while remaining as broadly applicable. We provide a computationally efficient gradient estimator for the TVO that applies to continuous, discrete, and non-reparameterizable distributions and show that the objective functions used in variational inference, variational autoencoders, wake sleep, and inference compilation are all special cases of the TVO. We evaluate the TVO for learning of discrete and continuous variational auto encoders, and find it achieves state of the art for learning in discrete variable models, and outperform VAEs on continuous variable models without using the reparameterization trick.

## 1 Introduction

Unsupervised learning in richly structured deep latent variable models (Kingma and Welling, 2014; Rezende et al., 2014) remains challenging. Fundamental research directions include low-variance gradient estimation for discrete and continuous latent variable models (Mnih and Gregor, 2014; Mnih and Rezende, 2016; Tucker et al., 2017; Naesseth et al., 2017; Figurnov et al., 2018), tightening variational bounds in order to obtain better model learning (Burda et al., 2016; Maddison et al., 2017a; Le et al., 2018a; Naesseth et al., 2018), and alleviation of the associated detrimental effects on the learning of the inference network (Rainforth et al., 2018).

We present the thermodynamic variational objective (TVO) based on a key connection we make between thermodynamic integration (TI) and variational inference (VI), namely that the “instantaneous elbo” (Blei) that appears in VI is equivalent to the first derivative of the potential function that appears in TI. We demonstrate that optimizing the TVO leads to improved learning of both discrete and continuous latent-variable deep generative models. The gradient estimator we derive for optimizing the TVO does not require the high-variance REINFORCE estimator for discrete latent variables (Williams, 1992), nor the reparameterization trick which is only applicable to a limited family of continuous latent variables.

The TVO is a lower bound to the log evidence which can be made arbitrarily tight. We empirically show that optimizing the TVO results in better inference networks than optimizing the importance

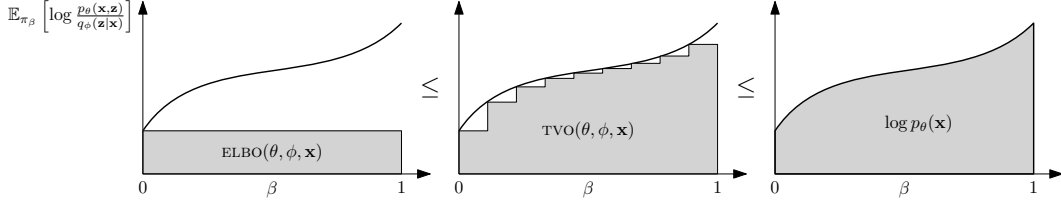


Figure 1: The Thermodynamic Variational Objective (center) is a finite sum numerical approximation to  $\log p_\theta(\mathbf{x})$ , defined by the Thermodynamic Variational Identity (right). The ELBO (left) is a single partition approximation of the same integral.

weighted autoencoder (IWAE) objective (Burda et al., 2016) for which tightening of the bound is known to make inference network learning worse (Rainforth et al., 2018). While this problem can be ameliorated by reducing the variance of the gradient estimator in the case of reparameterizable latent variables (Tucker et al., 2018), resolving it in the case of non-reparameterizable latent variables currently involves alternating optimization of model and inference networks (Hinton et al., 1995; Bornschein and Bengio, 2015; Le et al., 2018b).

## 2 The Thermodynamic Variational Objective

The evidence lower bound (ELBO), which is used in learning variational autoencoders (VAEs), lower bounds the log evidence of a generative model  $p_\theta(\mathbf{x}, \mathbf{z})$  parameterized by  $\theta$  of a latent variable  $\mathbf{z}$  and data  $\mathbf{x}$ , as it can be written as the log evidence minus a Kullback-Leibler (KL) divergence

$$\text{ELBO}(\theta, \phi, \mathbf{x}) := \log p_\theta(\mathbf{x}) - \text{KL}(q_\phi(\mathbf{z}|\mathbf{x})||p_\theta(\mathbf{z}|\mathbf{x})), \quad (1)$$

where  $q_\phi(\mathbf{z}|\mathbf{x})$  is an inference network parameterized by  $\phi$ . As illustrated in Figure 1, the TVO

$$\underbrace{\frac{1}{K} \left[ \text{ELBO}(\theta, \phi, \mathbf{x}) + \sum_{k=1}^{K-1} \mathbb{E}_{\pi_{\beta_k}} \left[ \log \frac{p_\theta(\mathbf{x}, \mathbf{z})}{q_\phi(\mathbf{z}|\mathbf{x})} \right] \right]}_{\text{TVO}(\theta, \phi, \mathbf{x})} \leq \underbrace{\int_0^1 \mathbb{E}_{\pi_\beta} \left[ \log \frac{p_\theta(\mathbf{x}, \mathbf{z})}{q_\phi(\mathbf{z}|\mathbf{x})} \right] d\beta}_{\text{THERMODYNAMIC VARIATIONAL IDENTITY}} = \log p_\theta(\mathbf{x}) \quad (2)$$

lower bounds the log evidence by using a Riemann sum approximation to the thermodynamic variational identity (TVI), a one-dimensional integral over a scalar  $\beta$  in a unit interval which evaluates to the log model evidence  $\log p_\theta(\mathbf{x})$ . The integrand, which is a function of  $\beta$ , is an expectation of the so-called “instantaneous ELBO” (Blei) under a geometric combination of  $p_\theta(\mathbf{x}, \mathbf{z})$  and  $q_\phi(\mathbf{z}|\mathbf{x})$  which we call  $\pi_\beta(\mathbf{z})$  and define in Section 3. Remarkably, at  $\beta = 0$ , the integrand equals the ELBO. This therefore allows us to view the ELBO as a single-term left Riemann sum of the TVI. At  $\beta = 1$ , the integrand equals to the evidence upper bound (EUBO). This sheds a new unifying perspective on the VAE and wake-sleep objectives, which we explore in detail in Section 5.

## 3 Connecting Thermodynamic Integration and Variational Inference

Suppose there are two unnormalized densities  $\tilde{\pi}_i(\mathbf{z})$  ( $i = 0, 1$ ) and corresponding normalizing constants  $Z_i := \int \tilde{\pi}_i(\mathbf{z}) d\mathbf{z}$ , which together define the normalized densities  $\pi_i(\mathbf{z}) := \tilde{\pi}_i(\mathbf{z})/Z_i$ . We can typically evaluate the unnormalized densities but cannot evaluate the normalizing constants. TI allows us to compute the log of the ratio of the normalizing constants,  $\log Z_1/Z_0$ . To do this, we first form a family of unnormalized densities (or a “path”) parameterized by  $\beta \in [0, 1]$  between the two distributions of interest

$$\tilde{\pi}_\beta(\mathbf{z}) := \tilde{\pi}_1(\mathbf{z})^\beta \tilde{\pi}_0(\mathbf{z})^{1-\beta} \quad (3)$$

with the corresponding normalizing constants and normalized densities

$$Z_\beta := \int \tilde{\pi}_\beta(\mathbf{z}) d\mathbf{z}, \quad \text{and} \quad \pi_\beta(\mathbf{z}) := \tilde{\pi}_\beta(\mathbf{z})/Z_\beta. \quad (4)$$

Following (Neal, 1993), we will find it useful to define a potential energy function  $U_\beta(\mathbf{z}) := \log \tilde{\pi}_\beta(\mathbf{z})$  along with its first derivative  $U'_\beta(\mathbf{z}) = \frac{dU_\beta(\mathbf{z})}{d\beta}$ . We can then estimate the log of the ratio of the normalizing constants via the identity central to TI, derived in Appendix B,

$$\log Z_1 - \log Z_0 = \int_0^1 \mathbb{E}_{\pi_\beta} [U'_\beta(\mathbf{z})] d\beta. \quad (5)$$

Our key insight connecting TI and VI is the following. If we set

$$\begin{aligned} \tilde{\pi}_0(\mathbf{z}) &:= q_\phi(\mathbf{z} | \mathbf{x}) & Z_0 &= \int q_\phi(\mathbf{z} | \mathbf{x}) d\mathbf{z} = 1 \\ \tilde{\pi}_1(\mathbf{z}) &:= p_\theta(\mathbf{x}, \mathbf{z}) & Z_1 &= \int p_\theta(\mathbf{x}, \mathbf{z}) d\mathbf{z} = p_\theta(\mathbf{x}) \end{aligned} \quad (6)$$

this results in a geometric path between the variational distribution  $q_\phi(\mathbf{z} | \mathbf{x})$  and the model  $p_\theta(\mathbf{x}, \mathbf{z})$

$$\tilde{\pi}_\beta(\mathbf{z}) = p_\theta(\mathbf{x}, \mathbf{z})^\beta q_\phi(\mathbf{z} | \mathbf{x})^{1-\beta} \quad \text{and} \quad \pi_\beta(\mathbf{z}) = \frac{p_\theta(\mathbf{x}, \mathbf{z})^\beta q_\phi(\mathbf{z} | \mathbf{x})^{1-\beta}}{Z_\beta}, \quad (7)$$

where the first derivative of the potential is equal to the ‘‘instantaneous ELBO’’ (Blei)

$$U'_\beta(\mathbf{z}) = \log \frac{p_\theta(\mathbf{x}, \mathbf{z})}{q_\phi(\mathbf{z} | \mathbf{x})}. \quad (8)$$

Substituting this identity and  $Z_0 = 1$  and  $Z_1 = p_\theta(\mathbf{x})$  into (5) results in the *thermodynamic variational identity*:

$$\log p_\theta(\mathbf{x}) = \int_0^1 \mathbb{E}_{\pi_\beta} \left[ \log \frac{p_\theta(\mathbf{x}, \mathbf{z})}{q_\phi(\mathbf{z} | \mathbf{x})} \right] d\beta. \quad (9)$$

This means that the log evidence  $\log p_\theta(\mathbf{x})$  can be expressed as a one-dimensional integral of an expectation of the instantaneous ELBO under  $\pi_\beta$  from  $\beta = 0$  to  $\beta = 1$  (see Figure 1 (right)).

To obtain the thermodynamic variational objective (TVO) defined in (2), we approximate the integral in (9) using a left Riemann sum. In order for the left Riemann sum to be a lower bound to the integral, we require the integrand to be increasing. The derivative of the integrand is equal to the variance of  $U'_\beta(\mathbf{z})$  under  $\pi_\beta(\mathbf{z})$  (see Appendix C) which means that it is non-negative, and so the integrand itself must be increasing. This is due to the fact that  $U'_\beta(\mathbf{z})$ , the instantaneous ELBO, is independent of  $\beta$  for our choice of the path in (7). For equal spacing of the partitions, where  $\beta_k = k/K$ , we arrive at the TVO in (2), illustrated in Figure 1 (middle). We present a generalized variant with non-equal spacing in Appendix A.

Maximization of the ELBO( $\theta, \phi, \mathbf{x}$ ) can be seen as a special case of the TVO, since for  $\beta = 0$ ,  $\pi_\beta(\mathbf{z}) = q_\phi(\mathbf{z} | \mathbf{x})$  so the integrand in (9) becomes  $\mathbb{E}_{q_\phi(\mathbf{z} | \mathbf{x})} \left[ \log \frac{p_\theta(\mathbf{x}, \mathbf{z})}{q_\phi(\mathbf{z} | \mathbf{x})} \right]$  which is equivalent to the definition of ELBO in (1). Because the integrand is increasing, we have

$$\text{ELBO}(\theta, \phi, \mathbf{x}) \leq \text{TVO}(\theta, \phi, \mathbf{x}) \leq \log p_\theta(\mathbf{x}), \quad (10)$$

which means that the TVO is an alternative to IWAE for tighten the variational bounds. In Appendix D we show maximizing the TVO is equivalent to minimizing a divergence between the variational distribution and the true posterior  $p_\theta(\mathbf{z} | \mathbf{x})$ .

The integrand in (9) is typically estimated by long running Markov chain Monte Carlo chains computed at different values of  $\pi_\beta(\mathbf{z})$  (Friel and Pettitt, 2008; Lartillot and Philippe, 2006). Instead, we propose a simple importance sampling mechanism that allows us reuse samples across an arbitrary number of discretizations and which is compatible with gradient-based learning.

## 4 Optimizing the TVO

We now provide a novel gradient estimator for the TVO which doesn’t involve the REINFORCE estimator nor the reparameterization trick.

**Gradient of TVO.** The gradient of TVO consists of terms of the form  $\nabla_\lambda \mathbb{E}_{\pi_{\lambda,\beta}}[f_\lambda(\mathbf{z})]$  for  $\pi_{\lambda,\beta}(\mathbf{z})$  and  $f_\lambda(\mathbf{z})$  which are parameterized by  $\lambda$  (say  $\lambda := (\theta, \phi)$ ). We can compute such terms as

$$\nabla_\lambda \mathbb{E}_{\pi_{\lambda,\beta}}[f_\lambda(\mathbf{z})] = \mathbb{E}_{\pi_{\lambda,\beta}}[\nabla_\lambda f_\lambda(\mathbf{z})] + \text{Cov}_{\pi_{\lambda,\beta}}[\nabla_\lambda \log \tilde{\pi}_{\lambda,\beta}(\mathbf{z}), f_\lambda(\mathbf{z})] \quad (11)$$

as shown in Appendix E. The covariance in (11) has the same dimensionality as  $\lambda \in \mathbb{R}^D$  because it is between  $\nabla_\lambda \log \tilde{\pi}_{\lambda,\beta}(\mathbf{z}) \in \mathbb{R}^D$  and  $f_\lambda(\mathbf{z}) \in \mathbb{R}$  and is defined as

$$\text{Cov}_{\pi_{\lambda,\beta}}(\mathbf{a}, b) := \mathbb{E}_{\pi_{\lambda,\beta}}[(\mathbf{a} - \mathbb{E}_{\pi_{\lambda,\beta}}[\mathbf{a}])(b - \mathbb{E}_{\pi_{\lambda,\beta}}[b])]. \quad (12)$$

Thus, estimating the gradient in (11) requires estimating expectations under  $\pi_\beta$ .

**Estimating Expectations.** We can estimate an expectation of a general function  $f(\mathbf{z})$  under  $\pi_\beta(\mathbf{z})$  using  $S$ -sample importance sampling using  $q_\phi(\mathbf{z}|\mathbf{x})$  as the proposal distribution

$$\mathbb{E}_{\pi_\beta}[f(\mathbf{z})] \approx \sum_{s=1}^S \overline{w_s^\beta} f(\mathbf{z}_s), \quad (13)$$

where  $\mathbf{z}_s \sim q_\phi(\mathbf{z}|\mathbf{x})$ ,  $\overline{w_s^\beta} := w_s^\beta / \sum_{s'=1}^S w_{s'}^\beta$  and  $w_s := \frac{p_\theta(\mathbf{x}, \mathbf{z}_s)}{q_\phi(\mathbf{z}_s|\mathbf{x})}$ , since each unnormalized importance weight can be expressed as

$$\frac{\tilde{\pi}_\beta(\mathbf{x}, \mathbf{z}_s)}{q_\phi(\mathbf{z}_s|\mathbf{x})} = \frac{p_\theta(\mathbf{x}, \mathbf{z}_s)^\beta q_\phi(\mathbf{z}_s|\mathbf{x})^{1-\beta}}{q_\phi(\mathbf{z}_s|\mathbf{x})} = \frac{p_\theta(\mathbf{x}, \mathbf{z}_s)^\beta}{q_\phi(\mathbf{z}_s|\mathbf{x})^\beta} = \left( \frac{p_\theta(\mathbf{x}, \mathbf{z}_s)}{q_\phi(\mathbf{z}_s|\mathbf{x})} \right)^\beta = w_s^\beta. \quad (14)$$

We can reuse the  $S$  samples  $\mathbf{z}_s \sim q_\phi(\mathbf{z}|\mathbf{x})$ , since evaluating the normalized weight  $\overline{w_s^{\beta_k}}$  only requires raising each weight to different powers of  $\beta_k$  before normalizing. Reusing  $S$  samples instead of sampling  $SK$  times reduces the variance of the estimator (Owen, 2013).

The gradient estimator in (11) doesn't require  $\mathbf{z}$  to be reparameterizable, nor does it use the potentially high-variance REINFORCE estimator. This means that it can be used in the case of non-reparameterizable continuous latent variables as well as in the case of discrete latent variables without modifying the model using continuous relaxations (Jang et al., 2017; Maddison et al., 2017b).

## 5 Generalizing Variational Objectives

The wake-sleep (ws) (Hinton et al., 1995) and reweighted wake-sleep (RWS) (Bornschein and Bengio, 2015; Le et al., 2018b) algorithms have traditionally been viewed as optimizing two separate objectives. We can now view ws as alternating between maximizing a lower bound and minimizing an upper bound of a Riemann approximation to the TVI.

Using the right Riemann sum to approximate the TVI, we can upper-bound  $\log p_\theta(\mathbf{x})$ . With  $K$  equally spaced partitions where  $\beta_k = k/K$ , we define the upper-bound variant of the TVO as

$$\text{TVO}_K^U(\theta, \phi, \mathbf{x}) := \frac{1}{K} \left[ \text{EUBO}(\theta, \phi, \mathbf{x}) + \sum_{k=1}^{K-1} \mathbb{E}_{\pi_{\beta_k}} \left[ \log \frac{p_\theta(\mathbf{x}, \mathbf{z})}{q_\phi(\mathbf{z}|\mathbf{x})} \right] \right] \geq \log p(\mathbf{x}) \quad (15)$$

where the EUBO is defined

$$\text{EUBO}(\theta, \phi, \mathbf{x}) := \mathbb{E}_{p_\theta(\mathbf{z}|\mathbf{x})} \left[ \log \frac{p_\theta(\mathbf{x}, \mathbf{z})}{q_\phi(\mathbf{z}|\mathbf{x})} \right] \quad (16)$$

and is analogous to the ELBO under the true posterior at  $\beta_k = 1$ .

Methods that optimize the ELBO—variational inference (Blei et al., 2017), variational autoencoders (Kingma and Welling, 2014; Rezende et al., 2014), and the wake-phase update of ws (Hinton et al., 1995)—can be viewed as optimizing a one-term left Riemann sum approximation of the TVI. Algebraically, this is equivalent to setting  $K = 1$  and dropping the additional terms in the summation of (2). We now show how the sleep-phase update of ws and the wake-phase  $\phi$  update of RWS can be seen as minimizing the  $\text{TVO}_K^U(\theta, \phi, \mathbf{x})$  with  $K = 1$ .

**Sleep-phase  $\phi$  update.** In the sleep phase  $\phi$  update, we consider  $\theta$  fixed and minimize  $\text{TVO}_1^U$  using “dream” data  $\{\mathbf{x}_i\}_{i=1}^N \sim p_\theta(\mathbf{x})$  and a single partition.

$$\phi^* = \arg \min_{\phi} \mathbb{E}_{\mathbf{x} \sim p_\theta(\mathbf{x})} [\text{TVO}_1^U(\theta, \phi, \mathbf{x})] \quad (17)$$

$$= \arg \min_{\phi} \int p_\theta(\mathbf{x}) \left[ \mathbb{E}_{p_\theta(\mathbf{z}|\mathbf{x})} \left[ \log \frac{p(\mathbf{x}, \mathbf{z})}{q_\phi(\mathbf{z}|\mathbf{x})} \right] \right] d\mathbf{x} \quad (18)$$

$$= \arg \min_{\phi} \mathbb{E}_{p_\theta(\mathbf{x}, \mathbf{z})} [-\log q_\phi(\mathbf{z}|\mathbf{x})] \quad (19)$$

This objective is the same as the inference compilation objective (Le et al., 2018b).

**Wake-phase  $\phi$  update.** In the wake phase  $\phi$  update, we instead use real data  $\{\mathbf{x}_i\}_{i=1}^N \sim p(\mathbf{x})$  and again minimize  $\text{TVO}_1^U$ :

$$\phi^* = \arg \min_{\phi} \mathbb{E}_{\mathbf{x} \sim p(\mathbf{x})} [\text{TVO}_1^U(\theta, \phi, \mathbf{x})] \quad (20)$$

$$= \arg \min_{\phi} \mathbb{E}_{\mathbf{x} \sim p(\mathbf{x})} \left[ \mathbb{E}_{p(\mathbf{z}|\mathbf{x})} \left[ \log \frac{p(\mathbf{x}, \mathbf{z})}{q_\phi(\mathbf{z}|\mathbf{x})} \right] \right] \quad (21)$$

The gradient estimator for performing this update given in (Le et al., 2018b) is equivalent to the gradient estimator obtained via equations (11) and (13).

## 6 Related Work

Thermodynamic integration was originally developed in physics to calculate the difference in free energy of two molecular systems (Evans, 1986). Neal (1993) and Gelman and Meng (1998) then introduced TI into the statistics community to calculate the ratios of normalizing constants of general probability models. TI is now commonly used in phylogenetics to calculate the Bayes factor  $B = p(x|M_1)/p(x|M_0)$ , where  $M_0, M_1$  are two models specifying (for instance) tree topologies and branch lengths (Lartillot and Philippe, 2006; Xie et al., 2010; Rodrigue and Aris-Brosou, 2011). We took inspiration from (Fan et al., 2010) who replaced the “power posterior”  $p(\theta|\mathbf{x}, M, \beta) = p(x|\theta, M)^\beta p(\theta, M)/Z_\beta$  of (Xie et al., 2010) with  $p(\theta|\mathbf{x}, M, \beta) = [p(\mathbf{x}|\theta, M)p(\theta|M)]^\beta [p_0(\theta|M)]^{1-\beta}/Z_\beta$ , where  $p_0(\theta|M)$  is a tractable reference distribution chosen to facilitate sampling.

To the best of our knowledge, our work is the first to explicitly connect TI and VI. Grosse et al. (2013) studied annealed importance sampling (AIS), a related technique that estimates partition functions using a sequence of intermediate distributions using a product of ratios of importance weights. They observe the geometric path taken in AIS is equivalent to minimizing a weighted sum of KL divergences and use this insight to motivate their alternative moment-averaged path. In contrast, we observe the instantaneous ELBO common to all the above method is equivalent to the first derivative of the potential (cf. Section 9) and use this to propose a new optimization procedure.

## 7 Experiments

### 7.1 Discrete Deep Generative Models

We optimize the TVO to learn the parameters of a deep generative model with discrete latent variables, also known as sigmoid belief networks (Neal, 1992). This family of models is used to evaluate objectives, continuous relaxations and control variate methods for learning discrete latent variable models (Mnih and Gregor, 2014; Maddison et al., 2017b; Jang et al., 2017; Mnih and Rezende, 2016; Bornschein and Bengio, 2015; Tucker et al., 2017; Grathwohl et al., 2018).

We use the binarized MNIST dataset with the standard train/validation/test split of 50000/10000/10000 (Salakhutdinov and Murray, 2008). The generative model is of the form  $p(\mathbf{z}_{1:L}, \mathbf{x}) = p(\mathbf{z}_L) \prod_{\ell=1}^{L-1} p(\mathbf{z}_\ell|\mathbf{z}_{\ell+1})p(\mathbf{x}|\mathbf{z}_1)$  where each conditional on  $\mathbf{z}_\ell$  is an independent Bernoulli whose parameters are a linear function of  $\mathbf{z}_{\ell+1}$ . The likelihood  $p(\mathbf{x}|\mathbf{z}_1)$  is also an independent Bernoulli whose parameters are a linear function of  $\mathbf{z}_1$ . We also parameterize the prior  $p(\mathbf{z}_L)$ . The inference network is factorized in an opposite way where  $q(\mathbf{z}|\mathbf{x}) = q(\mathbf{z}_1|\mathbf{x}) \prod_{\ell=2}^L q(\mathbf{z}_\ell|\mathbf{z}_{\ell-1})$ . Here, each

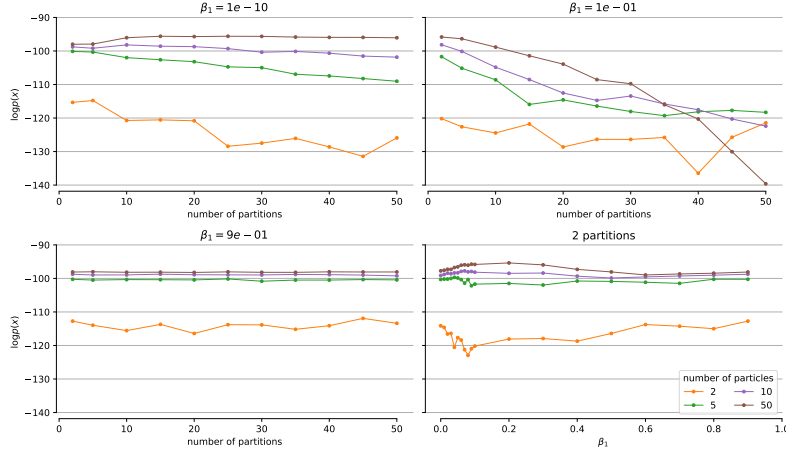


Figure 2: Investigation of how number of particles  $S$ , number of partitions  $K$ , and  $\beta_1$  affect learning of the generative model. In the first three plots, we vary  $S$  and  $K$  for three various  $\beta_1$  and see that while  $S$  should be as high as possible, there is a sweet spot for  $K$ . Often  $K = 2$  results in good performance, assuming  $\beta_1$  is well-chosen, as seen in the last plot.

conditional is an independent Bernoulli whose parameters are linear functions of the condition. We set  $L = 2$  and the dimensionality of each  $\mathbf{z}_\ell$  to 200. We optimize the TVO using the Adam optimizer with default parameters.<sup>1</sup>

**The effect of  $S$ ,  $K$ , and  $\beta$  locations** In the first set of experiments, we sought to understand the behavior of choosing the discretization  $\beta_{0:K}$ , number of partitions  $K$  and number of particles  $S$ .

We expect that increasing the number of partitions  $K$  makes the Riemann sum approximate the integral over  $\beta$  more tightly. However, with each addition term we increase bias due to the use of importance sampling to estimate the expectations. Further, importance sampling estimates of points on the curve further to the right are likely to be more biased, since  $q$  has to approximate an increasingly less annealed  $p$ . We found that combination of these two effects means that there is a “sweet spot”, or an optimal number of partitions beyond which additional of partitions will be detrimental to performance.

We also found that the location of the discretization points affects performance, as we have empirically observed the curve in Figure 1 often has a region of high curvature near  $\beta = 0$ . To investigate this, instead of searching over all configurations of  $\beta_{0:K}$ , we focus only on  $\beta_1$ , with the rest  $\beta_{2:K}$  being equally spaced on the logarithmic scale. We hypothesized that if  $\beta_1$  is located beyond the region of high curvature, adding additional partitions would incur a high cost of bias without capturing much additional area.

To investigate these effects, we performed a sweep over the grid of  $K \in \{2, 5, 10, 15, \dots, 50\}$  and  $S \in \{2, 5, 10, 50\}$  for a small ( $10^{-10}$ ), medium (0.1) and large (0.9)  $\beta_1$ . In figure Figure 2 we show the test log evidence  $\log p(x)$  at the last iteration, approximated by evaluating the IWAE loss with 5000 samples.

The first effect we see is that increasing number of samples  $S$ —which decreases importance sampling bias per partition—improves performance. We also that two partitions are often enough to obtain good performance. In the top left and bottom right plots, we also see that if  $\beta_1$  is chosen too large, the Riemann sum cannot recover the “lost” area even if number of partitions is increased.

In our second experiment, we fixed  $K = 2$  and investigate the quality of the learned generative model for different  $\beta_1$ , shown in the bottom right plot of Figure 2. For each  $S \in \{2, 5, 10, 50\}$ , there is an optimal  $\beta_1$  that is neither the smallest, nor the largest from the ones we searched over. We have empirically observed that the curve in Figure 1 is often rising sharply from  $\beta = 0$  until a point  $\beta_*$  after which it is almost flat until  $\beta = 1$ . Thus, we hope for  $\beta_1$  to match  $\beta_*$ .

<sup>1</sup>We include further details in Appendix H.

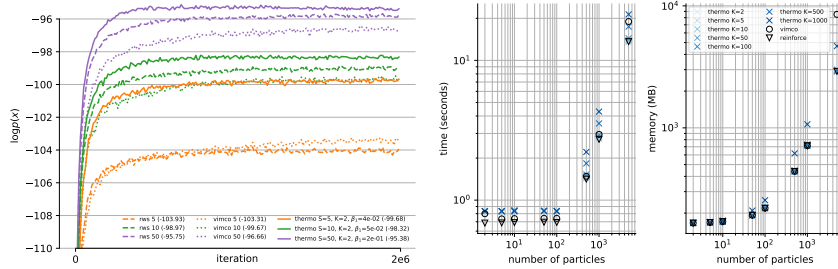


Figure 3: Comparisons with baselines on a held out test set. (Left) Learning curves for different methods. TVO outperforms other methods both in terms of speed of convergence and the learned model. (Right) Time and memory efficiency TVO with increasing number of partitions vs baselines, measured for 100 iterations of optimization. Increasing the number of partitions is much cheaper than increasing the number of particles.

Given these results, we recommend using as many particles  $S$  as possible and performing a hyperparameter search over  $\beta_1$  (with  $K = 2$  being fixed) when using the TVO objective.

**Baseline Comparison** In Figure 3 (left), we compare the TVO with RWS, where we use only the wake objective to learn  $\phi$  (which is the better one out of the two) and VIMCO which are state-of-the-art IWAE-based methods for learning discrete latent variable models (Le et al., 2018b). TVO outperforms both methods in terms of speed of convergence and the final test log evidence  $\log p_\theta(\mathbf{x})$  estimated using 5000 IWAE particles as before.

Since we use  $K = 2$  partitions for the same number of particles  $S$ , the time and memory complexity of TVO is double that of other methods. While this is true, in both time and memory cases, the constant factor for increasing  $S$  is much higher than for increasing  $K$ . As shown in Figure 3 (right), it is virtually free to increase number of partitions. This is because for each new particle, we must additionally sample from the inference network and score the sample under both  $p$  and  $q$  to obtain the weight. On the other hand, we can reuse the  $S$  samples and corresponding weights in estimating values for the  $K + 1$  terms in the Riemann sum. Thus, the part of the computation graph (which gets evaluated in both the forward and backward passes) that is dependent on  $K$  is after the expensive sampling and scoring. The part dependent on  $K$  only involves creating a few matrices of size  $S \times K$  and performing basic operations on them.

Finally, we investigate the quality of the learned inference network by plotting the KL divergence (averaged over the test set) between the current  $q$  and current  $p$  as training progresses (Figure 4 (left)). This indicates how well  $q$  “tracks”  $p$ . This is estimated as log evidence minus ELBO where the former is estimated as before and the latter is estimated using 5000 Monte Carlo samples. The KL is lowest for TVO. For all methods, increasing number of particles makes the KL worse. We hypothesize that because it is harder for  $q$  to track a  $p$  that is learning better. In Figure 4 (right), we plot the gradient variance of the  $\phi$  gradient estimator for all methods. The gradient estimator of TVO is lower variance than VIMCO but higher variance than RWS. This indicates that TVO is a new, potentially advantageous point on the bias-variance trade-off curve along with REINFORCE (high variance, zero bias), VIMCO (lower variance, zero bias), RWS (even lower variance, biased) but which allows us to learn better and faster.

## 7.2 Continuous Deep Generative Models

We have performed the same set of experiments on the same dataset but with a deep generative model with continuous latent variables. The generative model is of the form  $p(\mathbf{z})p_\theta(\mathbf{x}|\mathbf{z}) = \text{Normal}(\mathbf{z}|0, I)\text{Bernoulli}(\mathbf{x}|\text{decoder}_\theta(\mathbf{z}))$ , where  $\mathbf{z}$  is 200-dimensional and  $\text{decoder}_\theta$  is a three-layer multilayer perceptron with tanh activations and sigmoid output which parameterizes the probabilities of the independent Bernoulli distribution. The inference network is of the form  $q_\phi(\mathbf{z}|\mathbf{x}) = \text{Normal}(\mathbf{z}|\text{encoder}_\phi(\mathbf{x}))$ , where the encoder is a two-layer multilayer perceptron with tanh activations and whose output is passed through two separate linear layers which output the

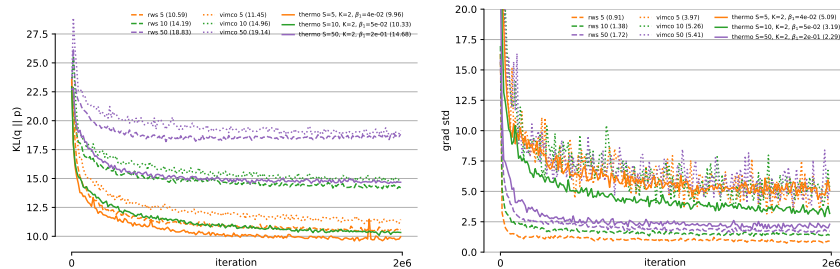


Figure 4: Investigating the inference network. (Left) KL divergence between current  $q$  and  $p$  (which measures how well  $q$  “tracks”  $p$ ) is lowest for TVO. (Right) Standard deviation of  $\phi$  gradient estimators. RWS is lowest variance, VIMCO is highest variance, TVO is in the middle.

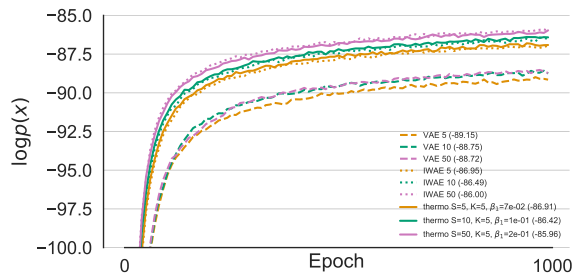


Figure 5: Learning curves for continuous deep generative models for different objectives. Despite not using the reparameterization trick, TVO outperforms VAEs and is competitive with IWAE.

means and the log standard deviations of the independent Normal distribution. We optimize the objectives using the Adam optimizer with default parameters.<sup>2</sup>

The influence of  $\beta_1$ , number of partitions  $K$  and number of particles  $S$  on performance are similar to the discrete case. In Figure 5, we compare TVO with IWAE and VAE with equivalent computational budget. In the case of VAE, we use  $S$  Monte Carlo terms to approximate the ELBO gradient while in IWAE, the sum over  $S$  instantaneous ELBO terms is inside the log. TVO outperforms VAE and performs competitively with IWAE despite the latter two using the reparameterization trick.

## 8 Conclusions

The thermodynamic variational objective represents a new way to tighten evidence bounds and is based on a novel choice of thermodynamic integration path that tightly connects variational inference and thermodynamic integration. We demonstrated that optimizing the TVO can have a positive impact on the learning of discrete deep generative models and can perform as well as using the reparameterization trick to learn continuous deep generative models. This is a subtle point worth emphasizing. Tightening evidences bound seems to nearly universally help model learning, but gradients estimated using evidence bounds from the importance sampling family seem to negatively impact the inference network. In the end the importance sampling family gradients are not only higher variance but also must be, according to our experimental results, subtly biased so as to lead to less optimal final models. Put another way, the joint optimization path provided by the TVO-derived gradients seems to implicitly regularize towards better models. This and the near equivalence of learning using TVO gradients without reparameterization to IWAE with reparameterization in deep continuous latent variable models encourages much closer examination and experimentation.

The approximate path integration perspective provided by our development of the TVO also sheds light on the connection between otherwise disparate deep generative model learning techniques. In particular, the TVO integration perspective points to ways to improve wake-sleep via tightening the EUBO using similar integral upper-bounding techniques. Further experimentation is warranted to

<sup>2</sup>We include further details in Appendix H.



explore how TVO insights can be applied to all special cases of the TVO including non-amortized variational inference and to the use of the TVO as a compliment to annealing importance sampling for final model evidence evaluation. Our unreported initial experiments along these lines show small but positive effects.

The TVO, specifically in the  $K = 2$  case that we use in our experiments, does come with a pesky requirement: choosing  $\beta_1$ . The choice of  $\beta_1$  is more important than we would ideally wish for it to be, and grid-searching for an optimal value that maximally fills the area is not ideal. This does, however, point out opportunities for future work wherein we adaptively select optimal positions of the  $\beta_{1:K}$  points, perhaps borrowing techniques from the Bayesian numerical quadrature literature (O’Hagan, 1991; Rasmussen and Ghahramani, 2003; Osborne et al., 2012).

## Acknowledgments

We would like to thank Trevor Campbell for his constant support throughout this project, as well as Adam Ścibior, Boyan Beronov, and Saifuddin Syed for their helpful comments on early drafts of this manuscript.

## References

- David Blei. Variational inference: Foundations and innovations. URL [http://www.cs.columbia.edu/~blei/talks/Blei\\_VI\\_tutorial.pdf](http://www.cs.columbia.edu/~blei/talks/Blei_VI_tutorial.pdf).
- David M Blei, Alp Kucukelbir, and Jon D McAuliffe. Variational inference: A review for statisticians. *Journal of the American Statistical Association*, 112(518):859–877, 2017.
- Jörg Bornschein and Yoshua Bengio. Reweighted wake-sleep. In *International Conference on Learning Representations*, 2015.
- Yuri Burda, Roger Grosse, and Ruslan Salakhutdinov. Importance weighted autoencoders. In *International Conference on Learning Representations*, 2016.
- Shinto Eguchi et al. A differential geometric approach to statistical inference on the basis of contrast functionals. *Hiroshima mathematical journal*, 15(2):341–391, 1985.
- DJ Evans. Molecular dynamics simulation of statistical mechanical systems. *International School of Physics, “Emico Fermi” (July 22-August 2, 1985), to be published*, 1986.
- Yu Fan, Rui Wu, Ming-Hui Chen, Lynn Kuo, and Paul O Lewis. Choosing among partition models in bayesian phylogenetics. *Molecular biology and evolution*, 28(1):523–532, 2010.
- Mikhail Figurnov, Shakir Mohamed, and Andriy Mnih. Implicit reparameterization gradients. In *Advances in Neural Information Processing Systems*, pages 441–452, 2018.
- Nial Friel and Anthony N Pettitt. Marginal likelihood estimation via power posteriors. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(3):589–607, 2008.
- Andrew Gelman and Xiao-Li Meng. Simulating normalizing constants: From importance sampling to bridge sampling to path sampling. *Statistical science*, pages 163–185, 1998.
- Will Grathwohl, Dami Choi, Yuhuai Wu, Geoff Roeder, and David Duvenaud. Backpropagation through the void: Optimizing control variates for black-box gradient estimation. In *International Conference on Learning Representations*, 2018.
- Roger B Grosse, Chris J Maddison, and Ruslan R Salakhutdinov. Annealing between distributions by averaging moments. In *Advances in Neural Information Processing Systems*, pages 2769–2777, 2013.
- 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.
- Eric Jang, Shixiang Gu, and Ben Poole. Categorical reparameterization with Gumbel-softmax. In *International Conference on Learning Representations*, 2017.
- Diederik P Kingma and Max Welling. Auto-encoding variational Bayes. In *International Conference on Learning Representations*, 2014.

- Nicolas Lartillot and Hervé Philippe. Computing bayes factors using thermodynamic integration. *Systematic biology*, 55(2):195–207, 2006.
- Tuan Anh Le, Maximilian Igl, Tom Rainforth, Tom Jin, and Frank Wood. Auto-encoding sequential Monte Carlo. In *International Conference on Learning Representations*, 2018a.
- Tuan Anh Le, Adam R Kosiorek, N Siddharth, Yee Whye Teh, and Frank Wood. Revisiting reweighted wake-sleep. *arXiv preprint arXiv:1805.10469*, 2018b.
- P L’Ecuyer. On the interchange of derivative and expectation. *Management Science*, to appear, 1993.
- Chris J Maddison, John Lawson, George Tucker, Nicolas Heess, Mohammad Norouzi, Andriy Mnih, Arnaud Doucet, and Yee Teh. Filtering variational objectives. In *Advances in Neural Information Processing Systems*, pages 6576–6586, 2017a.
- Chris J Maddison, Andriy Mnih, and Yee Whye Teh. The concrete distribution: A continuous relaxation of discrete random variables. In *International Conference on Learning Representations*, 2017b.
- Andriy Mnih and Karol Gregor. Neural variational inference and learning in belief networks. In *International Conference on Machine Learning*, pages 1791–1799, 2014.
- Andriy Mnih and Danilo Rezende. Variational inference for Monte Carlo objectives. In *International Conference on Machine Learning*, pages 2188–2196, 2016.
- Christian Naesseth, Francisco Ruiz, Scott Linderman, and David Blei. Reparameterization gradients through acceptance-rejection sampling algorithms. In *Artificial Intelligence and Statistics*, pages 489–498, 2017.
- Christian Naesseth, Scott Linderman, Rajesh Ranganath, and David Blei. Variational sequential Monte Carlo. In *International Conference on Artificial Intelligence and Statistics*, 2018.
- Radford M Neal. Connectionist learning of belief networks. *Artificial intelligence*, 56(1):71–113, 1992.
- Radford M Neal. Probabilistic inference using markov chain monte carlo methods. 1993.
- Anthony O’Hagan. Bayes–hermite quadrature. *Journal of statistical planning and inference*, 29(3):245–260, 1991.
- Michael Osborne, Roman Garnett, Zoubin Ghahramani, David K Duvenaud, Stephen J Roberts, and Carl E Rasmussen. Active learning of model evidence using Bayesian quadrature. In *Advances in neural information processing systems*, pages 46–54, 2012.
- Art B. Owen. *Monte Carlo theory, methods and examples*. 2013.
- 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. In *ICML*, 2018.
- Carl Edward Rasmussen and Zoubin Ghahramani. Bayesian Monte Carlo. *Advances in neural information processing systems*, pages 505–512, 2003.
- Danilo Jimenez Rezende, Shakir Mohamed, and Daan Wierstra. Stochastic backpropagation and approximate inference in deep generative models. In *International Conference on Machine Learning*, 2014.
- Nicolas Rodrigue and Stéphane Aris-Brosou. Fast bayesian choice of phylogenetic models: Prospecting data augmentation–based thermodynamic integration. *Systematic biology*, 60(6):881–887, 2011.
- 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. ACM, 2008.
- 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 2624–2633, 2017.
- George Tucker, Dieterich Lawson, Shixiang Gu, and Chris J Maddison. Doubly reparameterized gradient estimators for monte carlo objectives. *arXiv preprint arXiv:1810.04152*, 2018.
- Ronald J Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning*, 8(3-4):229–256, 1992.
- Wangang Xie, Paul O Lewis, Yu Fan, Lynn Kuo, and Ming-Hui Chen. Improving marginal likelihood estimation for bayesian phylogenetic model selection. *Systematic biology*, 60(2):150–160, 2010.

## A The generalized TVO

The TVO presented in Section 3 is a lower bound to  $\log p_\theta(\mathbf{x})$  using a left Riemann sum approximation to the thermodynamic variational identity. Using the right Riemann sum results in an upper bound which can be minimized (rather than maximized) during optimization. This loss is used in the inference compilation and during the sleep-phase  $\phi$  update in the Wake-Sleep algorithm. Below we present both the upper-bound and lower-bound variants of the TVO, with non-equally spaced partitions  $0 = \beta_0 < \beta_1 < \dots < \beta_K = 1$ ,  $\Delta_{\beta_k} = \beta_k - \beta_{k-1}$ ,  $K > 1$

$$\text{TVO}_K^L(\theta, \phi, \mathbf{x}) := \Delta_{\beta_1} \text{ELBO}(\theta, \phi, \mathbf{x}) + \sum_{k=2}^K \Delta_{\beta_k} \mathbb{E}_{\pi_{\beta_{k-1}}} \left[ \log \frac{p_\theta(\mathbf{x}, \mathbf{z})}{q_\phi(\mathbf{z} | \mathbf{x})} \right] \leq \log p(\mathbf{x}), \quad (22)$$

$$\text{TVO}_K^U(\theta, \phi, \mathbf{x}) := \Delta_{\beta_K} \text{EUBO}(\theta, \phi, \mathbf{x}) + \sum_{k=1}^{K-1} \Delta_{\beta_k} \mathbb{E}_{\pi_{\beta_k}} \left[ \log \frac{p_\theta(\mathbf{x}, \mathbf{z})}{q_\phi(\mathbf{z} | \mathbf{x})} \right] \geq \log p(\mathbf{x}), \quad (23)$$

where

$$\begin{aligned} \text{ELBO}(\theta, \phi, \mathbf{x}) &:= \mathbb{E}_{q_\phi(\mathbf{z} | \mathbf{x})} \left[ \frac{p_\theta(\mathbf{x}, \mathbf{z})}{q_\phi(\mathbf{z} | \mathbf{x})} \right], & \text{EUBO}(\theta, \phi, \mathbf{x}) &:= \mathbb{E}_{p_\theta(\mathbf{z} | \mathbf{x})} \left[ \frac{p_\theta(\mathbf{x}, \mathbf{z})}{q_\phi(\mathbf{z} | \mathbf{x})} \right], \\ \pi_{\beta_k}(\mathbf{z}) &:= p_\theta(\mathbf{x}, \mathbf{z})^\beta q_\phi(\mathbf{z} | \mathbf{x})^{1-\beta} / Z_\beta, & Z_\beta &:= \int p_\theta(\mathbf{x}, \mathbf{z})^\beta q_\phi(\mathbf{z} | \mathbf{x})^{1-\beta} d\mathbf{z}. \end{aligned}$$

## B Thermodynamic Integration

TI is a technique used in physics and phylogenetics to approximate intractable normalized constants of high dimensional distributions (Neal, 1993; Gelman and Meng, 1998). It is based on the observation that it is easier to calculate the ratio of two unknown normalizing constants than it is to calculate the constants themselves. More formally, consider two densities over space  $\mathcal{Z}$

$$\pi_i(\mathbf{z}) = \frac{\tilde{\pi}_i(\mathbf{z})}{Z_i}, \quad Z_i = \int_{\mathcal{Z}} \tilde{\pi}_i(\mathbf{z}) d\mathbf{z}, \quad i \in \{0, 1\}. \quad (24)$$

To apply TI, we form a continuous family (or “path”) between  $\pi_0(\mathbf{z})$  and  $\pi_1(\mathbf{z})$  via a scalar parameter  $\beta \in [0, 1]$

$$\pi_\beta(\mathbf{z}) = \frac{\tilde{\pi}_\beta(\mathbf{z})}{Z_\beta} = \frac{\tilde{\pi}_0(\mathbf{z})^\beta \tilde{\pi}_1(\mathbf{z})^{1-\beta}}{Z_\beta}, \quad Z_\beta = \int_{\mathcal{Z}} \tilde{\pi}_\beta(\mathbf{z}) d\mathbf{z}, \quad \beta \in [0, 1]. \quad (25)$$

The central identity that allows us to compute the ratio  $\log(Z_1/Z_0)$  is derived as follows. Assuming the legitimacy of exchanging integration with differentiation,

$$\begin{aligned} \frac{\partial \log Z_\beta}{\partial \beta} &= \frac{1}{Z_\beta} \frac{\partial}{\partial \beta} Z_\beta \\ &= \frac{1}{Z_\beta} \frac{\partial}{\partial \beta} \int \tilde{\pi}_\beta(\mathbf{z}) d\mathbf{z} \\ &= \int \frac{1}{Z_\beta} \frac{\partial}{\partial \beta} \tilde{\pi}_\beta(\mathbf{z}) d\mathbf{z} \\ &= \int \frac{\tilde{\pi}_\beta(\mathbf{z})}{Z_\beta} \frac{\partial}{\partial \beta} \log \tilde{\pi}_\beta(\mathbf{z}) d\mathbf{z}, \end{aligned}$$

which directly implies

$$\frac{\partial \log Z_\beta}{\partial \beta} = \mathbb{E}_{\pi_\beta} [U'_\beta(\mathbf{z})], \quad (26)$$

where the quantity  $U_\beta(\mathbf{z}) = \log \tilde{\pi}_\beta(\mathbf{z})$  is referred to as the “potential” in statistical physics and  $U'_\beta(\mathbf{z}) := \frac{\partial}{\partial \beta} U_\beta(\mathbf{z})$ .  $\beta$  can be interpreted as the inverse temperature parameter. Because one can typically compute  $\log \tilde{\pi}_\beta(\mathbf{z})$ , (26) allows us to exchange the first derivative of something we cannot

compute with an expectation over something we can compute. Then, to calculate the ratio  $\log(Z_1/Z_0)$  we integrate out  $\beta$  on both sides of (26)

$$\int_0^1 \frac{\partial \log Z_\beta}{\partial \beta} d\beta = \int_0^1 \mathbb{E}_{\pi_\beta} [U'_\beta(\mathbf{z})] d\beta, \quad \text{which results in} \quad (27)$$

$$\log(Z_1) - \log(Z_0) = \int_0^1 \mathbb{E}_{\pi_\beta} [U'_\beta(\mathbf{z})] d\beta. \quad (28)$$

## C The Increasing Integrand

### C.1 Notation

$$\log p_\theta(\mathbf{x}) = \int_0^1 g(\beta) d\beta \quad (29)$$

$$g(\beta) = \mathbb{E}_{\pi_\beta(\mathbf{z})} [U'(\mathbf{z})] \quad (30)$$

$$U'(\mathbf{z}) = \log \frac{p_\theta(\mathbf{x}, \mathbf{z})}{q_\phi(\mathbf{z} | \mathbf{x})} \quad (31)$$

Given our choice of geometric path  $\pi_\beta(\mathbf{z}) = \tilde{\pi}_\beta(\mathbf{z})/Z_\beta$ ,  $\tilde{\pi}_\beta(\mathbf{z}) = p(\mathbf{x}, \mathbf{z})^\beta q(\mathbf{z} | \mathbf{x})^{1-\beta}$ , the potential  $U'(\mathbf{z}) = \frac{\partial}{\partial \beta} \log \tilde{\pi}_\beta(\mathbf{z})$  it loses its dependency on  $\beta$  after differentiating. This allows us to show

$$\frac{\partial}{\partial \beta} g(\beta) = \text{Var}_{\pi_\beta(\mathbf{z})} [U'(\mathbf{z})] \quad (32)$$

which entails  $\frac{\partial}{\partial \beta} g(\beta) \geq 0, \forall \beta \in (0, 1)$  and therefore that  $g(\beta)$  is monotonically non-decreasing. Changes between lines are tracked in [blue](#).

*Proof of Equation (32).*

$$\begin{aligned} \frac{\partial}{\partial \beta} g(\beta) &= \frac{\partial}{\partial \beta} \mathbb{E}_{\pi_\beta(\mathbf{z})} [U'(\mathbf{z})] \\ &= \frac{\partial}{\partial \beta} \left[ \int \pi_\beta(\mathbf{z}) U'(\mathbf{z}) d\mathbf{z} \right] \\ &= \int U'(\mathbf{z}) \frac{\partial}{\partial \beta} \pi_\beta(\mathbf{z}) d\mathbf{z} \\ &= \int U'(\mathbf{z}) \frac{\partial}{\partial \beta} [Z_\beta^{-1} * \tilde{\pi}_\beta(\mathbf{z})] d\mathbf{z} \\ &= \int U'(\mathbf{z}) \left[ \tilde{\pi}_\beta(\mathbf{z}) * \frac{\partial}{\partial \beta} Z_\beta^{-1} + Z_\beta^{-1} * \frac{\partial}{\partial \beta} \tilde{\pi}_\beta(\mathbf{z}) \right] d\mathbf{z}. \end{aligned}$$

Now we use the ‘‘inverse log-derivative’’ trick  $\frac{\partial}{\partial x} (f(x)^{-1}) = -\frac{1}{f(x)} \frac{\partial}{\partial x} \log f(x)$  on the first term, and the log-derivative trick on the second

$$= \int U'(\mathbf{z}) \left[ \tilde{\pi}_\beta(\mathbf{z}) * \frac{-1}{Z_\beta} \frac{\partial}{\partial \beta} \log Z_\beta + \frac{1}{Z_\beta} * \tilde{\pi}_\beta(\mathbf{z}) \frac{\partial}{\partial \beta} \log \tilde{\pi}_\beta(\mathbf{z}) \right] d\mathbf{z} \quad (33)$$

$$= \int U'(\mathbf{z}) \left[ -\pi_\beta(\mathbf{z}) \frac{\partial}{\partial \beta} \log Z_\beta + \pi_\beta(\mathbf{z}) \frac{\partial}{\partial \beta} \log \tilde{\pi}_\beta(\mathbf{z}) \right] d\mathbf{z}, \quad (34)$$

Then we use (26) on the first term, and the definition of  $U'(\mathbf{z})$  in the second

$$= \int U'(\mathbf{z}) \left[ -\pi_\beta(\mathbf{z}) \mathbb{E}_{\pi_\beta(\mathbf{z})} [U'(\mathbf{z})] + \pi_\beta(\mathbf{z}) U'(\mathbf{z}) \right] d\mathbf{z} \quad (35)$$

$$= - \int \pi_\beta(\mathbf{z}) U'(\mathbf{z}) \mathbb{E}_{\pi_\beta(\mathbf{z})} [U'(\mathbf{z})] d\mathbf{z} + \int U'(\mathbf{z}) U'(\mathbf{z}) \pi_\beta(\mathbf{z}) d\mathbf{z} \quad (36)$$

Finally we rearrange, noting that the expectation is a scalar and can therefore come out of the integrand

$$= - \left[ \mathbb{E}_{\pi_\beta(\mathbf{z})} [U'(\mathbf{z})] \right] \left[ \int \pi_\beta(\mathbf{z}) U'(\mathbf{z}) d\mathbf{z} \right] + \int U'(\mathbf{z}) U'(\mathbf{z}) \pi_\beta(\mathbf{z}) d\mathbf{z} \quad (37)$$

$$= - \left[ \mathbb{E}_{\pi_\beta(\mathbf{z})} [U'(\mathbf{z})] \right]^2 + \mathbb{E}_{\pi_\beta(\mathbf{z})} [U'(\mathbf{z})^2] \quad (38)$$

$$= \text{Var}_{\pi_\beta(\mathbf{z})} [U'(\mathbf{z})]. \quad (39)$$

Therefore,

$$\frac{\partial}{\partial \beta} g(\beta) = \text{Var}_{\pi_\beta(\mathbf{z})} [U'(\mathbf{z})]. \quad (40)$$

□

## D Maximizing the TVO minimizes a divergence between the variational distribution and true posterior

We now show:

$$\text{TVO}(\theta, \phi, \mathbf{x}) = \log p_\theta(\mathbf{x}) - \mathcal{D}(q_\phi(\mathbf{z} | \mathbf{x}) || p_\theta(\mathbf{z} | \mathbf{x})) \quad (41)$$

Where  $\mathcal{D}(q_\phi(\mathbf{z} | \mathbf{x}) || p_\theta(\mathbf{z} | \mathbf{x}))$  is a divergence between the variational distribution  $q_\phi(\mathbf{z} | \mathbf{x})$  and true posterior  $p_\theta(\mathbf{z} | \mathbf{x})$ . We refer to the notation defined in Appendix C.1 and the definition of divergence defined by Eguchi et al. (1985).

*Proof.* The TVO is a left Riemann sum approximation of  $\log p_\theta(\mathbf{x}) = \int_0^1 g(\beta) d\beta$ , where  $g(\beta) = \mathbb{E}_{\pi_\beta(\mathbf{z})} [U'(\mathbf{z})]$  and  $g(\beta)$  is a differentiable monotonically non-decreasing function in  $\beta$  (cf. Equation (32)). The TVO is therefore a lower bound of  $\log p_\theta(\mathbf{x})$  and can be written

$$\begin{aligned} \text{TVO}(\theta, \phi, \mathbf{x}) &\leq \log p_\theta(\mathbf{x}) \\ \text{TVO}(\theta, \phi, \mathbf{x}) &= \log p_\theta(\mathbf{x}) - c(\theta, \phi, \mathbf{x}), \quad c(\theta, \phi, \mathbf{x}) \geq 0 \end{aligned} \quad (42)$$

We will show  $c(\theta, \phi, \mathbf{x}) = \mathcal{D}(q_\phi(\mathbf{z} | \mathbf{x}) || p_\theta(\mathbf{z} | \mathbf{x}))$ , which is equivalent to showing

- ①  $c \geq 0, \forall p_\theta(\mathbf{z} | \mathbf{x}), q_\phi(\mathbf{z} | \mathbf{x}) \in \mathcal{Z}$
- ②  $c = 0 \iff p_\theta(\mathbf{z} | \mathbf{x}) = q_\phi(\mathbf{z} | \mathbf{x})$

① is implied in the definition of  $c$  in 42. We now show ②.

**Forward direction** ( $c = 0$ )  $\Rightarrow$  ( $p_\theta(\mathbf{z} | \mathbf{x}) = q_\phi(\mathbf{z} | \mathbf{x})$ )

If  $c = 0$ , the left Riemann sum must be an exact approximation to  $\int_0^1 g(\beta) d\beta$ . Because  $g(\beta)$  is differentiable (and assuming it is finite), the Riemann approximation can only be exact when  $g(\beta)$  is flat (i.e.  $\frac{\partial g(\beta)}{\partial \beta} = 0$ ) in the region  $\beta \in [0, 1]$ . We first recall that by definition,  $\pi_0(\mathbf{z}) = q_\phi(\mathbf{z} | \mathbf{x})$  and  $\pi_1(\mathbf{z}) = p_\theta(\mathbf{z} | \mathbf{x})$ . Therefore

$$\int_0^1 \frac{\partial g(\beta)}{\partial \beta} d\beta = \int_0^1 0 d\beta \quad (43)$$

$$g(1) - g(0) = 0 \quad (44)$$

$$g(1) = g(0) \quad (45)$$

$$\mathbb{E}_{\pi_1(\mathbf{z})} [U'(\mathbf{z})] = \mathbb{E}_{\pi_0(\mathbf{z})} [U'(\mathbf{z})] \quad (46)$$

$$\mathbb{E}_{p_\theta(\mathbf{z} | \mathbf{x})} [U'(\mathbf{z})] = \mathbb{E}_{q_\phi(\mathbf{z} | \mathbf{x})} [U'(\mathbf{z})] \quad (47)$$

Which is only possible when  $p_\theta(\mathbf{z} | \mathbf{x}) = q_\phi(\mathbf{z} | \mathbf{x})$ .

**Reverse direction**  $(p_\theta(\mathbf{z} | \mathbf{x}) = q_\phi(\mathbf{z} | \mathbf{x})) \Rightarrow (c = 0)$

If  $p_\theta(\mathbf{z} | \mathbf{x}) = q_\phi(\mathbf{z} | \mathbf{x})$ , the TVO can be written as

$$\text{TVO}(\theta, \phi, \mathbf{x}) = \frac{1}{K} \sum_{k=0}^{K-1} \mathbb{E}_{\pi_{\beta_k}(\mathbf{z})} [\log \frac{p_\theta(\mathbf{x}, \mathbf{z})}{p_\theta(\mathbf{z} | \mathbf{x})}] \quad (48)$$

$$= \frac{1}{K} \sum_{k=0}^{K-1} \mathbb{E}_{\pi_{\beta_k}(\mathbf{z})} [\log p_\theta(\mathbf{x})] \quad (49)$$

$$= \log p_\theta(\mathbf{x}) \quad (50)$$

Therefore  $c = 0$ . □

## E Derivation of Gradient Estimator

We want to show that

$$\nabla_\lambda \mathbb{E}_{\pi_{\lambda, \beta}} [f(\mathbf{z}, \lambda)] = \mathbb{E}_{\pi_{\lambda, \beta}} [\nabla_\lambda f(\mathbf{z}, \lambda)] + \text{Cov}_{\pi_{\lambda, \beta}} [\nabla_\lambda \log \tilde{\pi}_{\lambda, \beta}(\mathbf{z}), f(\mathbf{z}, \lambda)]. \quad (51)$$

Our estimator holds under the common regularity conditions assumed for the score function estimator (L'Ecuyer, 1993). We begin with a simple lemma.

**Lemma 1.**

$$\nabla_\lambda \log Z_{\lambda, \beta}(\mathbf{x}) = \mathbb{E}_{\pi_{\beta}(\mathbf{z})} [\nabla_\lambda \log \tilde{\pi}_{\lambda, \beta}(\mathbf{z})] \quad (52)$$

*Proof of lemma 1.*

$$\nabla_\lambda \log Z_{\lambda, \beta}(\mathbf{x}) = \frac{1}{Z_{\lambda, \beta}(\mathbf{x})} \nabla_\lambda Z_{\lambda, \beta}(\mathbf{x}) \quad (53)$$

$$= \frac{1}{Z_{\lambda, \beta}(\mathbf{x})} \nabla_\lambda \int \tilde{\pi}_{\lambda, \beta}(\mathbf{z}) d\mathbf{z} \quad (54)$$

$$= \frac{1}{Z_{\lambda, \beta}(\mathbf{x})} \int \nabla_\lambda \tilde{\pi}_{\lambda, \beta}(\mathbf{z}) d\mathbf{z} \quad (55)$$

$$= \int \frac{\tilde{\pi}_{\lambda, \beta}(\mathbf{z})}{Z_{\lambda, \beta}(\mathbf{x})} \nabla_\lambda \log \tilde{\pi}_{\lambda, \beta}(\mathbf{z}) d\mathbf{z} \quad (56)$$

$$= \mathbb{E}_{\pi_{\beta}(\mathbf{z})} [\nabla_\lambda \log \tilde{\pi}_{\lambda, \beta}(\mathbf{z})] \quad (57)$$

□

To prove (51), we use the product rule and rearrange

$$\nabla_\lambda \mathbb{E}_{\pi_{\beta}(\mathbf{z})} [f(\mathbf{z}, \lambda)] = \mathbb{E}_{\pi_{\beta}(\mathbf{z})} [\nabla_\lambda f(\mathbf{z}, \lambda) + f(\mathbf{z}, \lambda) \nabla_\lambda \log \pi_{\lambda, \beta}(\mathbf{z} | \mathbf{x})] \quad (58)$$

$$= \mathbb{E}_{\pi_{\beta}(\mathbf{z})} [\nabla_\lambda f(\mathbf{z}, \lambda) + f(\mathbf{z}, \lambda) (\nabla_\lambda \log \tilde{\pi}_{\lambda, \beta}(\mathbf{z}) - \nabla_\lambda \log Z_{\lambda, \beta}(\mathbf{x}))] \quad (59)$$

$$= \mathbb{E}_{\pi_{\beta}(\mathbf{z})} [\nabla_\lambda f(\mathbf{z}, \lambda)] + \mathbb{E}_{\pi_{\beta}(\mathbf{z})} [f(\mathbf{z}, \lambda) \nabla_\lambda \log \tilde{\pi}_{\lambda, \beta}(\mathbf{z})] - \mathbb{E}_{\pi_{\beta}(\mathbf{z})} [f(\mathbf{z}, \lambda) \nabla_\lambda \log Z_{\lambda, \beta}(\mathbf{x})]. \quad (60)$$

Now using lemma 1 on the third term

$$= \mathbb{E}_{\pi_{\beta}(\mathbf{z})} [\nabla_\lambda f(\mathbf{z}, \lambda)] + \mathbb{E}_{\pi_{\beta}(\mathbf{z})} [f(\mathbf{z}, \lambda) \nabla_\lambda \log \tilde{\pi}_{\lambda, \beta}(\mathbf{z})] - \mathbb{E}_{\pi_{\beta}(\mathbf{z})} [f(\mathbf{z}, \lambda)] \mathbb{E}_{\pi_{\beta}(\mathbf{z})} [\nabla_\lambda \log \tilde{\pi}_{\lambda, \beta}(\mathbf{z})] \quad (61)$$

$$= \mathbb{E}_{\pi_{\beta}(\mathbf{z})} [\nabla_\lambda f(\mathbf{z}, \lambda)] + \text{Cov}_{\pi_{\lambda, \beta}(\mathbf{z} | \mathbf{x})} [\nabla_\lambda \log \tilde{\pi}_{\lambda, \beta}(\mathbf{z}), f(\mathbf{z}, \lambda)]. \quad (62)$$

## F Special Cases of the TVO

In Table 1, we summarize different ways the TVO generalizes existing variational objectives. In the main text, we have mentioned that the lower bound variant of the TVO with  $K = 1$  partition can be seen as the ELBO. For  $K > 1$ , we have a novel objective which we can optimize with respect to  $\theta$ ,  $\phi$  or both  $\theta, \phi$ . In the main text, we investigate in detail the latter.

As also mentioned in the main text, the upper bound variant of the TVO with  $K = 1$  partition can be seen as EUBO and hence connects TVO to WS, RWS and inference compilation. For  $K > 1$ , we have a novel objective which we can optimize with respect to  $\phi$ .

Table 1: The thermodynamic variational identity generalizes existing variational objectives.

Approximation	Left Riemann sum (lower bound—maximize)		Right Riemann sum (upper bound—minimize)	
	1	$> 1$	1	$> 1$
Number of partitions	1	$> 1$	1	$> 1$
$\theta$	wake in WS	THERMO- $\theta$	N/A	N/A
Optimize	$\phi$	VI	THERMO-VI	wake- $\phi$ in RWS, sleep in WS, inference compilation
	$\theta, \phi$	VAE	THERMO	N/A

## G Additional Illustrations of the Thermodynamic Variational Identity

In Figure 6, we provide illustrations of how the  $\mathbb{E}_{\pi_\beta}[U'(\mathbf{z})]$  curve relates to  $\log p_\theta(\mathbf{x})$ ,  $\text{KL}(q||p)$ ,  $\text{KL}(p||q)$ , ELBO and EUBO for the cases of  $\text{ELBO} < 0 < \text{EUBO}$  and  $\text{ELBO} < \text{EUBO} < 0$ . In the following, we provide derivations to justify the illustrations.

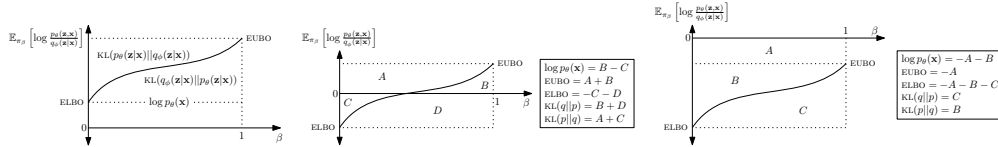


Figure 6: Different scenarios of the  $\mathbb{E}_{\pi_\beta}[U'(\mathbf{z})]$  curve where  $\text{ELBO} < 0$ . On the left,  $0 < \text{ELBO} < \text{EUBO}$ . In the middle,  $\text{ELBO} < 0 < \text{EUBO}$ . On the right  $\text{ELBO} < \text{EUBO} < 0$ .

**Case  $\text{ELBO} < 0 < \text{EUBO}$ .** The top-most point of the curve is the EUBO by definition which means that the area  $A + B$  is equal to the EUBO because of the unit length of the rectangle. In a similar manner, the ELBO is the negative of the area of  $C + D$ . Now, due to the thermodynamic identity,  $\log p_\theta(\mathbf{x}) = \int_{\beta=0}^1 \mathbb{E}_{\pi_\beta}[U'(\mathbf{z})] d\beta$ , it is equal to  $B - C$  which is the area denoted by the definite integral. To obtain the expressions for the KL, we use the identities  $\log p_\theta(\mathbf{x}) = \text{ELBO}(\mathbf{x}, \theta, \phi) + \text{KL}(q_\phi(\mathbf{z}|\mathbf{x})||p_\theta(\mathbf{z}|\mathbf{x})) = \text{EUBO}(\mathbf{x}, \theta, \phi) = \log p_\theta(\mathbf{x}) - \text{KL}(p_\theta(\mathbf{z}|\mathbf{x})||q_\phi(\mathbf{z}|\mathbf{x}))$ .

**Case  $\text{ELBO} < \text{EUBO} < 0$ .** The top-most point of the curve is the EUBO by definition which means that  $-A$  is equal to the EUBO because of the unit length of the rectangle. In a similar manner, the ELBO is  $-A - B - C$ . Due to the thermodynamic identity,  $\log p_\theta(\mathbf{x}) = \int_{\beta=0}^1 \mathbb{E}_{\pi_\beta}[U'(\mathbf{z})] d\beta$ , it is equal to  $-A - B$  which is the area denoted by the definite integral. We obtain expressions for the KL similarly as before.

Similar line of reasoning gives rise to the relationships in Figure 6 (left).

## H Details for Deep Generative Models

**Discrete latent variables.** The discrete generative model and inference network is as follows:

$$\begin{aligned}
 p_\theta(\mathbf{z}_L) &= \text{Bernoulli}(\mathbf{z}_L | \mathbf{b}_L), \\
 p_\theta(\mathbf{z}_\ell | \mathbf{z}_{\ell+1}) &= \text{Bernoulli}(\mathbf{z}_\ell | \text{decoder}_\ell(2\mathbf{z}_{\ell+1} - 1)) & \ell = L - 1, \dots, 1, \\
 p_\theta(\mathbf{x} | \mathbf{z}_1) &= \text{Bernoulli}(\mathbf{x} | \text{decoder}_x(2\mathbf{z}_1 - 1) + \tilde{\mathbf{x}}), \\
 q_\phi(\mathbf{z}_1 | \mathbf{x}) &= \text{Bernoulli}\left(\mathbf{z}_1 \mid \text{encoder}_1\left(\frac{\mathbf{x} - \tilde{\mathbf{x}} + 1}{2}\right)\right), \\
 q_\phi(\mathbf{z}_\ell | \mathbf{z}_{\ell-1}) &= \text{Bernoulli}(\mathbf{z}_\ell | \text{encoder}_\ell(2\mathbf{z}_{\ell-1} - 1)) & \ell = 2, \dots, L,
 \end{aligned}$$

where  $\mathbf{x} \in \{0, 1\}^{D_x}$  and  $\mathbf{z}_\ell \in \{0, 1\}^{D_z}$  for  $D_x = 784$  and  $D_z = 200$ . We used Pytorch’s default parameter initialization. The Bernoulli distributions are independent Bernoulli distributions whose parameters are logits, i.e. they get passed through a sigmoid function to obtain the probability.  $\tilde{\mathbf{x}}$  is the mean over training data set and  $\tilde{\mathbf{x}} = \log(\tilde{\mathbf{x}} - 1)$ . In the linear case, the encoders and decoders are linear functions of their inputs. In the non-linear case, they are a three-layer multilayer perceptrons with tanh nonlinearities of the form  $\text{input\_dim} \xrightarrow{\text{Lin+tanh}} D_z \xrightarrow{\text{Lin+tanh}} D_z \xrightarrow{\text{Lin}} \text{output\_dim}$ .

**Continuous latent variables.** The continuous generative model and inference network is as follows:

$$\begin{aligned}
 p(\mathbf{z}) &= \text{Normal}(\mathbf{z} | 0, I), \\
 p_\theta(\mathbf{x} | \mathbf{z}) &= \text{Bernoulli}(\mathbf{x} | \text{decoder}_\theta(\mathbf{z})), \\
 q_\phi(\mathbf{z} | \mathbf{x}) &= \text{Normal}(\mathbf{z} | \text{encoder}_\phi(\mathbf{x})),
 \end{aligned}$$

where  $\mathbf{x} \in \{0, 1\}^{D_x}$  and  $\mathbf{z} \in \mathbb{R}^{D_z}$  for  $D_x = 784$  and  $D_z = 200$ . The decoder is of the form  $D_z \xrightarrow{\text{Lin+tanh}} D_z \xrightarrow{\text{Lin+tanh}} D_z \xrightarrow{\text{Lin}} D_x$  and its output is passed through a sigmoid to obtain probabilities for the Bernoulli distribution. The encoder is of the form  $D_x \xrightarrow{\text{Lin+tanh}} D_z \xrightarrow{\text{Lin+tanh}} D_z$ . Its output is passed through two *separate* neural networks of the form  $D_z \xrightarrow{\text{Lin}} D_z$  which output the means and log standard deviations of the independent Normal distribution.

## I Notation



Table 2: Table of Notation

---

$\{\mathbf{x}_i\}_{i=1}^N$	:=	Data set consisting of N i.i.d samples $\mathbf{x}_i \in \mathcal{R}^D$
$\{\mathbf{z}_i\}_{i=1}^N$	:=	Unobserved latent random variables $\mathbf{z}_i \in \mathcal{R}^M$
$p_\theta(\mathbf{x}, \mathbf{z}) = p_\theta(\mathbf{x}   \mathbf{z})p_\theta(\mathbf{z})$	:=	The joint model parameterized by $\theta$ , which factorizes into a likelihood $p_\theta(\mathbf{x}   \mathbf{z})$ and prior $p_\theta(\mathbf{z})$
$p_\theta(\mathbf{z}   \mathbf{x}) = p_\theta(\mathbf{x}, \mathbf{z})/p_\theta(\mathbf{x})$	:=	The true (often intractable) posterior
$q_\phi(\mathbf{z}   \mathbf{x})$	:=	The variational distribution parameterized by $\phi$ . By assumption $q_\phi(\mathbf{z}   \mathbf{x})$ is correctly normalized.
$\tilde{\pi}_{\lambda, \beta}(\mathbf{z}) = p_\theta(\mathbf{x}, \mathbf{z})^\beta q_\phi(\mathbf{z}   \mathbf{x})^{1-\beta}$	:=	The unnormalized path distributions. By construction, $\tilde{\pi}_{\lambda, \beta=1}(\mathbf{z}) = p_\theta(\mathbf{x}, \mathbf{z})$ and $\tilde{\pi}_{\lambda, \beta=0}(\mathbf{z}   \mathbf{x}) = q_\phi(\mathbf{z}   \mathbf{x})$
$\pi_{\lambda, \beta}(\mathbf{z}   \mathbf{x}) = \tilde{\pi}_{\lambda, \beta}(\mathbf{z})/Z_{\lambda, \beta}(\mathbf{x})$	:=	The path distributions parameterized by $\lambda = \{ \theta, \phi \}$ and scalar parameter $\beta \in [0, 1]$ . By construction, $\pi_{\lambda, \beta=1}(\mathbf{z}   \mathbf{x}) = p_\theta(\mathbf{z}   \mathbf{x})$ and $\pi_{\lambda, \beta=0}(\mathbf{z}   \mathbf{x}) = q_\phi(\mathbf{z}   \mathbf{x})$
$Z_{\lambda, \beta}(\mathbf{x}) = \int \tilde{\pi}_{\lambda, \beta}(\mathbf{z}) d\mathbf{z}_{1:N}$	:=	The normalizing constant for the path distributions. By construction $Z_{\lambda, \beta=1}(\mathbf{x}) = p_\theta(\mathbf{x})$ and $Z_{\lambda, \beta=0}(\mathbf{x}) = 1$ (because $q_\phi(\mathbf{z}   \mathbf{x})$ is assumed to be correctly normalized).
$U_{\lambda, \beta}(\mathbf{z}) = \log \tilde{\pi}_{\lambda, \beta}(\mathbf{z})$	:=	The potential energy.
$U'_{\lambda, \beta}(\mathbf{z}) = \frac{\partial}{\partial \beta} U_{\lambda, \beta}(\mathbf{z})$	:=	The first derivative of the potential w.r.t $\beta$ , the inverse temperature.

---