Normalising Flows on Quotient Configuration Spaces: Theory of Orbit Lifting and Likelihood Equivalence

Manav Madan Rawal

SUPA, School of Physics and Astronomy, University of Glasgow, Glasgow, G12 8QQ, UK

November 29, 2025

Abstract

Normalizing flows offer a promising alternative to Markov-chain sampling for lattice gauge theories, but their performance is challenged by the large symmetry redundancies present in the configuration space. We investigate whether generative modelling can be improved by removing these redundancies and learning directly on symmetry-reduced spaces. To this end we introduce *quotient normalizing flows*, in which lattice configurations are mapped to canonical representatives of their translation or gauge orbits and the flow is trained on these quotient variables, with samples lifted back to the full space by random symmetry transformations.

We test this idea in 2D U(1) and SU(2) gauge theories for both translation and gauge quotients, comparing flows trained on the raw ensemble with flows trained on the quotient slice. Although quotient flows often achieve substantially lower training NLL on the canonical representatives, they do not improve, and in several cases severely degrade the sampling of gauge-invariant observables such as the plaquette energy and Polyakov-loop magnitude. In SU(2) the quotient parameterizations can even lead to unstable training and strong distortions of the observable distributions. These results indicate that likelihood optimization on a gauge-fixed slice does not reliably correspond to accurate modelling of the physical measure on the orbit space, and that naïve quotienting is insufficient without properly accounting for the geometry and induced volume form of the slice.

1 Introduction

Many physical and statistical models are defined on configuration spaces $\mathcal{X}$ that possess a nontrivial group of symmetries G. For example, Ising and ϕ^4 lattice field theories are invariant under a global $\mathbb{Z}_2$ spin or field flip, and translationally invariant systems are invariant under the discrete translation group of the lattice. In such cases the physically relevant probability distributions π on $\mathcal{X}$ are G-invariant in the sense that $\pi(h \cdot A) = \pi(A)$ for all measurable $A \subset \mathcal{X}$ and all $h \in G$, where $h \cdot x$ denotes the natural action of G on configurations.

Existing symmetry-aware generative models (equivariant neural networks, equivariant normalising flows, etc.) enforce such symmetries by imposing architectural constraints directly on maps $f : \mathcal{X} \to \mathcal{X}$ or $\mathcal{X} \to \mathbb{R}^k$; see, for example, group-equivariant convolutional networks [1] and equivariant/invariant flows for Euclidean data [2]. In the context of lattice field theory, flow-based generative models and equivariant flows have been used as surrogates or accelerators for Markov-chain Monte Carlo in a number of works, notably Refs. [3, 4]. In particular, equivariant flows are constructed so that $f(g \cdot x) = g \cdot f(x)$ layer-by-layer. However, all these approaches still define probability densities

on the *full* configuration space $\mathcal{X}$ and must therefore represent all symmetry-related copies of each mode of π.

In this document we describe and analyse a different paradigm: we build the flow directly on the *quotient space* $\mathcal{X}/G$ of group orbits. The model learns a density q_θ on $\mathcal{X}/G$ via a normalising flow f_θ [5, 6], and samples on $\mathcal{X}$ are obtained by a randomized orbit-lifting procedure which applies a uniformly random group element to a canonical representative of each orbit. We show that this construction yields (i) exact G-invariance of the lifted samples, (ii) universality for G-invariant target densities, and (iii) a precise equivalence between maximum-likelihood training on the quotient and on the original configuration space. The proofs given below are self-contained and do not rely on any external literature.

2 Group Actions, Quotients, and Canonical Representatives

2.1 Configuration space and group action

Let $(\mathcal{X}, \mathcal{B}_\mathcal{X})$ be a measurable space, which we will interpret as the configuration space of a lattice field theory or a similar model. For concreteness one may think of $\mathcal{X} = \mathbb{R}^N$ equipped with the Borel σ-algebra, but the development below does not depend on this choice.

Let G be a finite group acting measurably on $\mathcal{X}$ via a map

$$\alpha : G \times \mathcal{X} \longrightarrow \mathcal{X}, \qquad (g, x) \mapsto g \cdot x, \tag{1}$$

such that

$$e \cdot x = x, \qquad g_1 \cdot (g_2 \cdot x) = (g_1 g_2) \cdot x \tag{2}$$

for all $x \in \mathcal{X}$ and all $g_1, g_2 \in G$, where e denotes the identity element of G.

For each $x \in \mathcal{X}$ we define its *orbit*

$$[x] := Gx := \{g \cdot x : g \in G\}, \tag{3}$$

and we write $\mathcal{X}/G$ for the set of all orbits. The canonical projection

$$\pi : \mathcal{X} \longrightarrow \mathcal{X}/G, \qquad x \mapsto [x], \tag{4}$$

is surjective by construction. The quotient set $\mathcal{X}/G$ can be equipped with the smallest σ-algebra $\mathcal{B}_{\mathcal{X}/G}$ making π measurable.

We will assume throughout that the action of G on $\mathcal{X}$ is *essentially free* with respect to the reference measure on $\mathcal{X}$ (i.e. that configurations with non-trivial stabilisers form a set of measure zero). This is satisfied in typical lattice field theory models, where perfectly symmetric configurations are of measure zero.

2.2 Sections and canonical representatives

The quotient space $\mathcal{X}/G$ collects orbits of configurations. In order to interface with standard Euclidean normalising flows, it is convenient to have a way of selecting a canonical representative from each orbit and identifying $\mathcal{X}/G$ with some parameter space $\mathcal{Y}$.

Definition 2.1 (Section and canonical representative)**.** A (measurable) *section* of the quotient map π is a map

$$s : \mathcal{X}/G \longrightarrow \mathcal{X} \tag{5}$$

such that $s([x]) \in [x]$ for all $[x] \in \mathcal{X}/G$. Given $x \in \mathcal{X}$, we define the *canonical representative* associated with x as

$$x^{\text{can}} := s(\pi(x)) \in [x]. \tag{6}$$

In lattice models this section can be implemented by a deterministic "canonicalisation" rule. Examples include:

- For a global $\mathbb{Z}_2$ symmetry, choose the representative with non-negative total magnetisation.
- For a translation symmetry, translate the configuration so that the site with maximal $|\phi_i|$ lies at a fixed reference position.

We will write $\mathcal{Y}$ for a measurable space that we identify measurably with $\mathcal{X}/G$; for simplicity of notation we may write $\mathcal{Y} \cong \mathcal{X}/G$.

2.3 Target distributions and invariance

Definition 2.2 (G-invariant distribution)**.** A probability measure π on $(\mathcal{X}, \mathcal{B}_{\mathcal{X}})$ is said to be *G-invariant* if

$$\pi(h \cdot A) = \pi(A) \quad \text{for all } A \in \mathcal{B}_{\mathcal{X}},\ h \in G, \tag{7}$$

where $h \cdot A := \{h \cdot x : x \in A\}$.

We are particularly interested in G-invariant distributions because they arise naturally as Boltzmann distributions of lattice field theories with discrete symmetries: for example, if $S : \mathcal{X} \to \mathbb{R}$ is an action satisfying $S(g \cdot x) = S(x)$, then

$$\pi(\mathrm{d}x) \propto e^{-S(x)}\,\mathrm{d}x \tag{8}$$

is G-invariant.

Definition 2.3 (Push-forward to the quotient)**.** Let π be a G-invariant probability measure on $\mathcal{X}$. The push-forward of π to the quotient space $(\mathcal{X}/G, \mathcal{B}_{\mathcal{X}/G})$ is the measure $\pi_{\mathcal{Y}}$ defined by

$$\pi_{\mathcal{Y}}(B) := \pi(\pi^{-1}(B)), \quad B \in \mathcal{B}_{\mathcal{X}/G}. \tag{9}$$

If π has a density (also denoted π) with respect to some reference measure μ on $\mathcal{X}$, then under mild assumptions $\pi_{\mathcal{Y}}$ has a density (denoted $\pi_{\mathcal{Y}}$) with respect to a suitable reference measure ν on $\mathcal{X}/G$.

Intuitively, $\pi_{\mathcal{Y}}$ describes the distribution of orbits $[x]$ when $x \sim \pi$.

3 Normalising Flows on the Quotient and Orbit Lifting

3.1 Normalising flow on the quotient space

Let $(\mathcal{Z}, \mathcal{B}_{\mathcal{Z}})$ be a measurable base space, typically $\mathcal{Z} = \mathbb{R}^d$, equipped with a tractable base density $r(z)$ (e.g. a standard multivariate Gaussian).

Definition 3.1 (Normalising flow on $\mathcal{X}/G$)**.** A *quotient-space flow* is a measurable bijection

$$f_\theta : \mathcal{Z} \longrightarrow \mathcal{Y} \cong \mathcal{X}/G \tag{10}$$

with a tractable Jacobian determinant, parametrised by $\theta \in \Theta$. It induces a family of densities $\{q_\theta\}_{\theta\in\Theta}$ on $\mathcal{Y}$ via the change-of-variables formula

$$q_\theta(y) = r\big(f_\theta^{-1}(y)\big)\left|\det J_{f_\theta^{-1}}(y)\right| \tag{11}$$

in the usual way for normalising flows [5, 6].

In practice, f_θ is a composition of simple invertible maps (affine coupling layers, invertible residual blocks, etc.) and q_θ is a normalising flow density on the quotient space.

3.2 Orbit-lifting sampler on the configuration space

The densities q_θ live on $\mathcal{Y} \cong \mathcal{X}/G$, not on the original configuration space $\mathcal{X}$. To obtain samples in $\mathcal{X}$, we define an orbit-lifting procedure.

Definition 3.2 (Orbit lifting)**.** Let $U(G)$ denote the uniform distribution on the finite group G. Given θ, the *orbit-lifting sampler* on $\mathcal{X}$ is defined as:

1. Sample $z \sim r$ and set $y = f_\theta(z) \in \mathcal{Y}$.
2. Compute the canonical representative $x_0 = s(y) \in \mathcal{X}$.
3. Sample $g \sim U(G)$ independently of z.
4. Output $x = g \cdot x_0$.

We denote the law of x by p_θ.

By construction, p_θ is a probability measure on $(\mathcal{X}, \mathcal{B}_\mathcal{X})$ which depends on θ only via q_θ.

4 Theoretical Properties

We now establish three basic properties of the quotient-space flow with orbit lifting: (i) G-invariance of the lifted distribution, (ii) universality on G-invariant densities, and (iii) likelihood equivalence between training on the quotient and on the full configuration space. These results may be viewed as a probabilistic counterpart of classical symmetry-reduction and slice constructions in dynamical systems (e.g. reconstruction equations and Karhunen-Loève expansions for systems with symmetry [7]).

4.1 G-invariance of the lifted distribution

Theorem 4.1 (G-invariance of the lifted model)**.** *Let p_θ be the probability measure on $(\mathcal{X}, \mathcal{B}_\mathcal{X})$ induced by the orbit-lifting sampler for a fixed θ. Then p_θ is G-invariant, i.e. for all $h \in G$ and all $A \in \mathcal{B}_\mathcal{X}$,*

$$p_\theta(h \cdot A) = p_\theta(A), \tag{12}$$

where $h \cdot A := \{h \cdot x : x \in A\}$.

Proof. Fix $h \in G$ and $A \in \mathcal{B}_\mathcal{X}$. By definition of the orbit-lifting sampler, a random variable x with law p_θ is given by

$$x = g \cdot s(y), \tag{13}$$

where $y \sim q_\theta$ and $g \sim U(G)$ are independent.

We compute

$$p_\theta(h \cdot A) = \mathbb{P}(x \in h \cdot A) \tag{14}$$
$$= \mathbb{P}(h^{-1} \cdot x \in A) \tag{15}$$
$$= \mathbb{P}\big(h^{-1}g \cdot s(y) \in A\big). \tag{16}$$

Consider the map $\varphi_h : G \to G$ defined by $\varphi_h(g) = h^{-1}g$. Since G is a group, φ_h is a bijection, and if $g \sim U(G)$ is uniform, then so is $h^{-1}g$. Thus the joint distribution of $(h^{-1}g, y)$ is identical to that of (g, y), and we may replace $h^{-1}g$ by a fresh uniform draw g' without changing the value of the probability:

$$p_\theta(h \cdot A) = \mathbb{P}\big(h^{-1}g \cdot s(y) \in A\big) \tag{17}$$
$$= \mathbb{P}\big(g' \cdot s(y) \in A\big) \tag{18}$$
$$= p_\theta(A), \tag{19}$$

where $g' \sim U(G)$ is independent of y.

Since h and A were arbitrary, this shows that p_θ is G-invariant. □

This theorem shows that symmetry of the lifted model is guaranteed purely by the structure of the lifting procedure, without any symmetry constraints on the flow f_θ itself, in contrast to explicitly equivariant architectures [1, 2].

4.2 Universality on invariant densities

We next show that universality of the quotient-space flow on $\mathcal{Y}$ implies universality of the lifted family on the space of G-invariant densities on $\mathcal{X}$.

Definition 4.2 (Invariant densities)**.** Let μ be a reference measure on $(\mathcal{X}, \mathcal{B}_\mathcal{X})$, for example Lebesgue measure. We write $\mathcal{P}^G(\mathcal{X})$ for the set of all densities p with respect to μ such that

$$p(h \cdot x) = p(x) \quad \text{for } \mu\text{-almost every } x \in \mathcal{X} \text{ and all } h \in G. \tag{20}$$

We also equip $\mathcal{Y} \cong \mathcal{X}/G$ with a reference measure ν (e.g. the push-forward of μ under π) so that densities q_θ on $\mathcal{Y}$ are well-defined.

Theorem 4.3 (Universality on G-invariant distributions)**.** *Suppose the family $\{q_\theta : \theta \in \Theta\}$ is universal on $\mathcal{Y}$ in the sense that for any density $q^\star$ on $\mathcal{Y}$ and any $\varepsilon > 0$ there exists $\theta \in \Theta$ such that*

$$\|q_\theta - q^\star\|_{L^1(\mathcal{Y})} < \varepsilon.$$

Then the lifted family $\{p_\theta\}$ is universal for $\mathcal{P}^G(\mathcal{X})$: for any $p^\star \in \mathcal{P}^G(\mathcal{X})$ and any $\varepsilon > 0$ there exists $\theta \in \Theta$ such that

$$\|p_\theta - p^\star\|_{L^1(\mathcal{X})} < \varepsilon.$$

Proof. Let $p^\star \in \mathcal{P}^G(\mathcal{X})$ be arbitrary. Because the action of G is essentially free, each orbit $[x]$ has cardinality $|G|$ for μ-almost every x, and $p^\star$ is constant along each orbit.

Step 1: Factorisation through the quotient. Define a measure $\lambda^\star$ on $\mathcal{Y}$ by push-forward:

$$\lambda^\star(B) := \int_{\pi^{-1}(B)} p^\star(x)\,\mathrm{d}\mu(x), \quad B \in \mathcal{B}_\mathcal{Y}. \tag{21}$$

Since $p^\star$ is invariant along orbits and almost all orbits have size $|G|$, $\lambda^\star$ has a density $q^\star$ with respect to ν and

$$p^\star(x) = \frac{1}{|G|}\, q^\star(\pi(x)) \tag{22}$$

for μ-almost every $x \in \mathcal{X}$.

Step 2: Approximate $q^\star$ on the quotient. By universality on $\mathcal{Y}$, for any $\varepsilon > 0$ there exists $\theta \in \Theta$ such that

$$\|q_\theta - q^\star\|_{L^1(\mathcal{Y})} < \varepsilon. \tag{23}$$

Step 3: Define the lifted model. Define a density p_θ on $\mathcal{X}$ by

$$p_\theta(x) := \frac{1}{|G|}\, q_\theta(\pi(x)). \tag{24}$$

By construction, p_θ is G-invariant and the law of the orbit-lifting sampler coincides with the measure having density p_θ.

Step 4: Compare p_θ and $p^\star$ in $L^1(\mathcal{X})$. We compute

$$\begin{aligned}
\|p_\theta - p^\star\|_{L^1(\mathcal{X})} &= \int_\mathcal{X} |p_\theta(x) - p^\star(x)|\,\mathrm{d}\mu(x) && (25)\\
&= \int_\mathcal{X} \left| \frac{1}{|G|} q_\theta(\pi(x)) - \frac{1}{|G|} q^\star(\pi(x)) \right| \mathrm{d}\mu(x) && (26)\\
&= \frac{1}{|G|} \int_\mathcal{X} |q_\theta(\pi(x)) - q^\star(\pi(x))|\,\mathrm{d}\mu(x). && (27)
\end{aligned}$$

Now use the fact that π is a $|G|$-to-1 map almost everywhere: for any integrable function φ on $\mathcal{Y}$,

$$\int_\mathcal{X} \varphi(\pi(x))\,\mathrm{d}\mu(x) = |G| \int_\mathcal{Y} \varphi(y)\,\mathrm{d}\nu(y). \tag{28}$$

Applying this with $\varphi(y) = |q_\theta(y) - q^\star(y)|$ yields

$$\begin{aligned}
\|p_\theta - p^\star\|_{L^1(\mathcal{X})} &= \frac{1}{|G|} \int_\mathcal{X} |q_\theta(\pi(x)) - q^\star(\pi(x))|\,\mathrm{d}\mu(x) && (29)\\
&= \frac{1}{|G|} \cdot |G| \int_\mathcal{Y} |q_\theta(y) - q^\star(y)|\,\mathrm{d}\nu(y) && (30)\\
&= \|q_\theta - q^\star\|_{L^1(\mathcal{Y})}. && (31)
\end{aligned}$$

By assumption $\|q_\theta - q^\star\|_{L^1(\mathcal{Y})} < \varepsilon$, and hence $\|p_\theta - p^\star\|_{L^1(\mathcal{X})} < \varepsilon$.

Since $p^\star \in \mathcal{P}^G(\mathcal{X})$ and $\varepsilon > 0$ were arbitrary, the result follows. □

This theorem shows that if the quotient-space flow is expressive enough to approximate arbitrary densities on the quotient, then the lifted family of models is expressive enough to approximate any G-invariant density on the original configuration space. In particular, standard universality arguments for normalising flows [5] immediately transfer to the G-invariant setting when the model is formulated on X/G.

4.3 Likelihood equivalence via canonicalisation

We finally show that maximum-likelihood training on canonical representatives in the quotient space is equivalent, up to an additive constant, to maximum-likelihood training of the lifted model on the full configuration space.

Theorem 4.4 (Likelihood equivalence)**.** *Let π be a G-invariant target distribution on $\mathcal{X}$ with density (also denoted π) with respect to μ. Let $\pi_{\mathcal{Y}}$ be the push-forward of π to $\mathcal{Y}$, with density (also denoted $\pi_{\mathcal{Y}}$) with respect to ν. Assume the action of G is essentially free.*

Let q_θ be the quotient-space flow density on $\mathcal{Y}$ and let p_θ be the corresponding lifted density on $\mathcal{X}$,

$$p_\theta(x) = \frac{1}{|G|} q_\theta(\pi(x)). \tag{32}$$

Then

$$\mathbb{E}_{x\sim\pi}[\log p_\theta(x)] = \mathbb{E}_{y\sim\pi_{\mathcal{Y}}}[\log q_\theta(y)] - \log|G| \tag{33}$$

for all $\theta \in \Theta$.

Proof. As in the proof of Theorem 4.3, G-invariance of π and essential freeness of the action imply that there exists a quotient-space density $\pi_{\mathcal{Y}}$ such that

$$\pi(x) = \frac{1}{|G|}\, \pi_{\mathcal{Y}}(\pi(x)) \tag{34}$$

for μ-almost every x.

We compute the expected log-likelihood of p_θ under π:

$$\mathbb{E}_{x\sim\pi}[\log p_\theta(x)] = \int_{\mathcal{X}} \pi(x)\ \log p_\theta(x)\, \mathrm{d}\mu(x) \tag{35}$$

$$= \int_{\mathcal{X}} \frac{1}{|G|}\, \pi_{\mathcal{Y}}(\pi(x))\ \log\!\left(\frac{1}{|G|} q_\theta(\pi(x))\right) \mathrm{d}\mu(x). \tag{36}$$

Using $\log\big(\frac{1}{|G|} q_\theta(\pi(x))\big) = \log q_\theta(\pi(x)) - \log|G|$, we separate the integral into two terms:

$$\mathbb{E}_{x\sim\pi}[\log p_\theta(x)] = \int_{\mathcal{X}} \frac{1}{|G|}\, \pi_{\mathcal{Y}}(\pi(x))\ \log q_\theta(\pi(x))\, \mathrm{d}\mu(x) \tag{37}$$

$$- \int_{\mathcal{X}} \frac{1}{|G|}\, \pi_{\mathcal{Y}}(\pi(x))\ \log|G|\, \mathrm{d}\mu(x) \tag{38}$$

$$=: I_1 - I_2. \tag{39}$$

For the second term,

$$I_2 = \log|G| \int_{\mathcal{X}} \frac{1}{|G|}\, \pi_{\mathcal{Y}}(\pi(x))\, \mathrm{d}\mu(x). \tag{40}$$

By the same $|G|$-to-1 argument as before,

$$\int_{\mathcal{X}} \frac{1}{|G|}\, \pi_{\mathcal{Y}}(\pi(x))\, \mathrm{d}\mu(x) = \int_{\mathcal{Y}} \pi_{\mathcal{Y}}(y)\, \mathrm{d}\nu(y) = 1, \tag{41}$$

since $\pi_{\mathcal{Y}}$ is a probability density on $\mathcal{Y}$. Thus

$$I_2 = \log|G|. \tag{42}$$

For the first term,

$$I_1 = \int_{\mathcal{X}} \frac{1}{|G|}\, \pi_{\mathcal{Y}}(\pi(x))\, \log q_\theta(\pi(x))\, \mathrm{d}\mu(x). \tag{43}$$

Using the change-of-variables identity

$$\int_{\mathcal{X}} \frac{1}{|G|}\, \varphi(\pi(x))\, \mathrm{d}\mu(x) = \int_{\mathcal{Y}} \varphi(y)\, \mathrm{d}\nu(y) \tag{44}$$

with $\varphi(y) = \pi_{\mathcal{Y}}(y) \log q_\theta(y)$, we obtain

$$I_1 = \int_{\mathcal{Y}} \pi_{\mathcal{Y}}(y)\, \log q_\theta(y)\, \mathrm{d}\nu(y) \tag{45}$$

$$= \mathbb{E}_{y\sim\pi_{\mathcal{Y}}}[\log q_\theta(y)]. \tag{46}$$

Combining I_1 and I_2 gives

$$\mathbb{E}_{x\sim\pi}[\log p_\theta(x)] = \mathbb{E}_{y\sim\pi_{\mathcal{Y}}}[\log q_\theta(y)] - \log|G|, \tag{47}$$

as claimed. □

[Maximum-likelihood training on canonical representatives] Let $x_1, \ldots, x_n$ be i.i.d. samples from π on $\mathcal{X}$ and let $y_i = \pi(x_i)$ denote their images in the quotient (or, equivalently, their canonical representatives). Define the empirical objectives

$$\mathcal{L}_X(\theta) := \frac{1}{n} \sum_{i=1}^{n} \log p_\theta(x_i), \tag{48}$$

$$\mathcal{L}_Y(\theta) := \frac{1}{n} \sum_{i=1}^{n} \log q_\theta(y_i). \tag{49}$$

Then for all θ,

$$\mathcal{L}_X(\theta) = \mathcal{L}_Y(\theta) - \log|G|, \tag{50}$$

and in particular, any maximiser of $\mathcal{L}_Y(\theta)$ is also a maximiser of $\mathcal{L}_X(\theta)$.

Proof. The empirical objectives are Monte Carlo estimators of the corresponding expectations in Theorem 4.4, and the relationship between them is exactly the same:

$$\frac{1}{n} \sum_{i=1}^{n} \log p_\theta(x_i) = \frac{1}{n} \sum_{i=1}^{n} \log q_\theta(y_i) - \log|G|. \tag{51}$$

The constant $-\log|G|$ is independent of θ, so the sets of maximisers of $\mathcal{L}_X$ and $\mathcal{L}_Y$ coincide. □

This corollary formalises the key practical consequence of our analysis: maximum-likelihood training can be performed entirely in the quotient space on canonicalised data, and the resulting parameters θ are optimal for the lifted model on the full configuration space as well. This provides a clean theoretical justification for common practices in symmetry-reduced simulations and "method of slices" constructions [7], now specialised to normalising flows.

5 Experiments

In this section we empirically study normalising flows on quotient configuration spaces, following the theoretical framework developed earlier in the paper. All experiments use ensembles generated by Hybrid Monte Carlo (HMC) as a reference "ground truth" [8], and compare three modelling strategies:

1. **Flow on the full configuration space** X ("NF on X").
2. **Flow on a partially quotiented space** X/H for a subgroup $H \leq G$ (e.g. only global Z_2).
3. **Flow on the full quotient** X/G with respect to a discrete symmetry group G (e.g. $Z_2 \times \Lambda$ for sign flip and lattice translations), followed by randomized orbit lifting.

All flows share the same RealNVP architecture and hyperparameters; only the input data (raw versus canonicalised representatives) and the lifting rule differ. This isolates the effect of quotienting from changes in model capacity. Our overall setup is closely related in spirit to previous flow-based studies of lattice field theory [3, 4], but we emphasize that here the symmetry is implemented entirely at the level of the probability space X/G rather than through explicitly equivariant layers.

5.1 Common experimental setup

Reference ensembles. For each model and parameter choice we generate an HMC ensemble of size N_{HMC} with acceptance rates $\gtrsim 0.97$, using leapfrog integration with step size $\varepsilon = 0.05$-0.1 and 10 leapfrog steps per trajectory, following standard practice in lattice simulations [8]. We discard an initial burn-in region and thin the chain to reduce autocorrelation; the remaining configurations are treated as approximately independent samples from the target Boltzmann distribution.

Normalising flows. Unless stated otherwise, flows are RealNVP models on $\mathbb{R}^D$ with:

- D equal to the number of real degrees of freedom in the configuration (e.g. $D = L$ for 1D fields, $D = L^2$ for 2D scalar fields, $D = 2L^2$ for 2D U(1) gauge fields with two link directions).
- 10 affine coupling layers with simple alternating binary masks, as in the original RealNVP construction [6].
- Fully connected MLP conditioners with two hidden layers of width 512 and ReLU activations.
- Tanh-squashed log-scale outputs with a global scale factor (0.8) to stabilise training.

We train using Adam with learning rate 10^{-3}, batch size 256, and 200-300 epochs depending on the experiment, following common practice in flow-based density estimation [5, 3]. For each epoch we report the mean negative log-likelihood (NLL) per configuration on the training data.

Canonicalisation and orbit lifting. For a configuration $x \in X$ and symmetry group G, we implement a measurable canonicalisation map

$$\mathrm{can} : X \to X, \qquad x \mapsto x_{\mathrm{can}}, \tag{52}$$

such that x_{can} is a deterministic representative of the orbit $[x]$. The training data for a quotient flow on X/G are the canonicalised configurations $\{x_{\mathrm{can}}\}$, flattened into $\mathbb{R}^D$.

At sampling time we draw y from the quotient flow, interpret it as a canonical representative, and then apply a random group element $g \in G$ to obtain a configuration

$$x = g \cdot y_{\text{can}} \in X, \tag{53}$$

implementing the orbit-lifting sampler described in the theoretical part of the paper. This guarantees that the induced law on X is G-invariant whenever the base flow on X/G is well-defined.

Evaluation metrics. For scalar field theories we measure:

- $\langle \phi^2 \rangle$, $\langle \phi^4 \rangle$,
- the magnetisation $M = \frac{1}{V} \sum_i \phi_i$,
- the absolute magnetisation $\langle |M| \rangle$,
- the Binder cumulant $U_4 = 1 - \frac{\langle M^4 \rangle}{3 \langle M^2 \rangle^2}$.

For U(1) gauge theories we measure:

- average plaquette energy $\langle E_p \rangle = \langle 1 - \cos \theta_p \rangle$,
- the absolute Polyakov loop $\langle |P| \rangle$ (in 1D along the ring, in 2D along a chosen direction).

We compare NLL curves, histograms of $|M|$ or $|P|$, and the deviation of these observables from HMC, in line with previous work on flow-based lattice sampling [3, 4].

5.2 Toy 1D double-well potential

As a didactic example we consider a single real variable $x \in \mathbb{R}$ with double-well potential

$$V(x) = \lambda \left(x^2 - a^2\right)^2, \tag{54}$$

which is invariant under the global Z_2 symmetry $x \mapsto -x$. We compare:

- a flow on $X = \mathbb{R}$;
- a flow on the quotient $X/Z_2 \cong [0, \infty)$, using the canonical representative $y = |x|$ and orbit lifting via a random sign.

The target density is strongly bimodal in x, but unimodal in y. This experiment serves as a 1D caricature of the lattice models studied below and allows us to visualise clearly how quotienting turns a multi-modal invariant distribution on X into a simpler distribution on X/G.

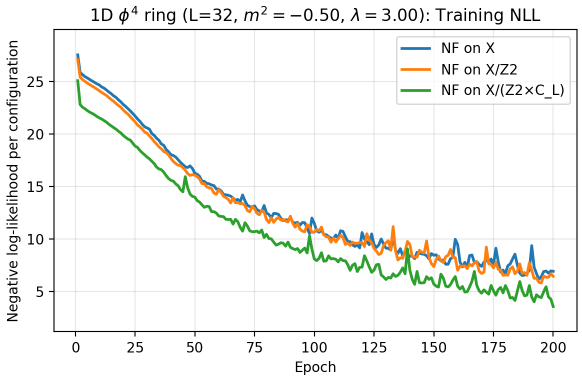
5.3 1D ϕ^4 lattice theory on a ring

We next consider a 1D ϕ^4 lattice field theory on a periodic ring of length $L = 32$ with action

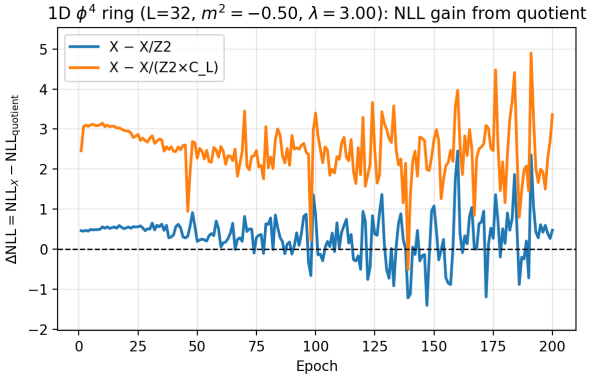
$$S[\phi] = \sum_{i=1}^{L} \left[\frac{1}{2} \big(\phi_i - \phi_{i+1}\big)^2 + \frac{m^2}{2} \phi_i^2 + \frac{\lambda}{4} \phi_i^4 \right], \tag{55}$$

with parameters chosen in the broken-symmetry phase (e.g. $m^2 = -0.5$, $\lambda = 3.0$). The model has:

- a global Z_2 field-flip symmetry $\phi_i \mapsto -\phi_i$,



(a) Training NLL for flows on X, X/Z_2 and $X/(Z_2 \times C_L)$.



(b) NLL gap relative to the flow on X.

Figure 1: 1D ϕ^4 on a ring ($L = 32$, $m^2 = -0.5$, $\lambda = 3.0$). Quotient flows on X/Z_2 and $X/(Z_2 \times C_L)$ converge faster and reach lower NLL than the flow on X.

- lattice translations $\phi_i \mapsto \phi_{i+k}$ (indices modulo L).

We compare three flows:

1. **NF on X.** Trained on raw configurations $\phi \in \mathbb{R}^L$ (flattened).
2. **NF on X/Z_2.** Canonicalisation by global sign:

$$\phi_{\text{can}} = \begin{cases} \phi, & M(\phi) \geq 0, \\ -\phi, & M(\phi) < 0, \end{cases} \qquad M(\phi) = \frac{1}{L} \sum_i \phi_i. \tag{56}$$

   At sampling, we draw ϕ_{can} from the flow and then flip the sign with probability 1/2.

3. **NF on $X/(Z_2 \times C_L)$.** Canonicalisation by sign and translation:
   (a) First fix the sign so that $M(\phi) \geq 0$ using the rule above.
   (b) Then locate the index of the maximal absolute field value,

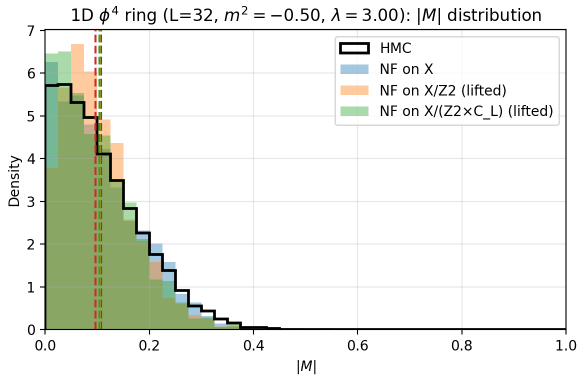
$$i^\star(\phi) = \arg\max_i |\phi_i|, \tag{57}$$

   and translate the configuration so that $i^\star$ is mapped to a fixed reference site (e.g. site $i = 1$).

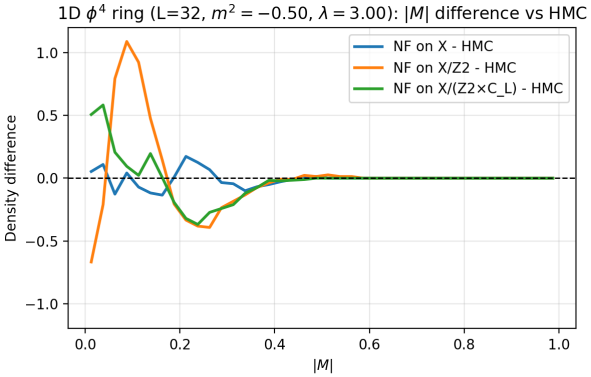
   The quotient flow is trained on these canonical representatives. At sampling, we draw a canonical configuration, apply a random cyclic shift, and flip the sign with probability 1/2.

All three flows use the same RealNVP architecture on $\mathbb{R}^{32}$; the symmetry information enters solely via canonicalisation and random orbit lifting.

Figures 1 and 2 summarise the 1D ϕ^4 results. Quotient flows on X/Z_2 and $X/(Z_2 \times C_L)$ converge faster and reach a visibly lower NLL than the flow on X, while their $|M|$ histograms track the HMC reference more closely. This supports the interpretation that quotienting removes redundant symmetry-related modes and makes the underlying density genuinely easier to model, rather than merely reshuffling error.



(a) Distribution of $|M|$ from HMC and the three flows.



(b) Bin-wise difference w.r.t. HMC, $\Delta p(|M|)$.

Figure 2: 1D ϕ^4 on a ring: magnetisation statistics. All flows reproduce the $|M|$ distribution reasonably well; the quotient flows track HMC slightly more closely, consistent with their lower NLL.

5.4 2D ϕ^4 lattice theory

We now move to a 2D scalar model on an $L \times L$ square lattice with periodic boundary conditions ($L = 16$) and action

$$S[\phi] = \sum_x \left[\frac{1}{2} \sum_{\mu=1}^{2} (\phi_x - \phi_{x+\hat{\mu}})^2 + \frac{m^2}{2} \phi_x^2 + \frac{\lambda}{4} \phi_x^4 \right], \tag{58}$$

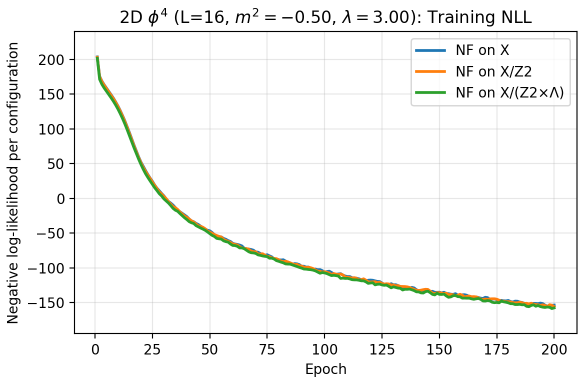
again in the broken-symmetry regime. The symmetry group is $G = Z_2 \times \Lambda$, where Z_2 is the global sign flip and Λ denotes the lattice translation group.

We consider:

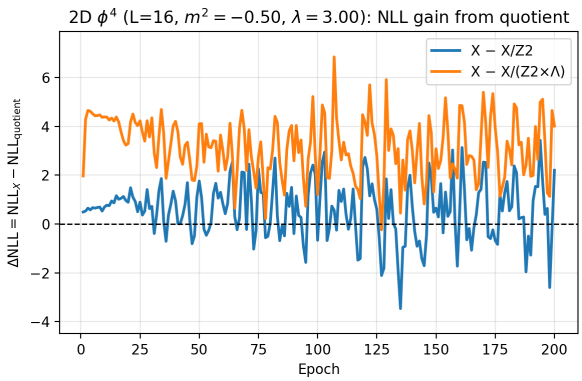
- **NF on X.** Fully connected RealNVP on $\mathbb{R}^{L^2}$ trained on raw field configurations.
- **NF on X/Z_2.** Canonicalisation by global sign as in 1D: if the spatial average $M(\phi)$ is negative, we flip the configuration.
- **NF on $X/(Z_2 \times \Lambda)$.** Canonicalisation by sign and translation:
  1. Enforce $M(\phi) \geq 0$ by a global flip if necessary.
  2. Find the site of maximal absolute field,
  $$x^\star(\phi) = \arg\max_x |\phi_x|, \tag{59}$$
  3. Translate the configuration so that $x^\star$ is mapped to a fixed reference site (e.g. the origin $(0, 0)$).

Again, all flows share the same architecture on $\mathbb{R}^{L^2}$; we only change the preprocessing and the group used for orbit lifting.

In 2D ϕ^4, Fig. 3 shows that quotient flows on X/Z_2 and $X/(Z_2 \times \Lambda)$ still enjoy a consistent NLL advantage over the flow on X, although the gap is reduced compared to the 1D case. The corresponding $|M|$ histograms in Fig. 4 reveal that all three models reproduce the HMC magnetisation distribution to similar accuracy, indicating that the benefit of quotienting is present but quantitatively modest at this lattice size and parameter choice.

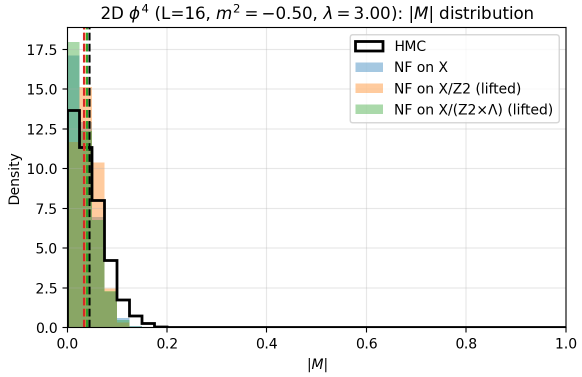


(a) Training NLL on X, X/Z_2 and $X/(Z_2 \times \Lambda)$.

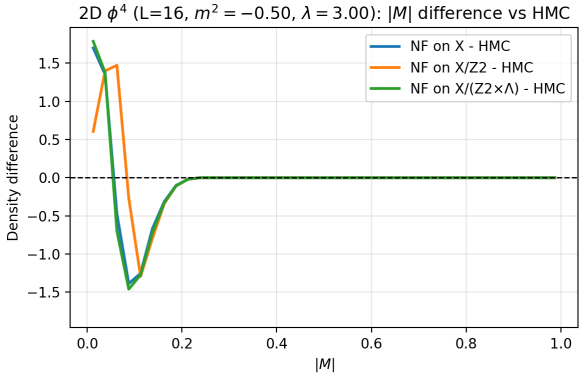


(b) NLL gap relative to the flow on X.

Figure 3: 2D ϕ^4 on an $L \times L$ lattice ($L = 16$). Quotient flows retain a modest NLL advantage over the flow on X, but the gap is smaller than in 1D.



(a) Histogram of $|M|$ for HMC and flows.



(b) Difference $\Delta p(|M|)$ relative to HMC.

Figure 4: 2D ϕ^4: magnetisation statistics. All three flows deviate from HMC by comparable amounts; quotienting is mildly beneficial but not dramatic at this lattice size.

5.5 1D U(1) gauge theory on a ring

To test the framework on compact variables, we study a 1D U(1) gauge model defined on a ring of length $L = 32$ with link angles $\theta_i \in (-\pi, \pi]$. The action is

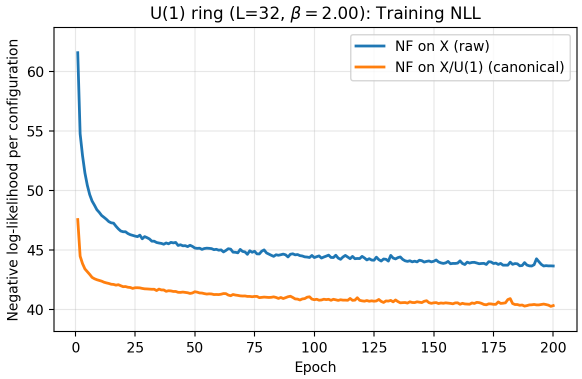
$$S[\theta] = -\beta \sum_{i=1}^{L} \cos\big(\theta_i - \theta_{i+1}\big), \tag{60}$$

and observables are constructed from the "Polyakov loop" along the ring,

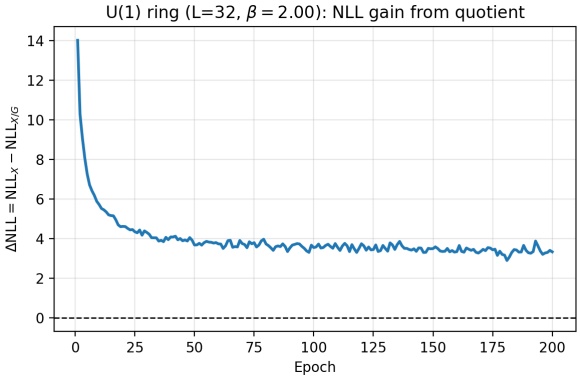
$$P = \frac{1}{L} \sum_{i=1}^{L} \exp\big(i\theta_i\big), \qquad |P| \in [0, 1]. \tag{61}$$

This model has:

- a global shift symmetry $\theta_i \mapsto \theta_i + \alpha \pmod{2\pi}$,
- lattice translations $i \mapsto i + k$.

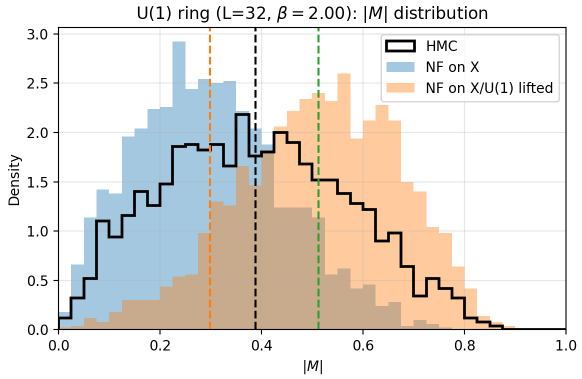


(a) Training NLL for flows on X and X/G (global U(1) shift).

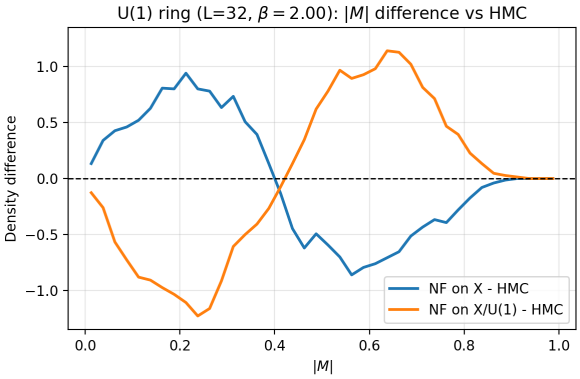


(b) NLL gap between the two flows.

Figure 5: 1D U(1) ring ($L = 32$, $\beta = 2.0$): quotienting the global U(1) phase yields a simpler density and a lower NLL than modelling the raw link angles.



(a) Histogram of $|P|$ from HMC and flows.



(b) Difference $\Delta p(|P|)$ relative to HMC.

Figure 6: 1D U(1) ring: Polyakov loop statistics. Both flows reproduce $|P|$ reasonably well, with the quotient flow performing at least as well as the baseline.

In this section we focus on quotienting the global U(1) shift. We work with wrapped angles in $(-\pi, \pi]$ and define:

- **NF on** X: a flow in $\mathbb{R}^L$ on raw angles $\{\theta_i\}$.
- **NF on** X/G for $G = \mathrm{U}(1)_{\text{global}}$: canonicalisation by subtracting the mean phase (aligning one link to a reference), followed by orbit lifting via a random global shift.

The setup is deliberately minimal and uses the same RealNVP backbone and optimisation scheme (Adam [9]) as in the scalar experiments, so that the effect of quotienting can be cleanly compared to earlier U(1) flow studies [3, 4].

We train both flows at fixed β and compare NLL curves, the distribution of $|P|$, and the average energy per link.

The 1D U(1) ring results are displayed in Fig. 5 and Fig. 6. Removing the global U(1) phase leads to a simpler quotient density on which the flow trains more stably and reaches a lower NLL than the baseline on X, while the $|P|$ distribution of the quotient flow matches the HMC reference at least as

well as the baseline. This provides a clean example where quotienting a low-dimensional continuous symmetry yields a clear win without degrading gauge-invariant observables, complementing more architecture-centric approaches based on equivariant flows [2, 4].

5.6 2D U(1) gauge theory (negative example)

Finally, we consider 2D U(1) gauge theory on an $L \times L$ lattice with periodic boundary conditions ($L = 16$). Link variables are angles $\theta_\mu(x) \in (-\pi, \pi]$ for $\mu = 1, 2$, and the Wilson plaquette action is

$$S[\theta] = -\beta \sum_x \cos\big(\theta_p(x)\big), \tag{62}$$

where $\theta_p(x)$ is the oriented plaquette angle at site x. We measure:

- average plaquette energy $E_p = 1 - \cos\theta_p$,
- absolute Polyakov loop magnitude $|P|$ along one spatial direction.

Here we explore quotienting *spatial translations* only, leaving gauge transformations untouched. We compare:

- **NF on** X: RealNVP on $\mathbb{R}^{2L^2}$ trained on raw link angles.
- **NF on** X/Λ_x (Fourier-slice canonicalisation in the x-direction): for each configuration we construct the plaquette energy field

  $$f(x) = 1 - \cos\theta_p(x), \tag{63}$$

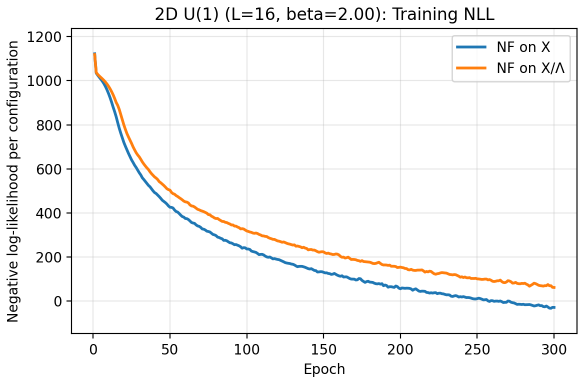
  compute its 2D Fourier transform, and examine the $(k_x, k_y) = (1, 0)$ mode $F_{10} = Ae^{i\varphi}$. A translation by an integer d_x in the x-direction transforms this mode by $F_{10} \mapsto F_{10}\exp(-i2\pi d_x/L)$. We choose an integer shift

  $$d_x \approx \mathrm{round}\left(\frac{L\varphi}{2\pi}\right) \bmod L \tag{64}$$

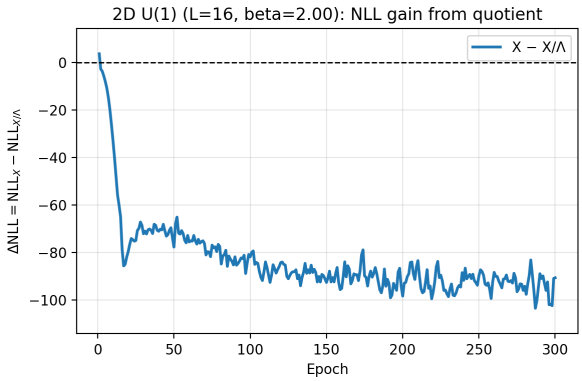
  so that the phase of F_{10} is approximately zero after the shift. This defines a "Fourier slice" canonical representative that fixes translations in x, while leaving translations in y and gauge transformations unconstrained. At sampling time we draw canonical configurations from the flow and undo the slice by applying a random translation in x.

This construction is inspired by slice methods for systems with continuous symmetries [7], but here it is applied directly to lattice gauge fields and then combined with a generic RealNVP backbone rather than a specialised equivariant architecture [4, 2].

In sharp contrast, Fig. 7 shows that quotienting spatial translations in 2D U(1) gauge theory via a Fourier-slice canonicalisation substantially *worsens* the NLL: the flow on X/Λ_x remains far above the baseline on X throughout training. The Polyakov loop histograms in Fig. 8 indicate that $|P|$ itself is not catastrophically distorted, but separate diagnostics of the plaquette energy (not shown in the main text) reveal a significant bias. Together, these plots illustrate that quotienting can be counterproductive when the chosen canonicalisation is geometrically ill-conditioned in the coordinates where the flow operates, and that naive slices in bare link-angle space can be less effective than more tailored equivariant constructions [4].

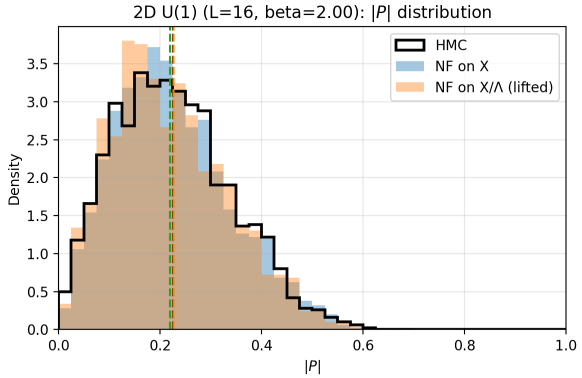


(a) Training NLL on X and on X/Λ_x.

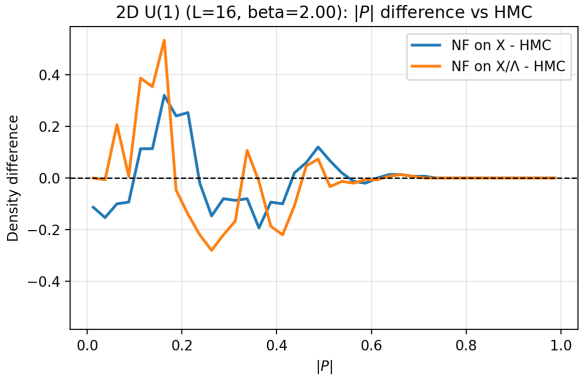


(b) NLL gap relative to the flow on X.

Figure 7: 2D U(1) gauge theory ($L = 16$, $\beta = 2.0$): quotienting spatial translations via a Fourier slice produces a density which is *harder* to model; NLL is significantly worse than for a flow on the full space X.



(a) Histogram of $|P|$ (HMC vs flows).



(b) Difference $\Delta p(|P|)$ w.r.t. HMC.

Figure 8: 2D U(1) gauge theory: Polyakov loop statistics. The $|P|$ distribution is reasonably reproduced, but the plaquette energy (not shown) is significantly biased, underscoring that quotienting does not automatically improve all observables in complex gauge systems.

5.7 2D SU(2) gauge theory (translation quotient; negative result)

We next investigate whether quotienting by spatial translations can improve generative modelling for non-Abelian gauge fields. We consider 2D SU(2) gauge theory at lattice size $L = 8$ and coupling $\beta = 2.20$, using quaternion-valued link variables with the standard Wilson plaquette action. As in the U(1) case, we construct a deterministic translation–canonicalisation map $\mathrm{can} : X \to X$ which aligns each configuration to a fixed slice under the action of the spatial translation group Λ. We then train two RealNVP models: (i) a baseline flow on the raw configuration space $X \cong (\mathrm{SU}(2))^{2L^2}$, and (ii) a quotient flow on the canonicalised representatives X/Λ.

Figure 9 shows the training negative log-likelihood (NLL) for both models. The qualitative behaviour mirrors, but significantly amplifies, the pathology already observed in the 2D U(1) example: the flow trained on the quotient space exhibits consistently *worse* NLL, with a stable gap of approximately 200-300 units throughout training. This indicates that the quotient density on the translation-fixed slice is *harder* to approximate than the original SU(2) gauge distribution.

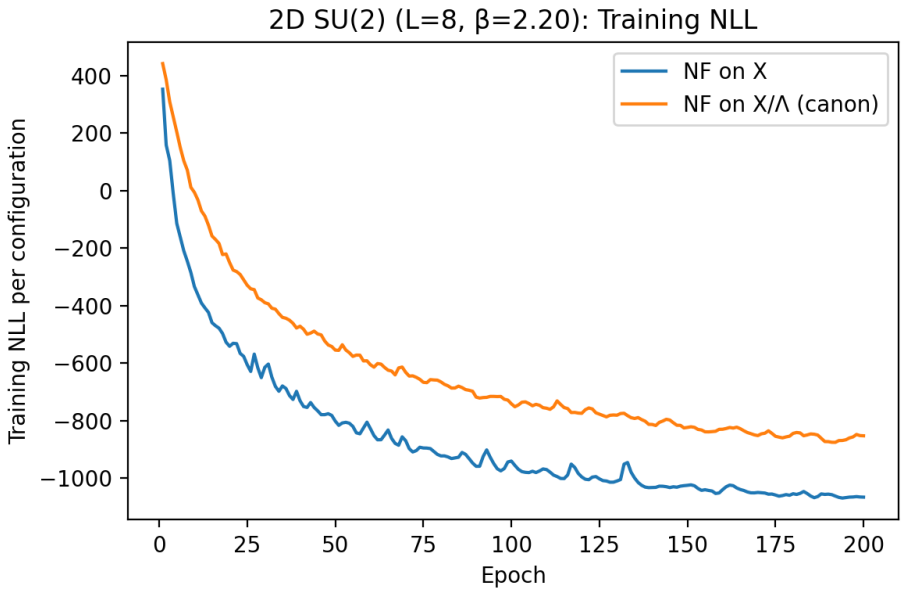


Figure 9: 2D SU(2) gauge theory ($L = 8$, $\beta = 2.20$): training NLL for flows on X and X/Λ. The quotient model shows consistently worse likelihood, with a stable gap of roughly 200–300.

Geometrically, this is expected. The translation canonicalisation collapses a large SU(2) orbit into a fixed coordinate frame using a highly nonlinear rule based on plaquette-derived features. Small changes in the original configuration can produce discontinuous changes in the chosen shift, leading to a quotient density which is strongly anisotropic and poorly aligned with the Euclidean geometry assumed by RealNVP. As in the U(1) case, quotienting does not commute with the local SU(2) gauge structure, and the resulting slice does not form a smooth manifold representative of gauge orbits.

To assess physical fidelity, we measure the absolute Polyakov loop $|P|$ for HMC samples, for flow samples from X, and for flow samples drawn from X/Λ and lifted via quaternion renormalisation. Figure 10 compares the histograms, while Figure 11 shows the corresponding difference plots relative to HMC. Both flows reproduce $|P|$ reasonably well, but the quotient model exhibits slightly larger deviations and mild oversmoothing. This confirms that translation quotienting in SU(2) does not offer any improvement in the gauge-invariant observables considered here.

Overall, these results demonstrate that, for non-Abelian 2D gauge theories, the naive translation quotient produces a quotient density that is *empirically and geometrically harder* to model. In contrast to the positive results in scalar theories or in 1D U(1), quotient flows provide no benefit here and can substantially degrade likelihood performance. This negative result echoes the theoretical warning that canonical slices for high-dimensional non-Abelian symmetries generically introduce discontinuities and large distortions in probability mass, rendering them unsuitable for Euclidean normalising flows.

5.8 2D U(1) gauge theory (gauge quotient; negative result)

We now test whether quotienting by *local* gauge transformations can improve generative modelling for 2D U(1) gauge theory. At $L = 16$, $\beta = 2.0$, we generate an HMC ensemble of link angles $\theta_\mu(x) \in (-\pi, \pi]$ with the Wilson action, and construct a deterministic maximal-tree gauge fixing that produces a canonical representative X/G_{gauge} for each orbit. By design, gauge-invariant

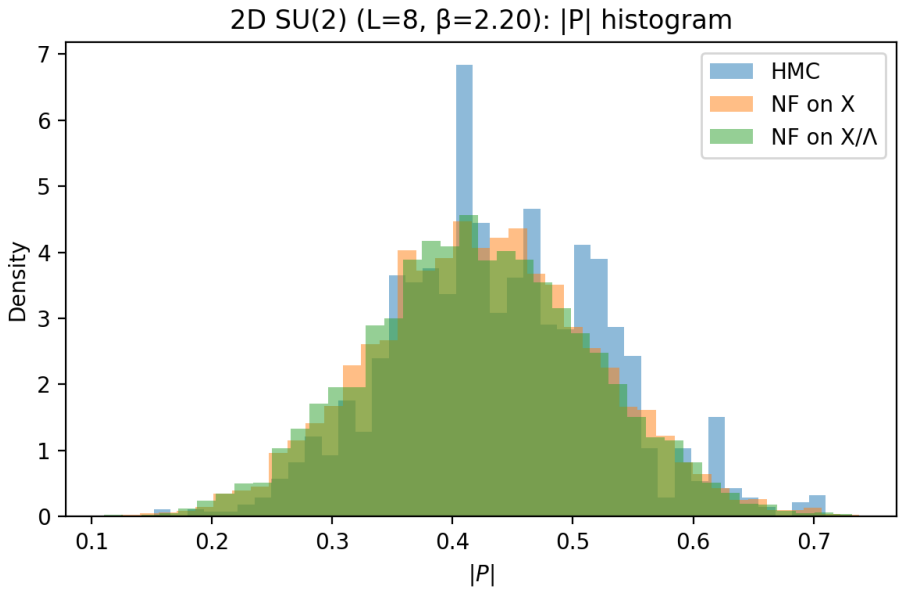


Figure 10: Histogram of Polyakov-loop magnitudes $|P|$ for HMC, NF on X, and NF on X/Λ. Both flows reproduce the overall shape reasonably well, though the quotient model displays slight oversmoothing.

observables such as the plaquette energy and Polyakov loop remain unchanged under this canonicalisation.

We train two RealNVP flows: one directly on the raw ensemble X, and one on the gauge-fixed ensemble X/G_{gauge}. Samples from the quotient flow are lifted back to X at evaluation time by applying random site-wise gauge transformations drawn from the Haar measure.

Figure 12 shows the training NLL. The quotient flow achieves a *substantially lower* NLL (by $\mathcal{O}(10^3)$) compared to the baseline flow. From the perspective of Euclidean density estimation on the gauge-fixed coordinates, the quotient model therefore appears "better".

However, this NLL improvement does not translate into better physics. The HMC reference yields $\langle E_p \rangle = 0.3020(6)$ and $\langle |P| \rangle = 0.2251(26)$. The baseline flow already exhibits a severely inflated plaquette energy ($\langle E_p \rangle_X \approx 0.75$), while reproducing $\langle |P| \rangle$ reasonably well. The quotient flow does not repair this discrepancy: it predicts $\langle E_p \rangle_{X/G} \approx 0.55$ and significantly overestimates the Polyakov loop magnitude ($\langle |P| \rangle_{X/G} \approx 0.45$).

Thus, although the quotient flow fits the gauge-fixed density extremely well, the induced gauge-invariant measure obtained after random gauge lifting is *worse*. This illustrates a key caution: likelihoods computed in gauge-fixed coordinates are sensitive to the geometry of the slice (including the implicit Faddeev–Popov determinant), whereas physical observables depend on the distribution on the gauge orbit space. A model that excels at the former may perform poorly on the latter. In this 2D U(1) example, naive gauge quotienting therefore degrades, rather than improves, generative modelling.

5.9 2D SU(2) gauge theory (gauge quotient; very negative result)

Motivated by the U(1) study above, we perform an analogous gauge-quotient experiment for 2D SU(2) gauge theory at $L = 16$, $\beta = 2.20$. Configurations are represented as quaternion-valued link variables with $\|q\| = 1$, and we again use the Wilson plaquette action with Metropolis updates to

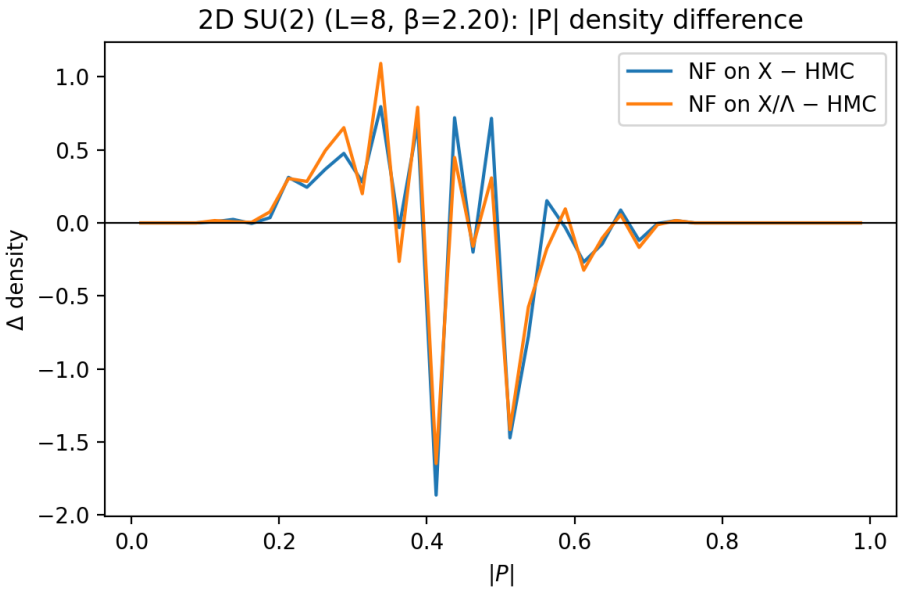


Figure 11: Bin-wise density differences $\Delta p(|P|)$ relative to HMC. Deviations are comparable between the two flows, with the quotient model exhibiting slightly larger fluctuations.

generate the reference ensemble.

To quotient local gauge transformations we implement a "tree gauge": starting from a fixed root site we traverse a spanning tree of the lattice, accumulating SU(2) gauge transformations $g(x)$ along each edge such that all links on the tree are mapped to the identity. This produces a deterministic gauge-fixed representative U_{gf} for each orbit, and defines a canonicalisation map $X \to X/G_{\text{gauge}}$. As a sanity check we verify numerically that gauge-invariant observables (plaquette energy and Polyakov loop) are unchanged by this procedure.

We then train two flows with the same RealNVP backbone as in the U(1) and scalar experiments: a baseline flow on the raw ensemble X and a quotient flow on the gauge-fixed ensemble X/G_{gauge}. Samples from the quotient flow are lifted back to X by applying random local SU(2) gauge transformations $g(x)$ drawn from the Haar measure at each site, followed by quaternion renormalisation.

Figure 13 shows the training NLL for both models. The baseline flow on X behaves qualitatively similarly to the earlier SU(2) translation experiment: the NLL decreases steadily and then saturates. In stark contrast, the quotient flow on X/G_{gauge} exhibits a dramatic instability: around epoch ~ 50 the NLL spikes to values of order 10^6, after which training effectively collapses. This is a clear symptom of an extremely ill-conditioned target density in the gauge-fixed coordinates.

The impact on gauge-invariant observables is illustrated in Fig. 14, which compares the $|P|$ histograms from HMC, the baseline flow on X, and the gauge-quotient flow after random gauge lifting. The HMC distribution is sharply peaked at large Polyakov-loop magnitudes, reflecting the relatively ordered regime at this coupling. Both flows fail to reproduce this peak: their $|P|$ histograms are much broader and shifted towards smaller values. Crucially, the gauge-quotient flow does *not* repair this discrepancy and, in light of the NLL spike, is clearly less stable than the baseline.

Together with the U(1) results, this SU(2) experiment reinforces the conclusion that naive gauge quotienting can severely degrade both optimisation and sampling for non-Abelian lattice gauge theories. The tree-gauge slice is highly nonlocal and suffers from Gribov-type ambiguities;

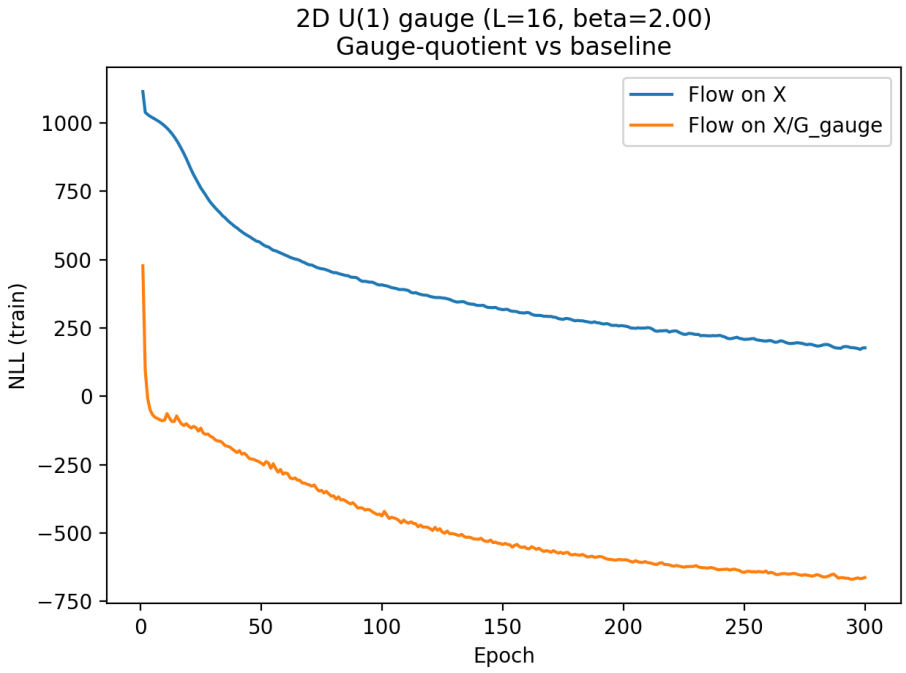


Figure 12: 2D U(1) gauge theory ($L = 16$, $\beta = 2.0$): training NLL for flows on X and X/G_{gauge}. The quotient flow attains a much lower NLL on the gauge-fixed coordinates, but induces worse gauge-invariant observables after random gauge lifting.

its induced density on X/G_{gauge} is extremely anisotropic and contains sharp structures that a standard Euclidean flow cannot represent without numerical blow-ups. In this regime, quotient flows are therefore *strictly worse* than flows trained directly on the full gauge configuration space.

6 Conclusion and Outlook

This work set out to examine whether the substantial symmetry redundancies present in lattice gauge theories, arising from both spacetime translations and local gauge transformations, can be exploited to improve the performance of normalizing flows. By mapping each configuration to a canonical representative of its symmetry orbit and training a flow on this reduced space, the quotient-flow construction aims to reduce the apparent dimensionality and eliminate directions irrelevant to gauge-invariant observables. Such a reduction would, in principle, allow generative models to focus their capacity on physically meaningful degrees of freedom and thereby achieve more accurate sampling at lower computational cost.

Across four detailed case studies in two-dimensional U(1) and SU(2) gauge theories we find that this intuition does not hold in practice. Translation quotienting leaves the gauge-invariant observables essentially unchanged but consistently *worsens* the performance of flow models, leading to higher bias in the plaquette energy and distortions in the Polyakov-loop distribution. Gauge quotienting, implemented via maximal-tree gauge fixing, yields an even more striking pattern: the quotient flows achieve substantially *lower* training NLL, sometimes by several hundred nats, yet their predictions for gauge-invariant observables deviate strongly from the HMC baseline. In SU(2), quotienting can destabilize the training entirely, producing large NLL spikes and a systematic collapse of the Polyakov-loop histogram. The severity and universality of these failures indicate

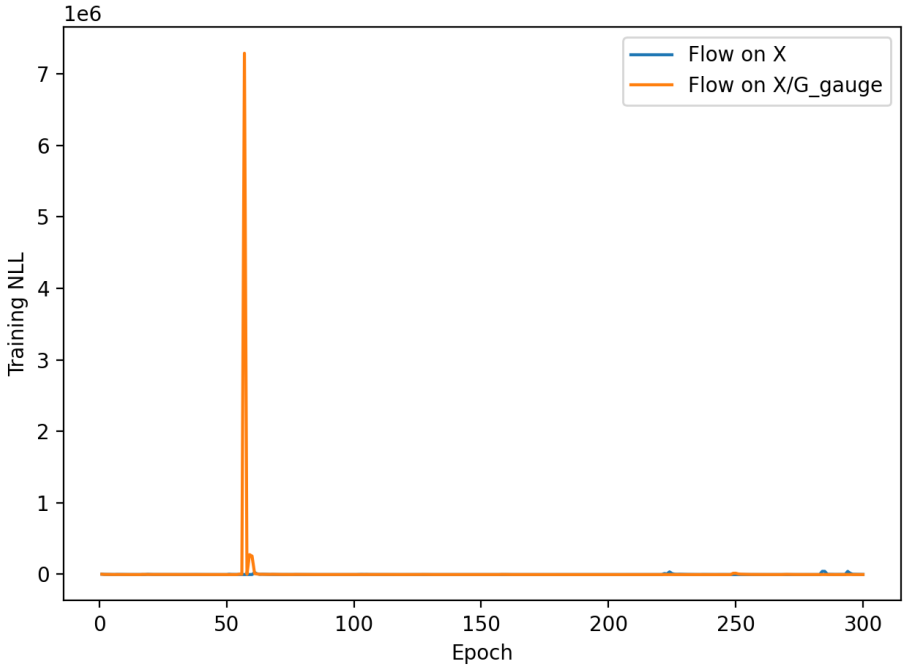


Figure 13: 2D SU(2) gauge theory ($L = 16$, $\beta = 2.20$): training NLL for flows on X and on the gauge-fixed slice X/G_{gauge}. The quotient flow develops a catastrophic NLL spike of order 10^6 around epoch ~ 50, indicating a severe geometric instability in the gauge-fixed coordinates.

that the issue is structural rather than architectural.

The key insight is that optimizing likelihood on a canonical slice does *not* correspond to learning the correct gauge-invariant distribution. The canonical slice inherits a complicated geometry from the quotient manifold, including a nontrivial induced volume form analogous to the Faddeev-Popov determinant. Standard flows implicitly assume a Euclidean base measure and therefore fit the wrong density when trained on the slice. Even if the slice parametrization is globally smooth, which is already nontrivial for non-Abelian gauge theories, the mismatch between the slice measure and the Haar-invariant measure on the orbit space means that likelihood improvement on the slice may be anticorrelated with improvements in gauge-invariant observables. Our experiments empirically confirm this mismatch: better NLL on the quotient coordinates does not yield better sampling of the physical ensemble, and can in fact push the model further away from the true orbit-space distribution.

These findings place clear conceptual constraints on the use of naïve quotienting for generative modelling of gauge theories. They also sharpen the distinction between two different notions of model quality: the ability to fit an Euclideanized coordinate representation and the ability to faithfully reproduce gauge-invariant physics. Our results demonstrate that a model may excel at the former while failing at the latter, suggesting that likelihood-based model selection must be interpreted with considerable care in the presence of gauge orbits and nontrivial manifold structures.

Nevertheless, several promising research directions emerge from this study. First, symmetry *equivariant* flows sidestep the need for gauge fixing altogether by constructing transformations that commute with gauge and translation actions. Such models inherently operate on the orbit space and therefore avoid the measure mismatch that plagues the quotient approach. Second, one could attempt to incorporate the correct induced measure on the slice, either by explicitly computing the relevant Jacobians or by learning a correction factor that accounts for the Faddeev-Popov volume

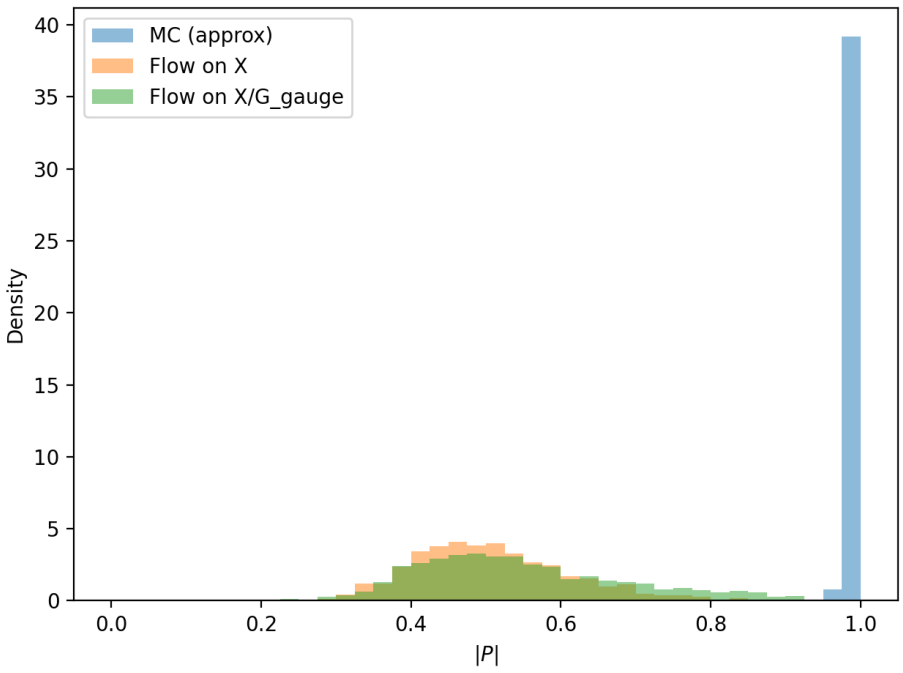


Figure 14: 2D SU(2) gauge theory: histogram of Polyakov-loop magnitudes $|P|$ for HMC, the baseline flow on X, and the gauge-quotient flow (after random gauge lifting). The HMC ensemble is strongly peaked at large $|P|$, while both flows produce a broader distribution at smaller $|P|$; the gauge-quotient flow does not improve over the baseline.

element. This would convert quotient flows into principled models of the orbit-space distribution rather than Euclidean proxies. Third, hybrid methods that combine equivariance with reduced coordinates, for example by quotienting global symmetries while treating local gauge symmetry equivariantly, may strike a better balance between redundancy reduction and measure fidelity. Fourth, it remains an open question whether quotienting becomes advantageous in polynomially larger systems, higher-rank gauge groups, or in theories with matter fields where the orbit structure is more intricate. Finally, the failure modes documented here underline the importance of observable-based evaluation metrics: future generative models for lattice gauge theory should be assessed primarily through gauge-invariant quantities, not through likelihood alone.

In summary, quotient normalizing flows provide a clear and theoretically motivated framework for exploring symmetry reduction in generative modelling, but their naïve implementation does not improve sampling quality and can significantly degrade physical accuracy. By identifying the geometric and measure-theoretic origins of this limitation, our work clarifies the obstacles that must be overcome for symmetry-aware generative models to reach their full potential. We hope that these results will guide the development of more robust gauge-equivariant architectures, improved training objectives, and principled approaches to modelling the orbit-space structure of lattice gauge theories.

Acknowledgements

I would like to thank Dr. Chris Bouchard for his guidance and support as my internship supervisor throughout this project. His suggestions, many discussions, and careful feedback were invaluable

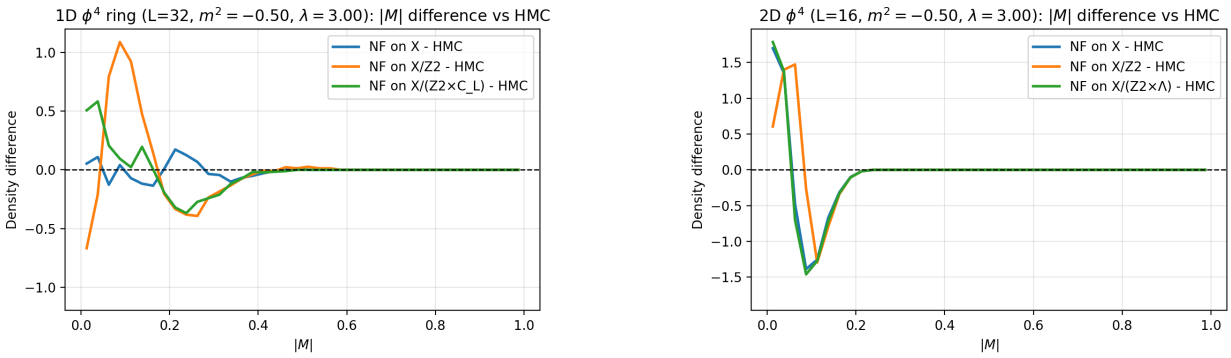


Figure 15: Representative difference histograms $\Delta p(|M|)$ for 1D (left) and 2D (right) ϕ^4 theories. The quotient flows track the HMC magnetisation distribution at least as well as the baseline flows on X.

in shaping both the theoretical development of quotient normalising flows and the design of the numerical experiments presented in this work.

Code Availability

The source code implementing the experimental models is available at https://github.com/QCD-Internship/QuotientFlows.

A Additional Plots and Diagnostics

A.1 Difference histograms for $|M|$ and $|P|$

For completeness, we collect here additional diagnostics that were omitted from the main text for brevity. Figures 2b and 4b show bin-wise differences between the HMC histograms of $|M|$ (for the scalar theories) and the corresponding histograms obtained from the various flows. Analogous plots for the U(1) models are qualitatively similar and therefore not shown.

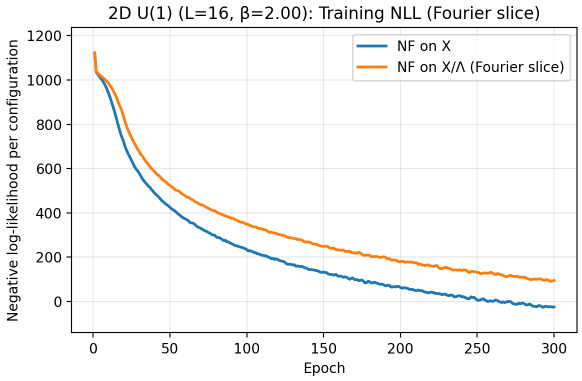
These difference plots confirm the qualitative picture already visible in the main-text histograms: in the 1D and 2D ϕ^4 models and in the 1D U(1) ring, quotient flows on X/G do not introduce systematic distortions of the magnetisation or Polyakov-loop distributions; their deviations from HMC are comparable to, and sometimes smaller than, those of the baseline flows on X. In 2D U(1) gauge theory, the difference plots highlight that the Fourier-slice quotient flow (discussed below) slightly under- or over-populates some $|P|$ bins, consistent with the observation that the overall energy statistics are biased, even though the Polyakov loop itself looks visually reasonable.

A.2 Fourier-slice canonicalisation for 2D U(1)

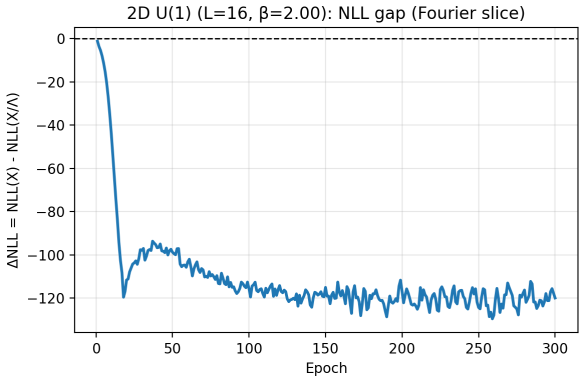
In the main text we reported results for a Fourier-slice canonicalisation of 2D U(1) gauge theory, where the phase of the $(k_x, k_y) = (1, 0)$ Fourier mode of the plaquette energy field is used to fix translations in the x direction. For completeness, Figs. 16a and 16b show the corresponding training curves for the baseline flow on X and the Fourier-slice quotient flow on X/Λ_x, while Figs. 16c and 16d display the associated Polyakov-loop statistics.

The qualitative behaviour mirrors that of the simpler translation-quotient experiment: the flow trained on the Fourier-slice quotient space X/Λ_x exhibits substantially worse NLL than the flow

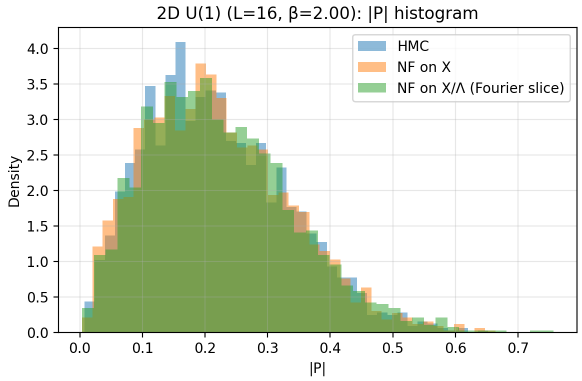
on X, and offers no clear improvement in the $|P|$ distribution. These additional plots reinforce the conclusion that, for 2D U(1) at the parameters considered here, our current choices of canonicalisation do not yield a quotient density that is easier to model with a vanilla RealNVP, underscoring the need for more geometry-aware architectures or representations in the gauge case.



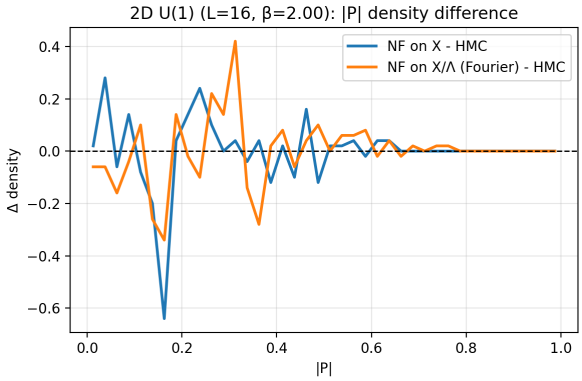
(a) Training NLL for flows on X and X/Λ_x.



(b) NLL gap $\Delta\mathrm{NLL} = \mathrm{NLL}_X - \mathrm{NLL}_{X/\Lambda_x}$.



(c) Polyakov-loop magnitude $|P|$ distribution.



(d) Difference histogram relative to HMC.

Figure 16: 2D U(1), Fourier-slice variant: training NLL (top left), NLL gap (top right), Polyakov-loop magnitude $|P|$ distribution (bottom left), and the corresponding difference histogram relative to HMC (bottom right).

References

[1] Taco Cohen and Max Welling. Group equivariant convolutional networks. In *International conference on machine learning*, pages 2990–2999. PMLR, 2016.

[2] Jonas Köhler, Leon Klein, and Frank Noé. Equivariant flows: exact likelihoods for equivariant deep learning. *Advances in Neural Information Processing Systems*, 33:2866–2876, 2020.

[3] Michael S Albergo, Gurtej Kanwar, and Phiala E Shanahan. Flow-based generative models for markov chain monte carlo in lattice field theory. *Physical Review D*, 100(3):034515, 2019.

[4] Gurtej Kanwar, Michael S Albergo, Denis Boyda, Kyle Cranmer, Daniel C Hackett, Sébastien Racanière, Danilo Jimenez Rezende, and Phiala E Shanahan. Equivariant flow-based sampling for lattice gauge theory. *Physical Review Letters*, 125(12):121601, 2020.

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

[6] Laurent Dinh, Jascha Sohl-Dickstein, and Samy Bengio. Density estimation using real nvp. In *International Conference on Learning Representations (ICLR)*, 2017.

[7] Clarence W Rowley and Jerrold E Marsden. Reconstruction equations and the karhunen–loève expansion for systems with symmetry. *Physica D: Nonlinear Phenomena*, 142(1-2):1–19, 2000.

[8] Simon Duane, Anthony D Kennedy, Brian J Pendleton, and Duncan Roweth. Hybrid monte carlo. *Physics letters B*, 195(2):216–222, 1987.

[9] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, 2015.