

# Detecting and mitigating mode-collapse for flow-based sampling of lattice field theories

Kim A. Nicoli<sup>1,2,3,\*</sup>, Christopher J. Anders<sup>3,4</sup>, Tobias Hartung<sup>5</sup>, Karl Jansen,<sup>6</sup>  
Pan Kessel,<sup>7,†</sup> and Shinichi Nakajima<sup>3,4,8</sup>

<sup>1</sup>Transdisciplinary Research Area (TRA) Matter, University of Bonn, 53115 Bonn, Germany

<sup>2</sup>Helmholtz Institute for Radiation and Nuclear Physics (HISKP), 53115 Bonn, Germany

<sup>3</sup>Berlin Institute for the Foundations of Learning and Data (BIFOLD), 10587 Berlin, Germany

<sup>4</sup>Machine Learning Group, Technische Universität Berlin, 10587 Berlin, Germany

<sup>5</sup>Northeastern University—London, London E1W 1LP, United Kingdom

<sup>6</sup>CQTA, Deutsches Elektronen-Synchrotron DESY, 15738 Zeuthen, Germany

<sup>7</sup>Prescient Design, gRED, Roche, 2807 Basel, Switzerland

<sup>8</sup>RIKEN Center for AIP, 103-0027 Tokyo, Japan



(Received 27 February 2023; revised 24 July 2023; accepted 2 November 2023; published 8 December 2023)

We study the consequences of mode-collapse of normalizing flows in the context of lattice field theory. Normalizing flows allow for independent sampling. For this reason, it is hoped that they can avoid the tunneling problem of local-update Markov Chain Monte Carlo algorithms for multimodal distributions. In this work, we first point out that the tunneling problem is also present for normalizing flows but is shifted from the sampling to the algorithm's training phase. Specifically, normalizing flows often suffer from mode-collapse for which the training process assigns vanishingly low probability mass to relevant modes of the physical distribution. This may result in a significant bias when the flow is used as a sampler in a Markov-Chain or with importance sampling. We propose a metric to quantify the degree of mode-collapse and derive a bound on the resulting bias. Furthermore, we propose various mitigation strategies in particular in the context of estimating thermodynamic observables, such as the free energy.

DOI: [10.1103/PhysRevD.108.114501](https://doi.org/10.1103/PhysRevD.108.114501)

## I. INTRODUCTION

Using normalizing flows for sampling in lattice field theory has gained significant attention over the last few years. Several works have been carried out in the domain of scalar field theories [1–10],  $U(1)$  [11,12] and  $SU(N)$  [13–15] pure gauge theories, and fermionic gauge theories [16,17]. This rapid development is attributed to the appealing conceptual properties of flow-based sampling. A well-trained flow approximately acts as a trivializing map [18] and therefore can significantly reduce the integrated auto-correlation time of physical observables. The practical obstruction to harnessing this conceptual advantage is that the training process becomes increasingly challenging as the dimensionality of the lattice increases, resulting in poor volume scaling [19–22]. Furthermore, it is well known that generative models struggle to learn long-range

correlations [23] which is crucial as a critical point is approached. When the continuum limit of the theory is taken, both challenges manifest simultaneously: the volume needs to be increased as the critical point is approached. As a result, it remains an open question whether useful architectures can be found for addressing critical slowing down in the continuum limit.

Another conceptually appealing property of normalizing flows is that they allow for independent sampling, thus making flow densities suitable for being combined with Metropolis-Hastings accept-reject schemes. This approach is often referred to as Neural-Markov Chain Monte Carlo (MCMC) [1,24–26]. As a result, it may be hoped that they can avoid the tunneling problem which arises when local update MCMC algorithms are applied to theories that have degenerate minima separated by high action barriers. However, normalizing flows are typically trained by self-sampling in the context of lattice field theory [1]. As we will discuss, this bears the risk that the training will assign vanishing low probability mass to some of the modes of the theory [2–4], since the training objective will not strongly penalize this. If mode-collapse happens, certain modes of the theory will not be probed by the sampler. This problem, therefore, leads to substantially biased estimators of physical observables as shown in Fig. 1.

\*kim.a.nicoli@gmail.com

†pan.kessel@gmail.com

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP<sup>3</sup>.

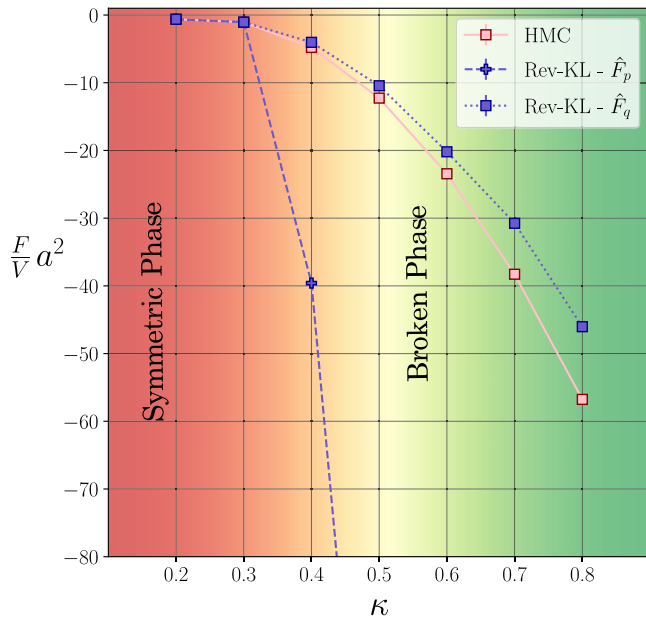


FIG. 1. Estimation of free energy density in broken and symmetric phases using a reverse-KL trained flow-estimation of the free energy density using the approach proposed in [2] and [4]. The second-order phase transition is represented with a color gradient from red to green in the background. The free energy density is estimated using samples drawn from the flow and the target distribution respectively. The flow estimates (purple) are compared to the HMC baseline (pink). The experiments consider lattices  $\Lambda = 64 \times 8$  at fixed coupling  $\lambda = 0.022$  for the  $\phi^4$ -theory as in [2]. In the broken phase (green) of the theory, e.g.,  $\kappa \geq 0.3$ , the target density has two modes with very little tunneling probability between the two. Using a deep generative sampler, i.e., a flow, trained with reverse-KL in the broken regime leads to biased estimates. This problem lays the foundation of the study presented in this work.

In our work we study mode-collapse and the more general mode-mismatch phenomenon, both theoretically and numerically. We first discuss in detail the mode-seeking nature of the standard self-sampling-based training procedure which corresponds to minimizing the reverse Kullback-Leibler (KL) divergence [3]. We compare this to an alternative training procedure which is based on minimizing the forward (as opposed to the reverse) KL divergence and review why it is equivalent to maximum likelihood training. This objective has the advantage that it is substantially less vulnerable to mode-collapse but has the disadvantage that it requires representative configurations sampled from the theory. In many applications, this prevents this objective from being of any use since if such configurations are available, we can directly measure physical observables on them and a flow is not necessary. However, we point out that there is an important exception to this: for thermodynamic observables, such as the free energy, it is still useful to train a flow. This is because these observables are typically obtained by integration through

the parameter space of the theory and thus require a significant number of Markov chains along a discretized trajectory in the parameter space. By training a flow on samples generated at a single point in parameter space, we can completely avoid the need for these additional Markov chains. In this important scenario, it is thus sensible and, as we argue, advisable to use forward KL training for the flow to significantly reduce mode-collapse. Besides modifying the training procedure, we also propose to mitigate mode-collapse by combining two flow-based estimators for the free energy. As a side remark, we note that concurrent works have been proposing strategies, alternative to the forward KL objective, trying to mitigate mode collapse. These include more stable path gradient estimators [27,28], learning deformed target distributions [29] and annealed importance sampling [30].

We then study the bias induced by mode-collapse theoretically. Specifically, we derive a bound on the bias of the estimator for physical observables. This allows us to propose a natural metric to quantify the degree of mode-collapse of the sampler.

The effectiveness of our proposed methods is then demonstrated on a two-dimensional  $\phi^4$  scalar theory.

We stress that our study focuses on the estimation of the free energy, an example of thermodynamic quantity involving the partition function [31], a crucial subset of physical observables in lattice field theory [32]. Estimating these observables with standard Markov-Chain-based methods requires sampling configurations at many different values in parameter space and integrating free energy differences from a known reference value, see previous works for more details [2,4]. This approach is computationally expensive since it often requires a significant number of HMC chains along the trajectory in parameter space, and crucially leads to high uncertainty, as errors from each chain accumulate upon integration. This problem becomes more severe when one needs to cross a phase transition. There, integrated autocorrelation times explode, thus resulting in larger errors for each Markov chain. For this reason, training a normalizing flow using a forward KL objective can often be advantageous: training a normalizing flow requires samples from only a single Markov chain at the target point in parameter space and thus allows us to circumvent the need for any additional chains along the trajectory through parameter space.

We emphasize that the intricacies of training a normalizing flow for multimodal distributions in the context of lattice field theories have been already discussed in [3]. Our work builds on this reference but is different in the sense that we consider thermodynamic observables. As explained above, these observables cannot be estimated on the Markov-Chain samples at the target point without the need for additional Markov chains for different coupling values. As a result, training of normalizing flows using the forward KL objective is particularly natural for the estimation of thermodynamic observables.

## II. TRAINING A GENERATIVE MODEL

Normalizing flows [33–35] are a particular class of generative models giving access to an analytic form of the likelihood. While this work focuses on flows for concreteness, we stress that the theoretical arguments made in the following sections hold for any generative model allowing for exact likelihood estimation. For such models, a variational density  $q_\theta$ , the *sampler*, parameterized by a set of weights  $\theta$ , is optimized to approximate the target density of the lattice field theory

$$p(\phi) = \frac{1}{Z} \exp(-S(\phi)), \quad (1)$$

where  $Z$  is the partition function and  $S$  is the action of the theory.

During the training of a normalizing flow, an efficient transformation to map a base density  $q_z$  into a nontrivial target is learned. In practice, the base distribution is chosen such that it allows for efficient sampling. Common choices for the base density are therefore normal or uniform distributions.

The flow uses a diffeomorphism  $f$  between the base space  $\mathcal{Z}$  and the configuration space  $\mathcal{X}$  hence

$$f_\theta: z \in \mathcal{Z} \sim q_z \rightarrow x = f_\theta(z) \in \mathcal{X} \sim q_\theta. \quad (2)$$

The diffeomorphism  $f_\theta$  is a composition of bijective transformations  $f_\theta^i$  referred to as *coupling blocks*. Each of these blocks satisfies the following requirements:

- (1)  $f_\theta^i$  is a bijection,
- (2) both  $f_\theta^i$  and its inverse are in  $\mathcal{C}^\infty$ .
- (3) the determinant of the Jacobian is efficient to evaluate.

The inverse of the transformation  $f_\theta^{-1}(x) = z$  therefore always exists by construction. Leveraging these properties, an analytic expression for the likelihood of the flow-based model reads

$$q_\theta(x) = q_z(f_\theta^{-1}(x)) \left| \frac{df_\theta}{dz} \right|^{-1} = q_z(z) \left| \frac{df_\theta}{dz} \right|^{-1}. \quad (3)$$

Different coupling blocks satisfying the requirements above have been proposed; these include nonlinear independent component estimation (NICE) [36], real nonvolume preserving (RealNVP) [37], and generative flow (GLOW) [38]. We refer to [34,35] for an overview of the existing coupling blocks and further technical details.

### A. The forward- and reverse-KL divergences

During training, the normalizing flow is optimized by density matching. It is common practice to minimize the so-called KL divergences to this end although other types of generalized divergences can be used [39–46]. As we will discuss in this section, and in Sec. III, choosing an

appropriate divergence is crucial to ensure successful training.

The so-called reverse-KL divergence reads

$$\text{KL}(q_\theta||p) = \int \mathcal{D}[\phi] q_\theta(\phi) \ln \frac{q_\theta(\phi)}{p(\phi)} \quad (4)$$

where  $\mathcal{D}[\phi]$  represents the measure of a high-dimensional integral. It is worth stressing that the KL divergence is not symmetric hence

$$\text{KL}(q_\theta||p) \neq \text{KL}(p||q_\theta). \quad (5)$$

The right-hand side of Eq. (5) is usually referred to as the *forward-KL* which can be written as an expectation value with respect to the target density  $p$

$$\text{KL}(p||q_\theta) = \int \mathcal{D}[\phi] p(\phi) \ln \frac{p(\phi)}{q_\theta(\phi)}. \quad (6)$$

These two choices for the divergence lead to different training procedures, as we will discuss in sections Secs. II A 1 and II A 2. We also note that the use of reverse and forward KL is not mutually exclusive. In the context of quantum chemistry [47], for instance, a combination of the two is typically chosen.

#### 1. Reverse-KL: Training by self-sampling

The reverse KL divergence is the standard choice for training normalizing flows on the lattice. This is because lattice field theory comes with an action  $S$  which is known in closed form—in contrast to many other machine learning applications.

The reverse KL divergence (4) can be approximated by a Monte-Carlo estimate [2,24] as follows

$$\begin{aligned} \text{KL}(q_\theta||p) &= \mathbb{E}_{q_\theta} \left[ \ln \frac{q_\theta(\phi)}{p(\phi)} \right] \\ &\approx \frac{1}{N} \sum_{i=1}^N (S(\phi_i) + \ln q_\theta(\phi_i)) + \text{const}. \end{aligned} \quad (7)$$

Here, the field configurations are sampled from the flow, i.e.,  $\phi_i \sim q_\theta$ , and thus the training relies on self-sampling. In particular, the partition function  $Z$  contributes by a shift term that is constant with respect to the parameters of the flow and thus can be ignored for optimization by gradient descent.

Training using a reverse-KL is therefore very efficient because it does not require samples from the target density  $p$  due to self-sampling. Unfortunately, this comes at a cost as this objective is known to be prone to mode-collapse. The sketch on the left-hand side of Fig. 2 shows that a reverse-KL-trained flow tends to focus its support on a subset of the modes when the target density is multimodal.

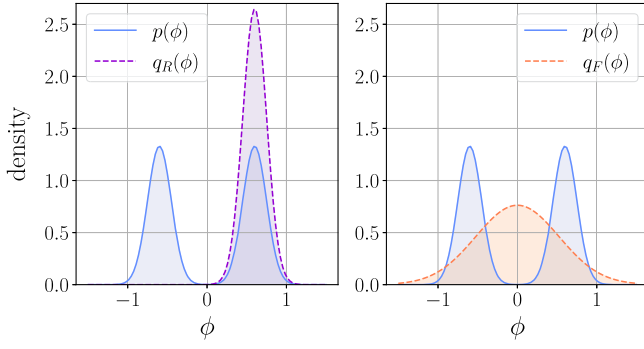


FIG. 2. Comparison between a reverse-KL (purple) and forward-KL (orange) optimization approach—a sketch of normalizing flows trained with both KL objectives as described in Sec. II A. Training with a reverse KL shows a mode-seeking behavior and is thus prone to mode-dropping. The forward-KL has instead a mode-covering behavior which leads to larger support over the sampling space.

This undesirable behavior strongly affects the reverse KL as the dropped modes are not probed in the self-sampling process.

This mode-seeking nature of the reverse KL represents a major drawback and limits the applicability of this framework when the physical density to be learned has more than one mode [3,4,29].

## 2. Forward-KL: Training by maximum likelihood

The forward KL divergences can be written as an expectation value with respect to the target density  $p$  and thus also be approximated by Monte-Carlo

$$\begin{aligned} \text{KL}(p||q_\theta) &= \mathbb{E}_p \left[ \ln \frac{p(\phi)}{q_\theta(\phi)} \right] \\ &= -\frac{1}{N} \sum_{i=1}^N \log q_\theta(\phi_i) + \text{const.} \end{aligned} \quad (8)$$

In contrast to the reverse KL divergence, the samples are here to be drawn from the density  $p$  of the theory, i.e.,  $\phi_i \sim p$ . As can be seen from the equation, the minimization of the forward KL corresponds to maximizing the likelihood of the model. In the machine learning literature, the forward training procedure is thus also known as *maximum likelihood training*. Indeed, this has already been explored in the context of lattice field theory [3].

This training procedure has the advantage that it is mode-covering since all modes of the physical target density  $p$  will necessarily be probed in training. It has however the disadvantage that it requires samples from the target  $p$ . In lattice applications, these are typically generated by a Monte-Carlo algorithm, such as HMC. However, if these configurations are available, one can directly measure physical observables on them and there is therefore no need to train a flow in the first place.

One may thus wonder if this training procedure is of any use then. For thermodynamic observables however, such as the free energy, one typically does not only require a single Markov chain for the target density  $p$  but a whole series of Markov chains along a discretized trajectory in the parameter space of the theory. As we will review in the next section, a flow allows us to completely avoid the need for these additional Markov chains. For the important class of thermodynamic observables, forward KL training is thus well justified and, as we will show, advisable.

## III. RELIABLE ESTIMATORS IN PRESENCE OF MODE-COLLAPSE

Combining deep generative models, e.g., normalizing flows, with neural importance sampling (NIS) has been shown to be a fruitful approach for estimating thermodynamic observables in lattice field theory [2,4], statistical mechanics [24], and chemistry [47–49]. This approach enables direct estimation of the free energy as well as other thermodynamic observables because flow-based sampling allows for estimating those observables at arbitrary points in the parameter space. Remarkably, this is in stark contrast to standard Markov-Chain Monte-Carlo methods which instead require nontrivial integration in the parameter space [2]. More specifically, NIS allows the computation of a direct Monte-Carlo estimate of the partition function which is crucial for many thermodynamic observables [2] such as entropy and free energy

$$F = -T \ln Z. \quad (9)$$

In the following, we revise two different estimators of the free energy, namely the  $p$ -estimator and the  $q$ -estimator. These allow estimation using samples drawn from the target and the generative model respectively [4]. We stress that these estimators are applicable for any generative model that has a tractable likelihood, such as normalizing flows [34], autoregressive neural networks [50], and diffusion models [51].

### A. Different estimators for the partition function

Given a generative model, irrespective of whether it has been trained using reverse-KL or forward-KL, the resulting sampler  $q_\theta$  is an approximation for the target density  $p$ . Leveraging this, recent works proposed to estimate the partition function of a physical system [2,24] directly at a given point in parameter space. This approach samples lattice configurations from the generative model  $q_\theta$  and estimates the partition function  $Z$  with a so-called  $q$ -estimator with

$$Z_q \equiv \mathbb{E}_{\phi \sim q_\theta} \left[ \frac{e^{-S(\phi)}}{q_\theta(\phi)} \right] \approx \frac{1}{N} \sum_{i=1}^N \frac{e^{-S(\phi_i)}}{q_\theta(\phi_i)} \equiv \hat{Z}_q, \quad \phi_i \sim q_\theta. \quad (10)$$

Alternatively, when samples from the target  $p$  are available, i.e., using a thermalized Markov chain, one can estimate the inverse partition function with the so-called  $p$ -estimator

$$Z_p^{-1} \equiv \mathbb{E}_{\phi \sim p} \left[ \frac{q_\theta(\phi)}{e^{-S(\phi)}} \right] \approx \frac{1}{N} \sum_{i=1}^N \frac{q_\theta(\phi_i)}{e^{-S(\phi_i)}} \equiv \hat{Z}_p^{-1}, \quad \phi_i \sim p. \quad (11)$$

Combining these results with Eq. (9), one immediately derives corresponding estimators for the free energy

$$\hat{F}_q = -T \log(\hat{Z}_q) \approx F_q = -T \log(Z_q), \quad (12)$$

$$\hat{F}_p = T \log(\hat{Z}_p^{-1}) \approx F_p = T \log(Z_p^{-1}). \quad (13)$$

Both the  $p$ -estimator and the  $q$ -estimator can be shown to be asymptotically consistent under the assumption that the supports of the flow  $q_\theta$  and the target density  $p$  match, i.e.,  $\text{supp}(q_\theta) = \text{supp}(p)$  [24]. By construction, the learned density,  $q_\theta$ , has full support over the entire domain of the base distribution—at least from a purely theoretical point of view. This implies  $\text{supp}(q_\theta) = \text{supp}(p)$  always holds in theory. However, in practice it is not unlikely to have regions of the domain where the density  $q_\theta(\phi)$  is vanishingly small. Hence, for a finite number of samples, it can effectively be zero. This leads to incorrect estimation of expectation values of physical observables for any reasonable number of samples  $N$ . Furthermore, ensuring that a normalizing flow is invertible also in practice, i.e., to numerical precision, can be very challenging [52].

To analyze the resulting implications for the estimation process, it is useful to define the following generalized notion of the support of the variational density

**Definition 1.** The effective support of the variational density  $q_\theta$  relative to  $p$  is given by

$$\widetilde{\text{supp}}_{p,\epsilon}(q_\theta) = \{\phi \in \text{supp}(q_\theta); q_\theta(\phi) > \epsilon p(\phi)\}, \quad (14)$$

for a given numerical threshold  $\epsilon$ . The mode dropping set is then given by

$$\mathcal{S} := \text{supp}(p) \setminus \widetilde{\text{supp}}_{p,\epsilon}(q_\theta). \quad (15)$$

This definition is useful for the following reason: if the flow is effectively mode dropping, i.e., the mode-dropping set  $\mathcal{S}$  is nonempty, the importance weighted estimator, with a finite number of samples  $N$ , will miss a contribution from the mass  $\int_{\mathcal{S}} p(\phi) d\phi$  with approximately the probability  $1 - \epsilon N \int_{\mathcal{S}} p(\phi) d\phi$ . We note that defining the threshold for the effective support *relative* to the target distribution is pivotal. This is because the absolute definition  $q(\phi) < \epsilon$  would have no meaning with regard to mode-dropping. As an example, suppose we have an area of size  $O(1/\epsilon)$  and  $p = q \leq \epsilon$  in that area. Then the corresponding area with  $O(1)$  probability mass would be considered “ $\epsilon$ -mode-dropped” even though  $q$  is an exact copy of  $p$ . We therefore

choose a definition for which mode-dropping only exists when  $q$  approximately vanishes *relative* to  $p$ .

It is also useful to define the *effective sampler distribution*

$$\tilde{q}_\theta(\phi) = \begin{cases} q_\theta(\phi)/\zeta & \text{if } \phi \in \widetilde{\text{supp}}_{p,\epsilon}(q_\theta), \\ 0 & \text{otherwise,} \end{cases} \quad (16)$$

where  $\zeta = \int_{\widetilde{\text{supp}}_{p,\epsilon}} \mathcal{D}[\phi] q_\theta(\phi) \leq 1$  represents the multiplicative renormalization factor necessary to guarantee the normalization of  $\tilde{q}_\theta$ , i.e., the probability mass out of the effective support  $\widetilde{\text{supp}}_{p,\epsilon}(q_\theta)$  is redistributed to the effective support proportionally to the original density  $q_\theta(\phi)$ . With this definition, we express the practical situation where the importance weighted estimator for a physical observable  $\mathcal{O}$  *typically* misses the contribution from the mode-dropping set  $\mathcal{S}$ , as the assumption that the following approximation holds:

$$\hat{\mathcal{O}} \equiv \frac{1}{N} \sum_{i=1}^N \frac{p(\phi_i)}{q_\theta(\phi_i)} \mathcal{O}(\phi_i) \approx \mathbb{E}_{\phi \sim \tilde{q}_\theta} \left[ \frac{p(\phi)}{q_\theta(\phi)} \mathcal{O}(\phi) \right] \equiv \bar{\mathcal{O}}, \quad (17)$$

where  $\phi_i \sim q_\theta$ , for the sample size large enough for Monte Carlo sampling but not too large to assume that  $\zeta^N \approx 1$ , i.e., the probability that all  $N$  samples drawn from  $q_\theta$  lie within the effective support is close to one. Since  $q_\theta$  has the full support, it holds that  $\text{supp}(\tilde{q}_\theta) = \widetilde{\text{supp}}_{p,\epsilon}(q_\theta)$ . Throughout the manuscript, we will indicate by a hat a (finite sample) estimator, by a bar the expectation over the effective distribution  $\tilde{q}$ —which corresponds to the average over *typical* samples—and by an asterisk the expectation over the original distribution  $q$ . Note that, under our assumption of mode-dropping (17), i.e.,  $\text{supp}(\tilde{q}_\theta) \not\supseteq \text{supp}(q_\theta) = \text{supp}(p)$ , the typical values of the estimator  $\hat{\mathcal{O}} \approx \bar{\mathcal{O}}$  can be significantly different from the true expectation value

$$\begin{aligned} \mathcal{O}^* &= \mathbb{E}_{\phi \sim p} [\mathcal{O}(\phi)] \\ &= \mathbb{E}_{\phi \sim q_\theta} \left[ \frac{p(\phi)}{q_\theta(\phi)} \mathcal{O}(\phi) \right] \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \frac{p(\phi_i)}{q_\theta(\phi_i)} \mathcal{O}(\phi_i), \quad \text{where } \phi_i \sim q_\theta. \end{aligned} \quad (18)$$

A more detailed discussion of the effective relative support is provided in Appendix A 2.

We also remark that  $\widetilde{\text{supp}}_{q_\theta,\epsilon}(p)$  is also interesting to consider, as  $\text{supp}(q_\theta) \setminus \widetilde{\text{supp}}_{q_\theta,\epsilon}(p) \neq \emptyset$  implies effective “*fake*” modes that are present in  $q_\theta$  but not in  $p$ .

The following theorem holds:

**Theorem 2.** Suppose the trained model is mode dropping, i.e., the approximation (17) holds. Then the  $q$ -estimator  $\hat{F}_q$  and the  $p$ -estimator  $\hat{F}_p$  for the free energy

approximate  $\bar{F}_q$  and  $\bar{F}_p$ , respectively, being bounds on the true free energy as

$$\bar{F}_q \geq F \geq \bar{F}_p.$$

Furthermore, if  $\widetilde{\text{supp}}_{p,\epsilon}(q_\theta) \supseteq \text{supp}(p)$  it follows

$$\bar{F}_q = F,$$

and similarly if  $\widetilde{\text{supp}}_{q_\theta,\epsilon}(p) \supseteq \text{supp}(q_\theta)$

$$\bar{F}_p = F.$$

We prove in Appendix A 3 that the estimators serve as upper and lower bounds of the free energy.

In the presence of mode-collapse, the flow has smaller effective support than the target, i.e.,

$$\text{supp}(p) \not\subseteq \text{supp}(\tilde{q}_\theta).$$

Crucially, this may also happen when the variational density  $q_\theta$  is a very *bad* approximation of the true density  $p$ , see bottom left of Fig. 5. While this is strictly not a common manifestation of mode collapse, the following discussion holds for such badly-trained models where the overlap between the support of  $q_\theta$  and  $p$  is very small. In this case, the q-estimator in Eq. (12) may thus lead to (possibly strongly) biased results since it may not have full effective support under the assumption that the approximation (17) holds. On the other hand, the p-estimator has the advantage that it is typically more efficient to sample directly from the flow while the p-estimator requires the (possibly costly) generation of configurations by a Markov Chain. Nevertheless, it is advisable to estimate the free energy with both estimators if there is a risk of mode mismatch and ensure that both lead to consistent results. The phenomenon of mode-collapse is a widely known issue in the field of density estimation [53–56]. In particular, when deploying generative models for physical systems, this becomes crucial as neglecting subsets of the modes of a target density would inevitably lead to highly biased estimation of physical quantities. Moreover, this may sometimes not even be detected unless appropriate estimators are used [4]. We want to stress that this problem is not restricted to lattice field theories [3,4] but is also found within other contexts, such as molecular systems [30,47,57]. Having an estimator which quantifies the amount of probability mass being missed by a variational ansatz is therefore highly desirable for more reliable and unbiased estimation of physical quantities.

When the trained model neglects some modes of the target density, hence missing full effective support over the target domain, estimates of physical observables may be biased. A desirable property of our framework is to detect such bias by providing reliable bounds on the error when

the model is mode-dropping. When  $q_\theta$  has full effective support on the domain of  $p$  the expected value of the importance weights  $w(\phi) = \frac{p(\phi)}{q_\theta(\phi)}$  reads

$$\begin{aligned} w^* &= \mathbb{E}_{q_\theta} \left[ \frac{p(\phi)}{q_\theta(\phi)} \right] = \int_{\text{supp}(q_\theta)} q_\theta(\phi) \frac{p(\phi)}{q_\theta(\phi)} \mathcal{D}[\phi] \\ &= \int_{\text{supp}(q_\theta)} p(\phi) \mathcal{D}[\phi] = 1. \end{aligned} \quad (19)$$

This expectation value thus measures the degree to which the support of the target density  $p$  is covered by the sampler  $q_\theta$ . Statistically, in the limit of infinite measurements,  $w^*$  is always equal to one. However, if the sampler is mode-dropping, hence the approximation (17) holds, then the estimator  $\bar{w}$  will be in  $[0, 1]$  providing us with a natural quantity to measure the sampler's ability to probe the entire support of the target density  $p$ .

We will now derive an estimator for this expectation value. To this end, we rewrite the above expression as

$$\bar{w} \equiv \frac{1}{Z} \mathbb{E}_{\phi \sim \tilde{q}_\theta} \left[ \frac{e^{-S(\phi)}}{q_\theta(\phi)} \right] \quad (20)$$

and we note that the expectation value is now taken with respect to  $\tilde{q}_\theta$ . As shown in the last section, the partition function  $Z$  can be approximated by the  $p$ -estimator (11) when samples from the target density are available. Thus, under the assumption that the approximation (17) holds, the following Monte-Carlo estimate approximates Eq. (20), i.e.,

$$\begin{aligned} \bar{w} &\approx \frac{1}{\hat{Z}_p} \left( \frac{1}{N} \sum_{i=1}^N \frac{e^{-S(\phi_i)}}{q_\theta(\phi_i)} \right) \\ &= \left( \frac{1}{N} \sum_{j=1}^N \frac{q_\theta(\phi_j)}{e^{-S(\phi_j)}} \right) \left( \frac{1}{N} \sum_{i=1}^N \frac{e^{-S(\phi_i)}}{q_\theta(\phi_i)} \right) \equiv \hat{w}, \end{aligned} \quad (21)$$

where  $\phi_i \sim q_\theta$  and  $\phi_j \sim p$  are sampled from the flow and the target density  $p$  respectively.

## B. Bounding the bias of physical observables

Following [2], given a physical observable  $\mathcal{O}$ , our goal is to compute the importance weighted estimator  $\hat{\mathcal{O}}$ , defined in the left-hand side of (17), which approximates the expectation value  $\bar{\mathcal{O}}$  over the effective sampler distribution (16). This estimator is not necessarily unbiased to the true value (18), if the model  $q_\theta$  is affected by mode-collapse, i.e.,  $\text{supp}(p) \not\subseteq \text{supp}(\tilde{q}_\theta)$  for the approximation (17) to hold. Similarly, the bias of the estimator evaluated over a finite number of trials should approximate

$$|\bar{\mathcal{O}} - \mathcal{O}^*| = \left| \int (1 - 1_{\text{supp}(\tilde{q}_\theta)}(\phi)) \mathcal{O}(\phi) p(\phi) \mathcal{D}[\phi] \right|, \quad (22)$$

i.e., the bias arises due to the insufficient effective support of the sampler.

#### IV. THE MODE-DROPPING ESTIMATOR

In the following, we aim to derive a bound on this bias. In the process, we will also obtain a natural measure for the degree of mode-collapse. To this end, we foliate the sampling space by disjunct sets

$$\Delta_n := \{\phi | e^{D_\vee+1-n} \geq e^{-S(\phi)} > e^{D_\vee-n}\} \quad (23)$$

for  $n \in \mathbb{N}$  and  $D_\vee \in \mathbb{R}_{\geq 0}$ . We refer to Fig. 3 for a visual illustration of the foliation. For fixed  $n \in \mathbb{N}$ , we can define the following weights

$$\begin{aligned} \alpha_n &= \int_{\Delta_n} (1 - 1_{\text{supp}(\tilde{q}_\theta)}(\phi)) p(\phi) \mathcal{D}[\phi] \\ &= \frac{1}{Z} \left( \int_{\Delta_n} e^{-S(\phi)} \mathcal{D}[\phi] - \int_{\Delta_n \cap \text{supp}(\tilde{q}_\theta)} e^{-S(\phi)} \mathcal{D}[\phi] \right). \end{aligned} \quad (24)$$

Leveraging these definitions, we derive a bound on the bias in Appendix A 5 which is summarized in the following theorem.

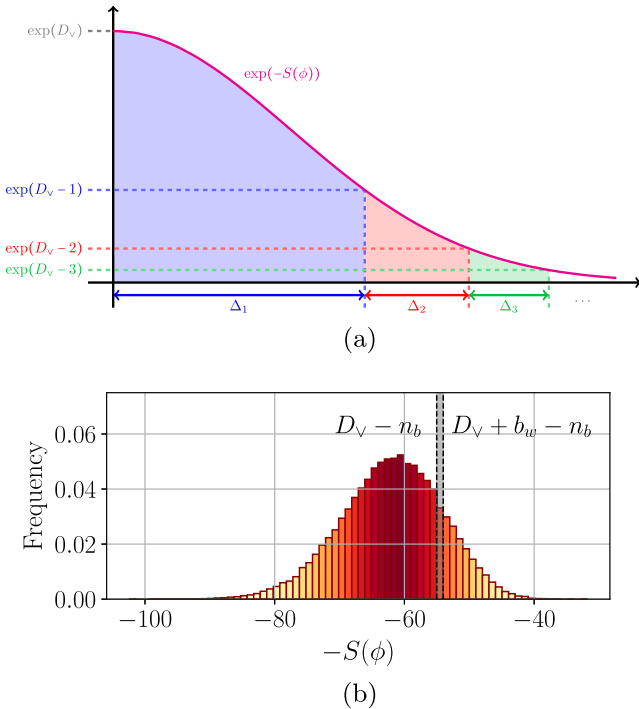


FIG. 3. Foliation—(a) visualization of the foliation according to Eq. (23) showing  $\Delta_1$ ,  $\Delta_2$ , and  $\Delta_3$  explicitly. (b) Field configurations are distributed into buckets. The labels  $n_b$  and  $b_w$  thereof represent the bucket index and the width of each bucket, respectively.

**Theorem 3.** Let the action  $S$  of the theory and the observable  $\mathcal{O}$  be polynomially bounded, i.e.,

$$C_\vee \|\phi\|^{\alpha_\vee} - D_\vee \leq S(\phi) < C_\wedge \|\phi\|^{\alpha_\wedge} + D_\wedge \quad (25)$$

for some  $C_{\vee,\wedge}, D_{\vee,\wedge}, \alpha_{\vee,\wedge} \in \mathbb{R}_{\geq 0}$  and

$$|\mathcal{O}(\phi)| \leq c \|\phi\|^\alpha + d \quad (26)$$

for some  $c, \alpha, d \in \mathbb{R}_{\geq 0}$ . The bias, i.e., the difference between the expectation value  $\bar{\mathcal{O}}$  and the true value  $\mathcal{O}^*$ , then satisfies

$$|\bar{\mathcal{O}} - \mathcal{O}^*| \leq \sum_{n \in \mathbb{N}} \sup_{\phi \in \Delta_n} |\mathcal{O}(\phi)| \cdot \alpha_n. \quad (27)$$

The bias is therefore bounded by a weighted sum over the  $\alpha_n$ . The weighting of each summand depends on the observable  $\mathcal{O}$  of interest. We note that the present discussion is only relevant for noncompact variables. Indeed, continuous functions living on *compact* manifolds  $\mathcal{M}$  are integrable, and therefore no additional care is required to handle indefinite forms of the type  $0 \cdot \infty$ . In particular, it follows that for a compact variable  $\mathcal{O}$  the bound is straightforward

$$|\bar{\mathcal{O}} - \mathcal{O}^*| \leq \sup_{\phi \in \mathcal{M}} |\mathcal{O}(\phi)| \cdot \bar{w}. \quad (28)$$

We note that many physical observables are simple powers of fields, i.e.,  $\mathcal{O}(\phi) = \|\phi\|^k$  for  $k \in \mathbb{N}$ . It can be shown that the foliation (23) along with the polynomial bound of the action  $S$  implies that

$$\|\phi\| < u_n, \quad (29)$$

where we have defined  $u_n = \left(\frac{n}{C_\vee}\right)^{\frac{1}{\alpha_\vee}}$ . We refer to Appendix A 4 for more details. For such observables, the bias can thus be bounded by

$$|\bar{\mathcal{O}} - \mathcal{O}^*| \leq \sum_{n \in \mathbb{N}} u_n^k \alpha_n. \quad (30)$$

The theorem also naturally relates to the quantity  $\bar{w}$  introduced in the last section which quantifies the degree of mode-collapse. In order to provide a single number for the degree of the mode-collapse of the sampler, it is natural to choose a uniform, i.e., observable agnostic, weighing. It then follows from the definition of the  $\alpha_n$ , see (24), that this weighing measures the mismatch in support between the sampler and the target density

$$\sum_{n \in \mathbb{N}} \alpha_n = \int (1 - 1_{\text{supp}(\tilde{q}_\theta)}(\phi)) p(\phi) \mathcal{D}[\phi] = 1 - \bar{w} \quad (31)$$

and is thus directly related to the mode-dropping estimator  $\bar{w}$  derived in the last section.

## V. NUMERICAL EXPERIMENTS

We evaluate our proposed methods to detect and mitigate mode-collapse using the two-dimensional scalar  $\phi^4$ -theory with action

$$S[\phi] = \sum_{x \in \Lambda} -2\kappa \sum_{\hat{\mu}=1}^2 \phi(x)\phi(x + \hat{\mu}) + (1 - 2\lambda)\phi(x)^2 + \lambda\phi(x)^4, \quad (32)$$

where  $\lambda$  is the bare coupling while  $\kappa$  is the hopping parameter. We refer to [2] for more details on this hopping parameterization of the action. Throughout all our experiments we keep the bare coupling fixed at  $\lambda = 0.022$  and vary  $\kappa$  such that the theory crosses the phase transition due to the spontaneous breaking of its  $\mathbb{Z}_2$  symmetry, i.e.,  $\phi \rightarrow -\phi$ . As the hopping parameter  $\kappa$  increases, spontaneous magnetization is observed. This is illustrated in Fig. 4 for which the hopping parameter takes values through the critical region around  $\kappa_c \approx 0.275$ . The curves show the density (top) and log density (bottom) of the normalized magnetization with different colors referring to different values of the hopping parameter  $\kappa$ . Spontaneous symmetry breaking is observed as the distribution of the magnetization changes from a wider single-mode to a bimodal density with a suppressed tunneling probability between the two modes. This suppression is accentuated as the value of  $\kappa$  increases.

### A. Free energy estimators

Our first numerical experiment analyzes the performance of two normalizing flows trained with both objectives described in Sec. II A. We refer to those flows as the

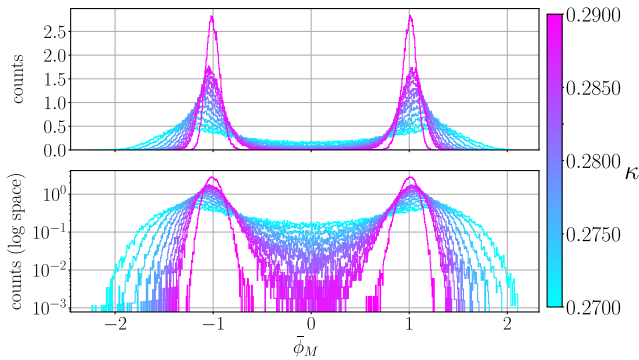


FIG. 4. Spontaneous symmetry breaking—spontaneous symmetry breaking for the  $\phi^4$  as a function of the hopping parameter using the action in Eq. (32). The top and bottom figures show a histogram of the magnetization—self-normalized by its absolute value—in both linear and log-scale respectively. The self-normalization makes the modes centered around  $+1$  and  $-1$ . From those histograms we thus observe that the probability between the modes reduces as the hopping parameter  $\kappa$  increases.

forward-KL flow and the reverse-KL flow if they were trained with maximum likelihood or self-sampling respectively. We train for a hopping parameter  $\kappa = 0.5$  such that the theory is in its broken phase, see Fig. 5. For maximum likelihood training, we use 50M samples generated by an overrelaxed HMC. Following [2], we choose an architecture for the (reverse-KL) normalizing flows such that those models are manifestly invariant under  $\mathbb{Z}_2$  symmetry (blue). In order to highlight the effects of mode-collapse, we also train reverse-KL flows without the  $\mathbb{Z}_2$  inductive bias (green) thus expecting these models to be prone to mode-collapse.

The reference estimates for the true free energy were obtained via HMC simulations. Similarly to the approach followed in [2] such estimates are obtained by discretizing the hopping parameter space so that free energy differences can be estimated via HMC along the trajectory. Those contributions are added, integrating such trajectory up to the desired point at which the free energy needs to be estimated. Further technical details on how deep generative models were trained and HMC reference values obtained, can be found in Appendix A 6.

As can be seen on the left-hand side of Fig. 5, the forward-KL flow (orange) very closely reproduces the reference distribution by HMC (pink). For the reverse-KL trained flows, we see that for smaller systems (top row in Fig. 5), leveraging the  $\mathbb{Z}_2$  inductive bias leads to a good approximation (blue) while the non- $\mathbb{Z}_2$  equivariant flow (green) fails to capture both modes. For larger systems, instead, both  $\mathbb{Z}_2$  equivariant, and nonequivariant, flows are not able to capture most of the support of the target density  $p$ , thereby resulting in poor approximations.

On the right-hand side of Fig. 5, we estimate the free energy density of the system using different flows. We use the same color scheme as on the left-hand side and measure the free energy using both the  $p$ -estimator (circle) Eq. (13) and the  $q$ -estimator (square) Eq. (12) of the free energy. Our numerical results indeed agree with the theoretical prediction of theorem 2. Specifically, with the model trained with maximum likelihood, both estimators lead to compatible predictions with the HMC estimator. This is consistent with the left-hand side of the plot which suggests that no mode-collapse took place for this model.

For the reverse-KL flow, however, such an agreement may not be expected as the left-hand side of Fig. 5 shows a mismatch in the support. The right-hand side plots show that the non- $\mathbb{Z}_2$  equivariant flow (green) in the  $\Lambda = 16 \times 8$  (top row) case is dropping the left-hand mode while its  $\mathbb{Z}_2$ -equivariant counterpart (blue) covers both modes. Nonetheless, as the dimensionality increases the density estimation task becomes increasingly more challenging thus preventing the  $\mathbb{Z}_2$ -equivariant reverse-KL flows to train effectively for  $\Lambda = 64 \times 8$ . As a result, we find that the  $q$ -estimator overestimates the true value  $F$  for both lattice sizes (top and bottom rows), while for the  $\Lambda = 64 \times 8$



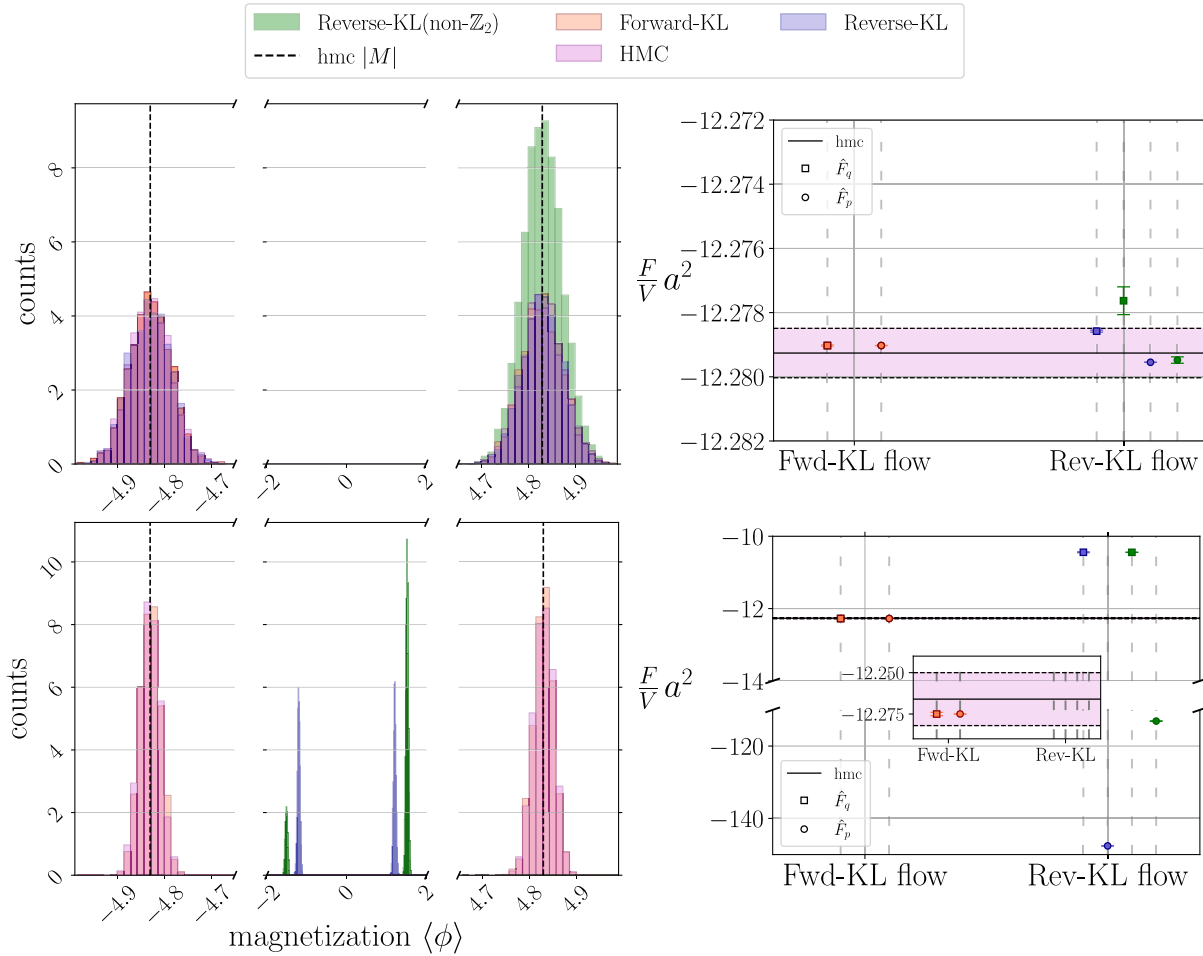


FIG. 5. Histogram of magnetization (left) and free energy estimates (right) for flow-based models trained with forward- and reverse KL at  $\kappa = 0.5$  for  $\Lambda = 16 \times 8$  (top) and  $\Lambda = 64 \times 8$  (bottom)—The left-hand side shows histograms of the magnetization for configurations sampled from a forward-KL trained normalizing flow (orange), a  $\mathbb{Z}_2$  equivariant reverse-KL flow (purple) a simple non- $\mathbb{Z}_2$ -equivariant reverse-KL flow (green) and an overrelaxed HMC (pink). The same choice of colors applies to the right hand side plots. Models were used to sample configurations at  $\kappa = 0.5$ , i.e., well in the broken phase. Dashed black lines are the expected values for the absolute magnetization (and its negative value) using HMC. The right-hand side plot shows estimates of free energy densities using the estimators introduced in Sec. III for both flow models. Their estimates are computed and then compared to the reference HMC baseline (denoted by the solid black line while the pink surroundings show the standard deviation) and following the method discussed in [2] and Appendix A 6. The  $q$ - and the  $p$ -estimators for the free energy are shown using different markers. Error bars—based on the standard deviation—for the flow-based estimate of the free energy density are often not visible as they tend to be several orders of magnitudes smaller than the plot scale. For the  $\Lambda = 64 \times 8$  case the mismatch between the reverse-KL flows and the HMC histograms, shown in the left-hand side plot, is reflected in highly biased estimates (dark blue markers).

case the  $p$ -estimator substantially underestimates  $F$ , i.e.,  $\hat{F}_{q_0} \geq F \geq \hat{F}_p$ , as predicted by theorem 2. This latter situation suggests that the reverse-KL flows (green and blue) are limited in approximating of the target density resulting in very different effective supports. While strictly speaking this is not usually referred to as mode-collapse, it can be understood through the same lenses.

These experiments thus illustrate that: (a) a mode-covering objective such as the forward-KL is more resilient when the target density is multimodal and shows sparse effective support, and (b) when the effective support does not match, both  $p$ - and  $q$ -estimators of the free energy from

Sec. III give upper and lower bound respectively. Moreover, we note that training using a forward KL objective does not worsen the performance compared to using a reverse KL. Practically, if the variational distribution presents some “fake” modes,  $\phi$  s.t.  $q_\theta(\phi)/p(\phi) \gg 1$ , field configurations sampled in these regions will always be exponentially suppressed in the reweighting phase. We emphasize once again that a significant drawback of training using the pure form of forward KL is the necessity for training samples. Although this limitation applies in general, it does not pose a problem for our specific task of estimating thermodynamic observables.

We repeated this analysis for a number of values of the hopping parameter  $\kappa$ . The results are summarized in Fig. 6 for the larger lattice with  $\Lambda = 64 \times 8$ . We evaluate the gap between the neural importance sampling (NIS) estimate  $\hat{F}$  and the HMC reference normalized by the total standard error. Namely, if the normalized gap is within the range  $[-1, +1]$ , both estimators are compatible (see inset in the top plot of Fig. 6). Dashed curves connect  $q$ -estimates (12), while solid curves connect  $p$ -estimates (13), of the free energy at different values of  $\kappa$ . The results obtained with  $\mathbb{Z}_2$ -equivariant reverse-KL and forward-KL flows are shown in blue and orange respectively. The inset shows close agreement of both estimators and both flow models for  $\kappa \in \{0.2, 0.3\}$ . However, deep in the broken phase, e.g.,  $\kappa \geq 0.4$ , the two modes of the target distribution start to lay further apart resulting in a failure of the mode-seeking objective, i.e., the reverse-KL, to properly capture the target

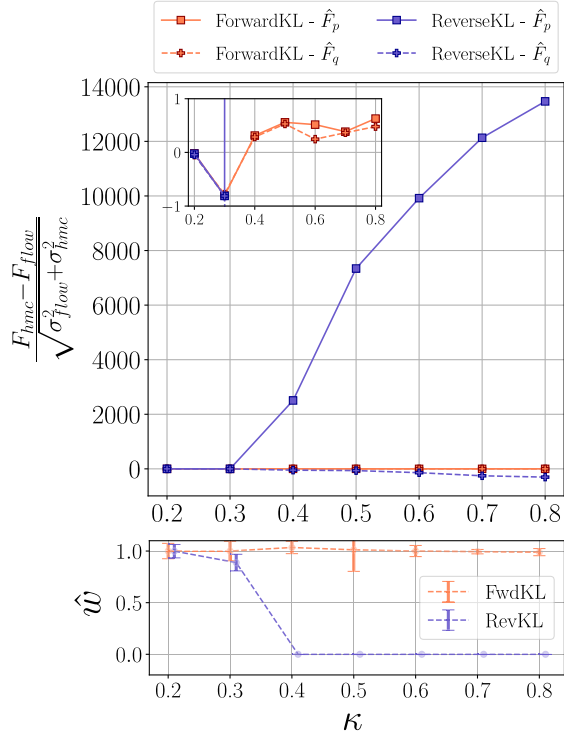


FIG. 6. Analysis of the free energy estimates for forward and reverse KL flows for a  $64 \times 8$  lattice and different hopping parameters  $\kappa$ —colors denote flow models trained with different objectives, namely the forward-KL (orange) and reverse-KL (purple). Markers refer to the two different estimators of the free energy introduced in Eqs. (12) and (13). Specifically, *plus* and *square* are used respectively. Every point in the inset shows that flow and HMC estimates are compatible, i.e., the gap is within the statistical uncertainty. The lower plot relates the results from above to the mode-dropping estimator Eq. (21). This demonstrates that when the flow is a good approximator for  $p$ , the estimator is close to one. When modes of the distribution are missed, e.g.,  $q_\theta$  is a bad approximator of  $p$ , the estimator of  $\bar{w}$  quickly decays to zero.

density, see also bottom left plot of Fig. 5. As a result, the probability mass transport induced by the normalizing flow fails to reproduce the correct target distribution  $p$ , leading to a larger gap between the  $p$ - and  $q$ -estimators. When using the forward-KL trained flow instead, the support of the sampler is closely matching the support of the target hence making free energy compatible with the HMC reference even at the higher values of the hopping parameter  $\kappa$ , when Eq. (17) holds. This effect is shown in the inset where there is a good agreement between both estimators of the free energy. This observation suggests that the mode-covering nature of the forward-KL is crucial to ensure that the flow leads to unbiased estimates of physical observables.

In Fig. 6, it is also shown that our proposed mode-dropping estimator (21) correlates well with the observed gap in the free energy estimation. Lastly, we use our estimator  $\bar{w}$  to evaluate the support-mismatch of forward and reverse flow models trained at several  $\kappa$  values and different lattice sizes as shown in Fig. 7. The top and bottom plots refer to lattices of size  $\Lambda = 16 \times 8$  and  $\Lambda = 64 \times 8$  respectively. These results demonstrate that the quality of the

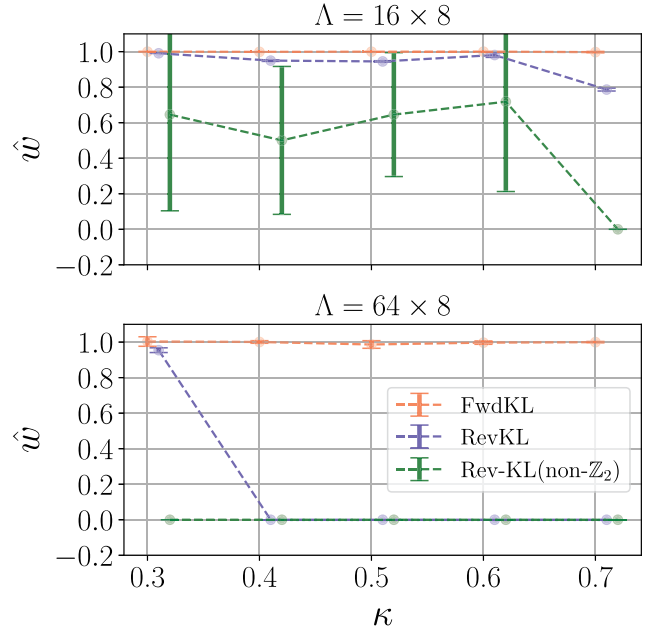


FIG. 7. Mode dropping evaluation as a function of the hopping parameter  $\kappa$  for fixed  $\lambda$  and two different lattice sizes—mode dropping estimator for different values of the hopping parameter and different lattice sizes. For each setup, three normalizing flows were trained with reverse- and forward-KL objectives as described in Sec. II A. The former models are again trained with (orange) and without (green) built-in  $\mathbb{Z}_2$  equivariance. The upper and lower plots show the mode-dropping estimator Eq. (21) for lattices  $\Lambda = 16 \times 8$  and  $\Lambda = 64 \times 8$ , respectively. Unsurprisingly, we observe that mode-collapse gets more severe as the lattice size increases. This is reflected in a stronger decay toward zero of the estimator for larger volumes.

sampler very quickly deteriorates in the broken phase due to mode-collapse for the model trained by self-sampling. This is not the case for models trained with the forward KL. Indeed, as shown in Fig. 7, these models scale significantly better in the volume of the system. Furthermore, the non- $\mathbb{Z}_2$  equivariant reverse-KL flow (green), is manifestly mode dropping for  $\Lambda = 16 \times 8$ , see Fig. 5, with values of  $\bar{w}$  around 0.5 for values  $0.4 \leq \kappa \leq 0.7$ . This agrees with the left-hand side of Fig. 5 where only half of the support is covered by the learned variational density in the top row.

## VI. OUTLOOK AND SUMMARY

Mode-collapse presents a significant limitation to flow-based sampling on the lattice because it may lead to inaccurate approximations of the target density, either partially or completely. Intuitively, it can be understood as being in a loose relation to the tunneling problem in local MCMC algorithms. Specifically, the algorithmic challenges in sampling from multimodal distributions are shifted from the sampling to the training phase for normalizing flows. In this work, we have studied this important limitation of flow-based sampling in great detail. We argue that in the important case of thermodynamic observables, there are practical and theoretically grounded mitigation strategies available. Specifically, the flow can be trained using the forward KL divergence and the free energy can be evaluated with two estimators that bound the true value. Furthermore, we have analyzed mode-mismatch theoretically and derived a bound on its induced bias as well as a quantitative measure for its severity. Normalizing flows are currently only limited to toy models. Encouragingly, we also observed as a side-product of our analysis, that the forward KL objective leads to better scaling in the system size. This observation may be worthwhile to be studied further as part of future work.

## ACKNOWLEDGMENTS

The authors thank the referee for stimulating discussions and useful suggestions that significantly improved the manuscript. K. A. N., C. J. A., S. N., and P. K. are supported by the German Ministry for Education and Research (BMBF) as BIFOLD—Berlin Institute for the Foundations of Learning and Data under the Grant No. BIFOLD23B. K. A. N. has been partially supported by the Einstein Research Unit Quantum (ERU) Project under Grant No. ERU-2020-607. This work is supported with funds from the Ministry of Science, Research, and Culture of the State of Brandenburg within the Centre for Quantum Technologies and Applications (CQTA). This work is funded by the European Union’s Horizon Europe Framework Program (HORIZON) under the ERA Chair scheme with Grant Agreement No. 101087126. This work

is funded by the European Union’s HORIZON MSCA Doctoral Networks programme and the AQTIVATE project (101072344). The authors acknowledge Lena Funcke and Paolo Stornati for helpful discussions.

## APPENDIX: SUPPLEMENTAL MATERIAL

### 1. Forward-KL training

Training a normalizing flow with forward-KL in the context of lattice field theory requires pre-generated samples at a given point in parameter space. Before training a flow model, one should instantiate a thermalized Markov chain at a fixed value of the coupling parameters and generate a sufficient number of Monte-Carlo configurations which are then used to train the flow. A pseudo-code for this approach is presented in algorithm 1. We note that practically this approach may not always be feasible. For example, the number of pre-generated configurations needed for training a flow to an acceptable accuracy increases as the size of the lattice grows. For instance, training a flow for a  $64 \times 8$  lattice in the context of the  $\phi^4$  field theory, in the broken phase, requires already more than fifty million samples. This problem, therefore, limits the practical deployment of forward-KL training schemes at larger scales. Moreover, another limitation of such an approach is that generating samples with HMC may not always be possible. Indeed, in the proximity of a phase transition, long-range autocorrelation will prevent to sample a necessary large amount of uncorrelated samples in time. One would therefore need to be very careful in generating a suitable dataset of HMC configurations to avoid incorporating any additional unwanted bias when training the flow.

### 2. Relative effective support

Let  $B(\phi, r)$  be the open ball centered at  $\phi$  with radius  $r$ . A point  $\phi$  is called  $\epsilon$ -dropped if and only if

$$\limsup_{r \rightarrow 0} \frac{\int_{B(\phi, r)} q_\theta(\phi') d\phi'}{\int_{B(\phi, r)} p(\phi') d\phi'} < \epsilon. \quad (\text{A1})$$

By the Lebesgue differentiation theorem, this implies that  $q_\theta(\phi) \leq \epsilon p(\phi)$  holds for almost every  $\epsilon$ -dropped  $\phi$ , and if  $p$  and  $q_\theta$  are continuous, it actually means  $q_\theta(\phi) \leq \epsilon p(\phi)$ . We recall the definition of effective relative support Eq. (14)

$$\widetilde{\text{supp}}_{p, \epsilon}(q_\theta) := \{\phi \in \text{supp}(q_\theta); \phi \text{ not } \epsilon\text{-dropped}\}. \quad (\text{A2})$$

Setting  $\mathcal{S} := \text{supp}(p) \setminus \widetilde{\text{supp}}_{p, \epsilon}(q_\theta)$  and assuming  $\mathcal{S} \neq \emptyset$  means that the importance weighted estimator with  $N$

Algorithm 1. To train an NF with forward-KL in the context of lattice field theory we need to generate training configurations from a thermalized HMC chain. This HMC pre-sampling process is made more efficient by running  $C_{\max}$  independent chains in parallel so that the runtime to sample the entire dataset  $\Phi$  is constant in the total number of samples  $B$ . The total number of samples will therefore be  $B = nC_{\max}$ . The sampling is done between line 2 and line 5 in the algorithm below. Once the dataset is sampled and stored on disk, one starts training up to  $T_{\max}$  iterations. Per iteration, one draws batches of  $m$  configurations from  $\Phi$  in line 7 and uses them to evaluate the expectation value of the log-density of  $q_\theta$  and compute the gradient of the forward KL objective in line 8 and line 9 respectively. The model weights are then updated and the learned bijection  $f_\theta$  along with the variational density  $q_\theta$  are returned by the algorithm at the end of its training steps.

---

**Input:**

- prior density, e.g.,  $q_z \sim \mathcal{N}(0, \mathbb{I})$
- parametric model with parameters  $\theta$
- parametric action  $S(\phi, \kappa, \lambda)$  with fixed coupling parameters  $\lambda$  and  $\kappa$
- empty tensor for storing a batch of  $B$  configurations  $\Phi \in \mathbb{R}^{B \times N_s \times N_T}$

**Results:**

- learned bijective transformation  $f_\theta$  s.t.  $\phi_i = f_\theta(z_i)$
- exact likelihood function  $q_\theta$

```

1: begin
   /* Generate training samples from HMC */
2: for  $c$  in  $\{1, \dots, C_{\max}\}$  do
3:    $\phi = \text{sampleHMC}(\lambda, \kappa)$ ; // sample configurations  $\phi \in \mathbb{R}^{n \times N_s \times N_T}$ 
4:    $\Phi = \text{concatenate}(\Phi, \phi)$ ; // concatenate configurations
5: end
   /* Flow training using the generated dataset  $\Phi$  */
6: for  $t$  in  $1, \dots, T_{\max}$  do
   /* Iterate over dataset to collect batches of configurations */
7:   draw samples  $\phi = \{\phi_i\}_{i=1}^m$  from  $\Phi$  where  $\phi_i \in \mathbb{R}^{N_s \times N_T} \forall i \in \{1, \dots, m\}$ 
8:   use  $\phi$  to evaluate  $\frac{1}{m} \sum_{i=1}^m \ln q_\theta(\phi_i)$ ;
9:    $\Delta\theta = \arg \min_{\theta} \mathbb{E}_p[\ln q_\theta(\phi)]$ ;
10:  update  $f_\theta$  with  $\theta \leftarrow \theta + \eta \Delta\theta$ ;
11: end
12: end
13: return  $f_\theta, q_\theta$ 

```

---

samples lacks a contribution from the mass  $\int_S p(\phi) d\phi$  with probability

$$\begin{aligned} \left(1 - \int_S q_\theta(\phi) d\phi\right)^N &> \left(1 - \epsilon \int_S p(\phi) d\phi\right)^N \\ &\approx 1 - \epsilon N \int_S p(\phi) d\phi. \end{aligned} \quad (\text{A3})$$

### 3. Proof of Theorem 2

**Theorem 2.** Suppose the trained model is mode dropping, i.e., the approximation (17) holds. Then the  $q$ -estimator  $\hat{F}_q$  and the  $p$ -estimator  $\hat{F}_p$  for the free energy approximate  $\bar{F}_q$  and  $\bar{F}_p$ , respectively, being bounds on the true free energy as

$$\bar{F}_q \geq F \geq \bar{F}_p.$$

Furthermore, if  $\widetilde{\text{supp}}_{p,\epsilon}(q_\theta) \supseteq \text{supp}(p)$  it follows

$$\bar{F}_q = F,$$

and similarly if  $\widetilde{\text{supp}}_{q_\theta,\epsilon}(p) \supseteq \text{supp}(q_\theta)$

$$\bar{F}_p = F.$$

*Proof.* From the definition of the free energy  $F = -T \ln Z$  we first note that  $\bar{F}_q \geq F$  is equivalent to  $\bar{Z}_{q_\theta} \leq Z$ . Using the fact that  $\text{supp}(e^{-S(\phi)}) = \text{supp}(p)$ , we obtain [58]

$$\begin{aligned} \bar{Z}_{q_\theta} &\equiv \mathbb{E}_{\tilde{q}_\theta}[\tilde{w}(\phi)] \\ &= \int_{\widetilde{\text{supp}}(q_\theta)} \mathcal{D}[\phi] q_\theta(\phi) \frac{e^{-S(\phi)}}{q_\theta(\phi)} \\ &= \int_{\widetilde{\text{supp}}(q_\theta)} \mathcal{D}[\phi] e^{-S(\phi)} \\ &= \int_{\widetilde{\text{supp}}(q_\theta) \cap \text{supp}(p)} \mathcal{D}[\phi] e^{-S(\phi)} \leq \int_{\text{supp}(p)} \mathcal{D}[\phi] e^{-S(\phi)} = Z, \end{aligned}$$

where the last inequality holds because  $e^{-S(\phi)} \geq 0$ . Thus, we conclude  $\bar{F}_q \geq F$  with the corollary that  $\widetilde{\text{supp}}(q_\theta) \supseteq \text{supp}(p)$  implies equality  $\bar{F}_q = F$ .

Similarly,

$$\begin{aligned}
\bar{Z}_p^{-1} &\equiv \mathbb{E}_{\bar{p}} \left[ \frac{q_\theta(\phi)}{e^{-S(\phi)}} \right] \\
&= \frac{1}{Z} \int_{\widetilde{\text{supp}}(p)} \mathcal{D}[\phi] e^{-S(\phi)} \frac{q_\theta(\phi)}{e^{-S(\phi)}} \\
&= \frac{1}{Z} \int_{\widetilde{\text{supp}}(p)} \mathcal{D}[\phi] q_\theta(\phi) \\
&= \frac{1}{Z} \int_{\widetilde{\text{supp}}(p) \cap \text{supp}(q_\theta)} \underbrace{\mathcal{D}[\phi] q_\theta(\phi)}_{\geq 0} \leq \frac{1}{Z} \int_{\text{supp}(q_\theta)} \mathcal{D}[\phi] q_\theta(\phi) \\
&= Z^{-1}
\end{aligned}$$

shows  $\bar{F}_p \leq F$ , in general, and  $\bar{F}_p = F$  given  $\widetilde{\text{supp}}(p) \supseteq \text{supp}(q_\theta)$ .

Hence, by combining the inequalities we can conclude

$$\bar{F}_q \geq F \geq \bar{F}_p. \quad \blacksquare$$

#### 4. Bound on the configuration

Let us assume  $S(\phi)$  to be polynomial bounded

$$C_\vee \|\phi\|^{\alpha_\vee} - D_\vee \leq S(\phi) < C_\wedge \|\phi\|^{\alpha_\wedge} + D_\wedge \quad (\text{A4})$$

with non negative coefficients  $C_\vee, C_\wedge, D_\vee, D_\wedge, \alpha_\vee, \alpha_\wedge \in \mathbb{R}_{\geq 0}$ . The left-hand and right-hand sides represent the lower and the upper bounds on the action  $S$ . One needs to find appropriate coefficients  $C_{\vee,\wedge}, \alpha_{\vee,\wedge}, D_{\vee,\wedge}$  such that the inequalities are satisfied. We now do a foliation of the sampling space

$$\Delta_n := \{\phi | e^{(D_\vee+1-n)} \geq e^{-S(\phi)} > e^{(D_\vee-n)}\} \quad (\text{A5})$$

which can be seen as a re-distribution of the lattice configurations  $\phi$  into infinitely many buckets labeled by the index  $n$ . Combining Eqs. (A4) and (A5) one can rewrite a condition on the norm of  $\phi$  [59]. For a configuration  $\phi \in \Delta_n$

$$C_\vee \|\phi\|^{\alpha_\vee} - D_\vee \leq S(\phi) < -D_\vee + n. \quad (\text{A6})$$

This implies

$$C_\vee \|\phi\|^{\alpha_\vee} < n \Rightarrow \|\phi\| < \left( \frac{n}{C_\vee} \right)^{\frac{1}{\alpha_\vee}} \quad (\text{A7})$$

for  $n > 0$ . On the other side, it follows that

$$C_\wedge \|\phi\|^{\alpha_\wedge} + D_\wedge \geq S(\phi) \geq -D_\vee - 1 + n, \quad (\text{A8})$$

which implies that

$$\begin{aligned}
C_\wedge \|\phi\|^{\alpha_\wedge} + D_\wedge &\geq -D_\vee - 1 + n \\
\Rightarrow \|\phi\| &\geq \left( \frac{-D_\vee - D_\wedge - 1 + n}{C_\wedge} \right)^{\frac{1}{\alpha_\wedge}} \quad (\text{A9})
\end{aligned}$$

for  $C_\wedge > 0, n > 0$ . Combining Eqs. (A7) and (A9) we obtain the following bounds on the norm of the lattice configuration

$$\underbrace{\left[ \frac{-D_\vee - D_\wedge - 1 + n}{C_\wedge} \right]^{\frac{1}{\alpha_\wedge}}}_{l_n} \leq \|\phi\| < \underbrace{\left( \frac{n}{C_\vee} \right)^{\frac{1}{\alpha_\vee}}}_{u_n}. \quad (\text{A10})$$

In particular, this can be shown to imply a bound on the volume of  $\Delta_n$ , e.g., making the volume finite

$$\text{vol}(\Delta_n) \leq \int_{l_n \leq \|\phi\| < u_n} d\phi = \frac{\pi^{\frac{N}{2}}}{\Gamma(\frac{N}{2} + 1)} \cdot (u_n^N - l_n^N), \quad (\text{A11})$$

where we used the volume of the  $N$ -ball,

$$\frac{\pi^{\frac{N}{2}}}{\Gamma(\frac{N}{2} + 1)} \cdot r^N.$$

#### 5. Proof of Theorem 3

We now leverage the result from Appendix A 4 to derive a bound on the bias for the general observable  $\mathcal{O}$  when  $\widetilde{\text{supp}}(q_\theta) \subset \text{supp}(p)$  and therefore the importance sampling estimator may not be unbiased when Eq. (17) holds true.

**Theorem.** Let the action  $S$  of the theory and the observable  $\mathcal{O}$  be polynomially bounded, i.e.,

$$C_\vee \|\phi\|^{\alpha_\vee} - D_\vee \leq S(\phi) < C_\wedge \|\phi\|^{\alpha_\wedge} + D_\wedge \quad (\text{A12})$$

for  $C_{\vee,\wedge}, D_{\vee,\wedge}, \alpha_{\vee,\wedge} \in \mathbb{R}_{\geq 0}$  and

$$|\mathcal{O}(\phi)| \leq c \|\phi\|^\alpha + d \quad (\text{A13})$$

for  $c, \alpha, d \in \mathbb{R}_{\geq 0}$ . Then, the bias between the estimated observable  $\bar{\mathcal{O}}$  and the true value  $\mathcal{O}^*$  is given by

$$|\bar{\mathcal{O}} - \mathcal{O}^*| \leq \sum_{n \in N} \sup_{\phi \in \Delta_n} |\mathcal{O}(\phi)| \cdot \alpha_n, \quad (\text{A14})$$

where  $\bar{\mathcal{O}} = \mathbb{E}_{\phi \sim q_\theta} [\hat{\mathcal{O}}(\phi)]$ .

*Proof.* Let us assume the generic observable  $\mathcal{O}$  to be polynomially bounded

$$|\mathcal{O}(\phi)| \leq c \|\phi\|^\alpha + d. \quad (\text{A15})$$

When  $\widetilde{\text{supp}}(q_\theta) \subseteq \text{supp}(p)$ , Eq. (A14) can be bounded as

$$|\bar{\mathcal{O}} - \mathcal{O}^*| = \left| \int_{\widetilde{\text{supp}}(q_\theta)} \mathcal{D}[\phi] \mathcal{O}(\phi) \frac{p(\phi)}{q_\theta(\phi)} q_\theta(\phi) - \int \mathcal{D}[\phi] \mathcal{O}(\phi) p(\phi) \right| \quad (\text{A16})$$

$$= \left| \int (1 - 1_{\widetilde{\text{supp}}(q_\theta)}(\phi)) \mathcal{O}(\phi) p(\phi) \mathcal{D}[\phi] \right| \quad (\text{A17})$$

$$\leq \sum_{n \in \mathbb{N}} \left| \int_{\Delta_n} \mathcal{O}(\phi) (1 - 1_{\widetilde{\text{supp}}(q_\theta)}(\phi)) p(\phi) \mathcal{D}[\phi] \right| \quad (\text{A18})$$

$$\leq \sum_{n \in \mathbb{N}} \sup_{\phi \in \Delta_n} |\mathcal{O}(\phi)| \cdot \underbrace{\int_{\Delta_n} (1 - 1_{\widetilde{\text{supp}}(q_\theta)}(\phi)) p(\phi) \mathcal{D}[\phi]}_{=: \alpha_n}, \quad (\text{A19})$$

where  $\alpha_n$  is the coefficient defined for each bucket, i.e.,

$$\begin{aligned} \alpha_n &= \int_{\Delta_n} (1 - 1_{\widetilde{\text{supp}}(q_\theta)}(\phi)) p(\phi) \mathcal{D}[\phi] \\ &= \frac{1}{Z} \int_{\Delta_n} (1 - 1_{\widetilde{\text{supp}}(q_\theta)}(\phi)) e^{-S(\phi)} \mathcal{D}[\phi] \\ &= \frac{1}{Z} \left( \int_{\Delta_n} e^{-S(\phi)} \mathcal{D}[\phi] - \int_{\Delta_n \cap \widetilde{\text{supp}}(q_\theta)} e^{-S(\phi)} \mathcal{D}[\phi] \right), \end{aligned} \quad (\text{A20})$$

such that the following relation holds

$$\sum_{n \in \mathbb{N}} \alpha_n = \int (1 - 1_{\widetilde{\text{supp}}(q_\theta)}(\phi)) p(\phi) \mathcal{D}[\phi] = 1 - \bar{w}. \quad (\text{A21})$$

One, therefore, concludes that the bias is bounded by the following series

$$|\bar{\mathcal{O}} - \mathcal{O}^*| \leq \sum_{n \in \mathbb{N}} \sup_{\phi \in \Delta_n} |\mathcal{O}(\phi)| \cdot \alpha_n. \quad (\text{A22})$$

In order to obtain convergence of this series, we observe that polynomial boundedness of the observable implies

$$\sup_{\phi \in \Delta_n} |\mathcal{O}(\phi)| \leq cu_n^\alpha + d,$$

i.e.,  $\sup_{\phi \in \Delta_n} |\mathcal{O}(\phi)|$  grows polynomially in  $n$ . Similarly, from Eq. (A20), it follows

$$\alpha_n \leq \frac{1}{Z} \text{vol}(\Delta_n) e^{D_\vee + 1 - n}, \quad (\text{A23})$$

showing that  $\alpha_n$  decays exponentially in  $n$ . Thus,  $\sup_{\phi \in \Delta_n} |\mathcal{O}(\phi)| \cdot \alpha_n$  decays exponentially in  $n$ . This implies convergence of the series in the right-hand side of Eq. (A22).

It is important to note that each  $\alpha_n$  is weighted by the maximum of the observable on the corresponding volume

$\Delta_n$  which makes the bias inherently dependent on the observable while the  $\alpha_n$  coefficients are universal and represent the amount of mode dropping per bucket. ■

As an explicit example, let us consider  $S: \mathbb{R} \rightarrow \mathbb{R}$ ;  $S(\phi) = \phi^2$  and the observable  $\phi^3$ . This means that the true value of the observable is

$$\mathcal{O}^* = \mathbb{E}_{\phi \sim p}[\mathcal{O}(\phi)] = \frac{1}{Z} \int_{-\infty}^{\infty} \mathcal{D}[\phi] \phi^3 e^{-\phi^2} = 0. \quad (\text{A24})$$

If we assume a mode dropping model  $q$ , with  $q = 2p \cdot 1_{\mathbb{R}_{\geq 0}}$ , then

$$|\bar{\mathcal{O}} - \mathcal{O}^*| = \bar{\mathcal{O}} = \frac{2}{\sqrt{\pi}} \int_0^{\infty} \mathcal{D}[\phi] \phi^3 e^{-\phi^2} = \frac{1}{2}.$$

For the definition of the  $\Delta_n$ , we can choose  $D_\vee = 1$  and thus  $\Delta_n = [-\sqrt{n}, -\sqrt{n-1}] \cup [\sqrt{n-1}, \sqrt{n}]$  and obtain  $\alpha_n = 0.5(\text{erf}(\sqrt{n}) - \text{erf}(\sqrt{n-1}))$ . We note that since  $u_n = \sup\{|\phi|, \phi \in \Delta_n\} = \sqrt{n}$  and the observable is  $\phi^3$  it follows that

$$\sup\{|\mathcal{O}(\phi)|, \phi \in \Delta_n\} \leq (\sqrt{n})^3.$$

Hence, the bias is within the bound given by the theorem

$$\frac{1}{2} = |\bar{\mathcal{O}} - \mathcal{O}^*| \leq \sum_{n \in \mathbb{N}} n^{3/2} \alpha_n \approx 0.73.$$

## 6. Details on the numerical experiments

In the following, we summarize the details and setup used to perform the training of both forward- and reverse-KL normalizing flows as well as to estimate the HMC reference values. For our experiments, we focused on the action  $S(\phi)$  from Eq. (32) as a function of  $\kappa$  while keeping the coupling  $\lambda = 0.022$  fixed throughout the analysis.

### a. HMC sampling

For estimating the HMC reference values of the free energy density reported in Sec. V, we followed the same approach as in [2]. The idea is to discretize the trajectory in the  $\kappa$ -space (hopping parameter) into a sequence of finite steps where free energy differences can be calculated by running an HMC. The target free energy at an arbitrary point  $\kappa^*$  is then obtained by summing up all the free energy differences from  $\kappa = 0$  to  $\kappa = \kappa^*$ . We note that the higher the kappa values, the more steps one needs to make in order to discretize the trajectory up to the target point in parameter space. It follows that the uncertainty on the estimates also grows when  $\kappa^*$  increases as more terms are combined to obtain the free energy at the desired  $\kappa^*$ . Specifically, in our experiments, we chose a regular step-size between two subsequent  $\kappa$  values in the trajectory to be  $\Delta\kappa = 0.01$ . Such step-size is used to discretize the trajectory starting from  $\kappa = 0.0$  all the way up to the target. For instance, measuring the free energy density at  $\kappa^* = 0.3$  would therefore require thirty steps, hence 30 independent HMC chains. Each of these chains is initialized around the vacuum expectation value (vev), has an overrelaxation every 10 steps, and a total of 10k thermalization steps, i.e., discarded configuration updates, followed by 500k sampling steps. Those configurations, from the equilibrium distribution, are used to estimate the free energy difference at a single given point of the trajectory. The total number of HMC samples needed to estimate the free energy at an arbitrary point thus depends on  $\kappa^*$ . For instance, referring to the previous example, 30 chains with 500k steps each add up to 15M HMC samples.

### b. Reverse-KL flow

To train the reverse-KL normalizing flows we followed the same strategy presented in [2] with the same setup of hyperparameters. We used a batch size of 8k samples and a learning rate update according to the ReduceLRonPlateau scheduler of PyTorch with an initial learning rate of  $5 \times 10^{-4}$  and patience of 3k steps. The flows have the same number of coupling blocks and the same type of checkerboard partitioning discussed in [2]. Models were

trained for 700k steps in total and the last saved checkpoint is used for sampling. Every reverse-KL model was trained on two GPUs (in parallel), either P100 or A100 NVIDIA devices. Depending on the lattice volume and the model type the training took up to 50 hrs of wall time.

### c. Forward-KL flow

Training a forward-KL flow requires a different procedure which was outlined in Appendix A 1. For every flow model, we used 50M pre-sampled HMC configurations as input data. These were sampled in batches of 100 independent HMC chains each of which had 10k equilibration (discarded) and 500k sampling (stored) steps. The stored configurations from each chain in the batch were concatenated to generate the full training set.

At the stage of training, the 50M configurations are loaded in batches of 8k samples per iteration (training step). When the entire dataset is processed once, the full set of configurations is reshuffled and reused (as it is standard practice in deep learning) until the desired number of training iterations is reached. Again, forward-KL models were trained for 700k steps on two GPUs (in parallel), either P100 or A100 NVIDIA devices. Depending on the lattice volume and the model type the training took up to 55 hrs of wall time.

### d. Flow sampling

For sampling configurations from both forward- and reverse-KL normalizing flows we proceed as follows. In order to have a fair comparison with HMC one would need to sample as many configurations as those needed to integrate the trajectory in the hopping parameter space, as discussed in Appendix 6 a. However, for our flow estimates, we took only 1M configurations and used the estimators for mean and variance introduced in [2,24] and proposed in Sec. III A. Though 1M is in general a lower bound on the total amount of configurations used to compute HMC estimates, we empirically observed this was sufficient to obtain estimates with errors several orders of magnitude smaller than HMC. Therefore, we took this as a sufficient number of samples for comparing the two sampling approaches.

- 
- [1] M. S. Albergo, G. Kanwar, and P. E. Shanahan, *Phys. Rev. D* **100**, 034515 (2019).  
 [2] K. A. Nicoli, C. J. Anders, L. Funcke, T. Hartung, K. Jansen, P. Kessel, S. Nakajima, and P. Stornati, *Phys. Rev. Lett.* **126**, 032001 (2021).  
 [3] D. C. Hackett, C.-C. Hsieh, M. S. Albergo, D. Boyda, J.-W. Chen, K.-F. Chen, K. Cranmer, G. Kanwar, and P. E. Shanahan, [arXiv:2107.00734](https://arxiv.org/abs/2107.00734).

- [4] K. A. Nicoli, C. J. Anders, L. Funcke, T. Hartung, K. Jansen, P. Kessel, S. Nakajima, and P. Stornati, *Proc. Sci. LATTICE2021* (2022) 338.  
 [5] M. Gerdes, P. de Haan, C. Rainone, R. Bondesan, and M. C. Cheng, [arXiv:2207.00283](https://arxiv.org/abs/2207.00283).  
 [6] M. Caselle, E. Cellini, A. Nada, and M. Panero, *J. High Energy Phys.* **07** (2022) 015.

- [7] M. Caselle, E. Cellini, A. Nada, and M. Panero, *Proc. Sci. LATTICE2022* (2023) 005.
- [8] A. Singha, D. Chakrabarti, and V. Arora, *Phys. Rev. D* **107**, 014512 (2023).
- [9] D. Albandea, L. Del Debbio, P. Hernández, R. Kenway, J. M. Rossney, and A. Ramos, *Proc. Sci. LATTICE2022* (2023) 001.
- [10] J. M. Pawłowski and J. M. Urban, *Phys. Rev. D* **108**, 054511 (2023).
- [11] G. Kanwar, M. S. Albergo, D. Boyda, K. Cranmer, D. C. Hackett, S. Racanière, D. J. Rezende, and P. E. Shanahan, *Phys. Rev. Lett.* **125**, 121601 (2020).
- [12] J. Finkenrath, [arXiv:2201.02216](https://arxiv.org/abs/2201.02216).
- [13] D. Boyda, G. Kanwar, S. Racanière, D. J. Rezende, M. S. Albergo, K. Cranmer, D. C. Hackett, and P. E. Shanahan, *Phys. Rev. D* **103**, 074504 (2021).
- [14] M. Favoni, A. Ipp, D. I. Müller, and D. Schuh, *Phys. Rev. Lett.* **128**, 032003 (2022).
- [15] S. Bacchio, P. Kessel, S. Schaefer, and L. Vaitl, *Phys. Rev. D* **107**, L051504 (2023).
- [16] M. S. Albergo, G. Kanwar, S. Racanière, D. J. Rezende, J. M. Urban, D. Boyda, K. Cranmer, D. C. Hackett, and P. E. Shanahan, *Phys. Rev. D* **104**, 114507 (2021).
- [17] R. Abbott, M. S. Albergo, D. Boyda, K. Cranmer, D. C. Hackett, G. Kanwar, S. Racanière, D. J. Rezende, F. Romero-López, P. E. Shanahan *et al.*, *Phys. Rev. D* **106**, 074506 (2022).
- [18] M. Lüscher, *Commun. Math. Phys.* **293**, 899 (2010).
- [19] L. Del Debbio, J. M. Rossney, and M. Wilson, *Proc. Sci. LATTICE2021* (2022) 059.
- [20] L. Del Debbio, J. Marsh Rossney, and M. Wilson, *Phys. Rev. D* **104**, 094507 (2021).
- [21] P. de Haan, C. Rainone, M. Cheng, and R. Bondesan, [arXiv:2110.02673](https://arxiv.org/abs/2110.02673).
- [22] R. Abbott, M. S. Albergo, A. Botev, D. Boyda, K. Cranmer, D. C. Hackett, A. G. Matthews, S. Racanière, A. Razavi, D. J. Rezende, F. Romero-López, P. E. Shanahan, and J. M. Urban, *Eur. Phys. J. A* **59**, 257 (2023).
- [23] Y. Bengio, P. Simard, and P. Frasconi, *IEEE Trans. Neural Networks* **5**, 157 (1994).
- [24] K. A. Nicoli, S. Nakajima, N. Strodthoff, W. Samek, K.-R. Müller, and P. Kessel, *Phys. Rev. E* **101**, 023304 (2020).
- [25] M. Gabrié, G. M. Rotskoff, and E. Vanden-Eijnden, *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2109420119 (2022).
- [26] L. Grenioux, A. Durmus, É. Moulines, and M. Gabrié, in *Proceedings of the 40th International Conference on Machine Learning* (PMLR, 2023), 11698–11733.
- [27] L. Vaitl, K. A. Nicoli, S. Nakajima, and P. Kessel, in *Proceedings of the 39th International Conference on Machine Learning* (PMLR, Baltimore, Maryland, USA, 2022), Vol. 162, pp. 21945–21959.
- [28] L. Vaitl, K. A. Nicoli, S. Nakajima, and P. Kessel, *Mach. Learn.* **3**, 045006 (2022).
- [29] B. Máté and F. Fleuret, *Transactions on Machine Learning Research* (2023), <https://openreview.net/forum?id=TH6YrEcbth>.
- [30] L. I. Midgley, V. Stimper, G. N. Simm, B. Schölkopf, and J. M. Hernández-Lobato, in *the Eleventh International Conference on learning representations* (2023), <https://openreview.net/forum?id=XCTVFJwS9LJ>.
- [31] Other examples of such thermodynamic observables not directly accessible with HMC are, for instance, entropy and pressure.
- [32] M. Caselle, A. Nada, and M. Panero, *Phys. Rev. D* **98**, 054513 (2018).
- [33] D. Rezende and S. Mohamed, in *Proceedings of the 32nd International Conference on Machine Learning* (PMLR, Lille, France, 2015), Vol. 37, pp. 1530–1538.
- [34] I. Kobyzev, S. J. Prince, and M. A. Brubaker, *IEEE Trans. Pattern Anal. Mach. Intell.* **43**, 3964 (2020).
- [35] G. Papamakarios, E. Nalisnick, D. J. Rezende, S. Mohamed, and B. Lakshminarayanan, *J. Mach. Learn. Res.* **22**, 1 (2021).
- [36] L. Dinh, D. Krueger, and Y. Bengio, in *3rd International Conference on Learning Representations (Workshop Track Proceedings)* (2015).
- [37] L. Dinh, J. Sohl-Dickstein, and S. Bengio, in *5th International Conference on Learning Representations* (2017), <https://openreview.net/forum?id=HkpbH9lx>.
- [38] D. P. Kingma and P. Dhariwal, in *Advances in Neural Information Processing Systems* (Curran Associates, Montreal, 2018), Vol. 31, p. 10236–10245.
- [39] H. Zhu and R. Rohwer, *Information Geometric Measurements of Generalisation*, Working Paper (Aston University, Birmingham, UK, 1995).
- [40] T. Minka, *Divergence measures and message passing*, Technical Report No. MSR-TR-2005-173, Microsoft Research Cambridge, 2005.
- [41] S.-I. Amari, *Differential-Geometrical Methods in Statistics* (Springer, New York, 1985), Vol. 28, p. 1.
- [42] S. Kullback and R. A. Leibler, *Ann. Math. Stat.* **22**, 79 (1951).
- [43] M. D. Reid and R. C. Williamson, *J. Mach. Learn. Res.* **12**, 731 (2011).
- [44] T. van Erven and P. Harremoës, *IEEE Trans. Inf. Theory* **60**, 3797 (2014).
- [45] J. Lin, *IEEE Trans. Inf. Theory* **37**, 145 (1991).
- [46] A. Banerjee, S. Merugu, I. S. Dhillon, and J. Ghosh, *J. Mach. Learn. Res.* **6**, 1705 (2005).
- [47] F. Noé, S. Olsson, J. Köhler, and H. Wu, *Science* **365**, eaaw1147 (2019).
- [48] P. Wirmsberger, A. J. Ballard, G. Papamakarios, S. Abercrombie, S. Racanière, A. Pritzel, D. Jimenez Rezende, and C. Blundell, *J. Chem. Phys.* **153**, 144112 (2020).
- [49] J. Köhler, M. Invernizzi, P. de Haan, and F. Noé, in *Proceedings of the 40th International Conference on Machine Learning* (Honolulu, Hawaii, USA, 2023).
- [50] A. van den Oord, N. Kalchbrenner, and K. Kavukcuoglu, in *Proceedings of The 33rd International Conference on Machine Learning* (PMLR, New York, New York, USA, 2016), Vol. 48, pp. 1747–1756.
- [51] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, in *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, 2021* (2021).



- [52] J. Behrmann, P. Vicol, K.-C. Wang, R. Grosse, and J.-H. Jacobsen, in *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics*, Proceedings of Machine Learning Research (PMLR, 2021), Vol. 130, pp. 1792–1800.
- [53] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, Cambridge, Massachusetts, 2016) Chap. 3.1.3.
- [54] A. Agrawal, D. R. Sheldon, and J. Domke, in *Advances in Neural Information Processing Systems* (Curran Associates, Vancouver, 2020), Vol. 33, pp. 17358–17369.
- [55] A. K. Dhaka, A. Catalina, M. Welandawe, M. R. Andersen, J. H. Huggins, and A. Vehtari, in *International Conference on Machine Learning (Workshop Track Proceedings)* (2021).
- [56] A. K. Dhaka, A. Catalina, M. Welandawe, M. R. Andersen, J. Huggins, and A. Vehtari, in *Advances in Neural Information Processing Systems* (Curran Associates, 2021), Vol. 34, pp. 7787–7798.
- [57] H. Wu, J. Köhler, and F. Noe, in *Advances in Neural Information Processing Systems* (2020), Vol. 33, pp. 5933–5944.
- [58] To make the notation more compact, in the integrals, we drop the subscript in the effective supports of both  $q_\theta$  and  $p$ .
- [59] We take the 2-norm for the field configuration and drop the subscript for notation convenience. We implicitly assume  $\|\phi\| \equiv \|\phi\|_2$ .