NEURAL NON-EQUILIBRIUM HAMILTONIAN MONTE CARLO FOR CORRECTED BOLTZMANN SAMPLING

Anonymous authors
Paper under double-blind review

ABSTRACT

Learned dynamical proposals can generate configurations without providing a tractable endpoint density. Nonequilibrium path probabilities offer a way to correct such proposals, but correction alone does not determine their ability to connect separated regions. We introduce Neural Non-Equilibrium Hamiltonian Monte Carlo (NHMC), which combines conditional momentum distributions with reversible, volume-preserving dynamics. The forward–reverse path ratio gives the work used for training, importance weighting, normalizer estimation, and Metropolis correction. We then construct a configuration-space round-trip kernel whose reverse and forward paths share an intermediate configuration. Conditional on that configuration, its path-record update is independence Metropolis–Hastings. We compare its stationary inter-region flow with the flow obtained under exact conditional matching and bound their difference by the conditional path mismatch. This separates errors in the conditional proposal from dependence between the endpoint region and the intermediate configuration. Many-well experiments test normalizers and mode probabilities; a controlled lattice ϕ^4 experiment relates the predicted flow to sector relaxation. Two-dimensional $U(1)$, $\mathrm{SU}(2)$, and $\mathrm{SU}(3)$ experiments compare four proposal constructions sharing a structured reference, including paths with analytic and learned forces.

1 INTRODUCTION

Sampling from an unnormalized Boltzmann target requires effective exploration and controlled statistical error. Hamiltonian Monte Carlo (HMC) provides an exact target-preserving transition under standard conditions, while transitions between separated regions can remain slow (Duane et al., 1987). Learned dynamical proposals can generate more flexible finite-time moves, but their endpoint density may be unavailable and their raw endpoints need not follow the target distribution. In lattice field theory, this distinction affects both precision observables and transitions between topological sectors (Lüscher & Schaefer, 2011; Nicoli et al., 2023).

Nonequilibrium work identities connect forward and reverse path probabilities to free-energy differences (Jarzynski, 1997; Crooks, 1999). Annealed importance sampling and sequential Monte Carlo use extended-state weights for estimation (Neal, 2001; Del Moral et al., 2006), while tempered transitions and nonequilibrium candidate Monte Carlo use forward–backward paths in exact Markov updates (Neal, 1996; Nilmeier et al., 2011). Stochastic normalizing flows and annealed flow transport extend path-based weighting to learned transformations (Wu et al., 2020; Arbel et al., 2021). These constructions show that a complete stochastic path record can retain the probability information needed for statistical correction even when the marginal endpoint density is not available.

Recent diffusion-path methods for Markov chain Monte Carlo (MCMC) reach a related correction principle from a different dynamical construction. The Metropolis-adjusted diffusion path (MAD-Path) sampler uses an Ornstein–Uhlenbeck noising process and an approximate reverse diffusion whose intermediate scores must be learned, while Denoising Diffusion Monte Carlo (DDMC) similarly corrects a denoising-diffusion proposal using its forward and reverse path densities (Chen et al., 2026; Hill, 2026). Their stochastic paths are diffusion trajectories. NHMC instead works in Hamiltonian phase space: each stage samples a conditional momentum and applies a reversible, volume-preserving kick–drift map. Repeated momentum refreshes make the endpoint law implicit,

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

while separate forward and reverse conditional momentum laws give tractable densities to the complete Hamiltonian record.

This construction is also distinct from learned HMC transition kernels, which parameterize a current-state Markov move, and from Hamiltonian annealing, which uses time-dependent Hamiltonians together with nonequilibrium weighting (Levy et al., 2018; Foreman et al., 2021; Sohl-Dickstein & Culpepper, 2012; Cohn-Gordon et al., 2026). In NHMC, the ratio of the complete forward and reverse record probabilities defines a nonequilibrium work used consistently for training, importance weighting, normalizer and free-energy estimation, and Metropolis correction.

Using the same forward and reverse path laws, we construct a round-trip Markov update. Starting from the current configuration, a reverse path reaches an intermediate configuration, and a fresh forward path from that same configuration proposes the next state. Conditional on the intermediate configuration, the path-record update is an independence Metropolis–Hastings (IMH) step. Under exact conditional matching, the update is a classical data-augmentation chain whose inter-region flow is determined by the conditional region probabilities given the intermediate configuration (Tanner & Wong, 1987; Liu et al., 1994). We bound the deviation of the accepted flow from this ideal value using the mismatch between the proposal and target conditional path laws, and obtain a further bound from the Kullback–Leibler (KL) divergence between the full forward and reverse path laws.

The experiments follow the same separation. Many-well targets test path correction against analytic normalizers and unequal mode probabilities. A controlled lattice ϕ^4 experiment changes inter-sector flow by more than a factor of four while the mean acceptance remains nearly unchanged, and directly compares the accepted flow with the value obtained under exact conditional matching. In $U(1)$, $\mathrm{SU}(2)$, and $\mathrm{SU}(3)$ gauge theories, we compare the effective sample size (ESS) computed from path weights across four proposal constructions using the same structured reference and evaluation grid. Corrected Wilson loops provide an independent observable check.

2 Hamiltonian Path Construction and Correction

Let $\gamma(x) = \exp[-E(x)]$, $Z = \int \gamma(x)\,\mu(dx) \in (0, \infty)$, and $\pi(x) = \gamma(x)/Z$. Here μ is Lebesgue measure for Euclidean configurations and product Haar measure for compact gauge fields. Let q_0 denote a normalized reference density with respect to μ. We first present the continuous path construction. Recorded discrete moves enter through their forward and reverse probabilities, as specified in Appendix A.2.

2.1 Forward and reverse path laws

Starting from $x_0 \sim q_0$, stage $t = 0, \ldots, T-1$ draws a momentum and applies the deterministic map

$$\begin{aligned} p_t &\sim q^F_{\theta,t}(\cdot \mid x_t), \\ (x_{t+1}, \bar{p}_{t+1}) &= \Phi_{\theta,t}(x_t, p_t). \end{aligned} \tag{1}$$

Here T is the number of momentum draws, and $\bar{p}_{t+1}$ denotes the momentum produced by $\Phi_{\theta,t}$; the next momentum p_{t+1} is sampled at the following stage. Let ν denote Lebesgue measure on the momentum space; for compact gauge fields this is Lebesgue measure on the Lie-algebra coordinates. We require $\Phi_{\theta,t}$ to preserve $\mu \otimes \nu$ and to satisfy

$$\Phi^{-1}_{\theta,t} = \mathcal{R}\Phi_{\theta,t}\mathcal{R}, \qquad \mathcal{R}(x,p) = (x,-p). \tag{2}$$

The corresponding reverse input is therefore $-\bar{p}_{t+1}$.

The implemented maps are palindromic kick–drift compositions with state-independent step sizes. A position-dependent kick need not be a potential gradient: at fixed x it is a translation of momentum. On $\mathrm{SU}(N_c)$, with N_c the number of colors and Hermitian traceless momenta P_ℓ, the drift is $U_\ell \leftarrow \exp(i\epsilon P_\ell)U_\ell$ and preserves Haar measure. The gauge implementations and their symmetry assumptions are specified in Appendices F.5 and G.

The complete record Γ contains x_0, every stochastic draw, and the states determined by the maps. Define normalized reverse laws $q^R_{\theta,t}$ and let the reverse process start from $x_T \sim \pi$. For fixed maps,

both path laws can be written on the same record coordinates:

$$Q_\theta(d\Gamma) = q_0(x_0) \prod_{t=0}^{T-1} q^F_{\theta,t}(p_t \mid x_t)\, d\lambda_\theta(\Gamma), \tag{3}$$

$$P_\theta(d\Gamma) = \pi(x_T) \prod_{t=0}^{T-1} q^R_{\theta,t}(-\bar{p}_{t+1} \mid x_{t+1})\, d\lambda_\theta(\Gamma). \tag{4}$$

The measure λ_θ is induced by the fixed deterministic maps and the product measure on the random records. The endpoint marginal of P_θ is π by construction. The reverse path law, unlike its endpoint target, depends on θ.

2.2 Work, weighting, and Metropolis correction

The recorded work is

$$W_\theta(\Gamma) = \log q_0(x_0) - \log \gamma(x_T) + \sum_{t=0}^{T-1} \log \frac{q^F_{\theta,t}(p_t \mid x_t)}{q^R_{\theta,t}(-\bar{p}_{t+1} \mid x_{t+1})}. \tag{5}$$

An invertible discrete move contributes its forward-minus-reverse log probability; its position in the stage fixes the conditioning states and reverse order. Appendix A.2 gives the full formula. A map that does not preserve measure would additionally require its Jacobian.

Theorem 1 (Recorded forward–reverse path ratio)**.** *Fix θ, normalized stochastic kernels, and reversible, measure-preserving path maps. Assume the two path laws have positive finite densities on a common support. With $\Delta F = -\log Z$,*

$$\log \frac{dQ_\theta}{dP_\theta} = W_\theta - \Delta F, \qquad \frac{dP_\theta}{dQ_\theta} = \frac{e^{-W_\theta}}{Z}. \tag{6}$$

Consequently $\mathbb{E}_{Q_\theta}[e^{-W_\theta}] = Z$ and self-normalized importance sampling (SNIS) with path weights is consistent for π-integrable endpoint observables.

This is the path-weight identity specialized to NHMC; the proof and weaker one-sided support conditions are given in Appendix A. For N independent paths, write $\widetilde{w}_i = e^{-W_i}$ and $\bar{w}_i = \widetilde{w}_i / \sum_j \widetilde{w}_j$. The estimators are

$$\widehat{\mathbb{E}_\pi[O]} = \sum_{i=1}^{N} \bar{w}_i O(x_T^{(i)}), \qquad \widehat{Z} = \frac{1}{N} \sum_{i=1}^{N} \widetilde{w}_i. \tag{7}$$

The first estimator is generally biased at finite N but consistent; $\widehat{Z}$ is unbiased, whereas $-\log \widehat{Z}$ has finite-sample bias. For an equilibrium reference known only through γ_0, replacing $\log q_0$ by $\log \gamma_0$ instead estimates Z/Z_0.

Independent path Metropolis correction. A fresh path $\Gamma' \sim Q_\theta$ is accepted with $1 \wedge \exp[W_\theta(\Gamma) - W_\theta(\Gamma')]$. Rejection retains the current path and its weight. This is independence Metropolis–Hastings on path space (Hastings, 1970; Mengersen & Tweedie, 1996); its invariant law is P_θ, and its stationary endpoint law is π. It produces a Markov chain, not independent target samples.

2.3 Training

The training objective combines mean work with its variance,

$$\mathcal{L}_{\text{train}} = \mathbb{E}_{Q_\theta}[W_\theta] + \lambda_{\text{var}} \operatorname{Var}_{Q_\theta}(W_\theta), \qquad D_\theta := D_{\text{KL}}(Q_\theta \| P_\theta) = \mathbb{E}_{Q_\theta}[W_\theta] - \Delta F. \tag{8}$$

The KL identity concerns the full path laws, not only their endpoint marginals. The variance penalty is an optimization choice. Continuous parameters are differentiated through the path; the treatment of discrete probabilities in the archived training runs is stated in Appendix C.3. All parameters are frozen for evaluation.

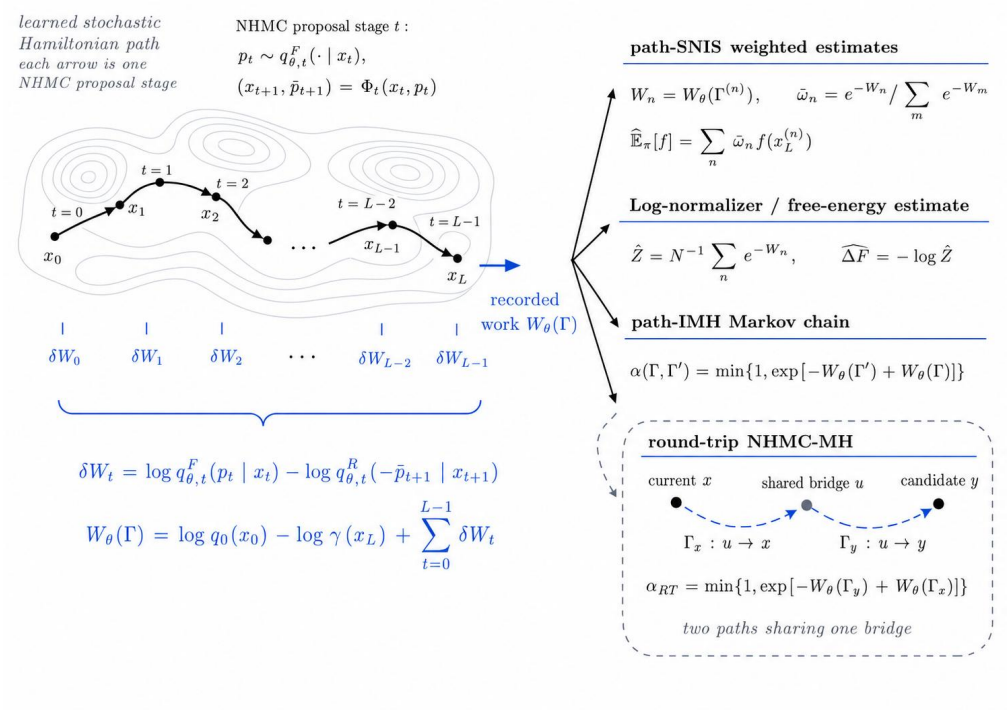

Figure 1: NHMC path construction and recorded-work corrections. Each proposal stage samples a momentum and applies a reversible, volume-preserving map. The recorded work combines endpoint target/reference terms with forward–reverse log-ratios; optional discrete-move terms are omitted from the stage formula shown. The same work defines path-SNIS weights, normalizer/free-energy estimates, and path-IMH acceptance. For the round-trip update, the current configuration x is pulled back to u along Γ_x, then a candidate y is proposed along Γ_y and accepted or rejected once (Section 3). The schematic denotes path length by L; the text uses T.

3 Shared-Bridge Round Trips and Inter-Region Flow

The round-trip analysis separates the geometry of the intermediate configuration from conditional proposal error. Endpoint–intermediate dependence determines the probability flow between regions when the conditional path distributions match exactly, whereas conditional path mismatch controls the deviation of the accepted flow from that ideal. Consequently, a large Metropolis acceptance probability does not by itself imply substantial transport between prescribed regions.

3.1 A configuration-space Metropolis update

The same path laws define a transition from the current configuration, without retaining a previous path. Let u be an intermediate configuration, a a forward random record, x its endpoint, and b the corresponding reverse record. The fixed path maps induce a measure-preserving bijection

$$\Psi_\theta : (u, a) \longleftrightarrow (x, b). \tag{9}$$

Write $f_\theta(a \mid u)$ and $r_\theta(b \mid x)$ for the normalized forward and reverse record densities. Starting from x, draw $b \sim r_\theta(\cdot \mid x)$ and invert Ψ_θ to recover (u, a). Draw a fresh $a' \sim f_\theta(\cdot \mid u)$ and map it to (y, b'). Accept y with probability

$$\alpha_{\mathrm{RT}} = 1 \wedge \frac{\gamma(y) r_\theta(b' \mid y) f_\theta(a \mid u)}{\gamma(x) r_\theta(b \mid x) f_\theta(a' \mid u)} = 1 \wedge e^{W_\theta(\Gamma_x) - W_\theta(\Gamma_y)}. \tag{10}$$

Both $\Gamma_x : u \to x$ and $\Gamma_y : u \to y$ use the forward orientation. Their reference term $\log q_0(u)$ cancels. The direct ratio remains the primary definition when that term is unavailable.

Theorem 2 (Round-trip reversibility)**.** *For fixed normalized record laws and a bimeasurable, measure-preserving bijection Ψ_θ, the update in* (10)*, with the usual zero-density Metropolis conventions, is reversible with respect to π. The work form requires finite work values on the common bridge support.*

The proof swaps the two forward records in the augmented space (x, b, a') and projects the symmetric accepted measure to (x, y) (Appendix A.5). Forward–backward path Metropolis moves have precedents (Neal, 1996; Nilmeier et al., 2011; Neklyudov et al., 2020).

3.2 Actual and ideal conditional flow

To characterize the stationary update, start with $X \sim \pi$ and $B \sim r_\theta(\cdot \mid X)$. Their inverse image has law

$$(\Psi_\theta^{-1})_\#[\pi(dx) r_\theta(db \mid x)] = \rho_\theta(du) h_\theta(da \mid u). \tag{11}$$

Thus ρ_θ is the induced distribution of the intermediate configuration, and h_θ is the target-induced conditional record law. At fixed u, substitution into (10) cancels $\rho_\theta(u)$: the record update is independence Metropolis–Hastings with target $h_\theta(\cdot \mid u)$ and proposal $f_\theta(\cdot \mid u)$.

Let $\{A_k\}$ be a measurable partition of configuration space, let $\mathsf{E}_\theta(u, a)$ denote the endpoint of a forward record, and set $C_k(u) = \{a : \mathsf{E}_\theta(u, a) \in A_k\}$. Using the total variation (TV) distance $d_{\mathrm{TV}}(p, q) = \frac{1}{2}\int |p - q|$, define

$$m_k(u) = h_\theta(C_k(u) \mid u), \qquad \Omega_{k\ell} = \mathbb{E}_{\rho_\theta}[m_k(U) m_\ell(U)], \qquad \delta_\theta(u) = d_{\mathrm{TV}}(h_\theta(\cdot \mid u), f_\theta(\cdot \mid u)). \tag{12}$$

The stationary mean rejection at fixed u lies between $\delta_\theta(u)$ and $2\delta_\theta(u)$. The quantity $\Omega_{k\ell}$ is the stationary flow of an *ideal* update that refreshes the record from h_θ. It is not generally the actual proposal flow: writing $n_k(u) = f_\theta(C_k(u) \mid u)$, the latter is $G_{k\ell} = \mathbb{E}_{\rho_\theta}[m_k(U) n_\ell(U)]$.

Theorem 3 (Stationary flow comparison)**.** *Assume the conditional record laws have densities with respect to a common measure $m(da)$. For $k \neq \ell$, the accepted stationary inter-region flow is*

$$\begin{aligned} F_{k\ell}^{\mathrm{RT}} &:= \int_{A_k} \pi(dx) K_{\mathrm{RT}}(x, A_\ell) \\ &= \int \rho_\theta(du) \int_{C_k(u)} m(da) \int_{C_\ell(u)} m(da') \min\{h_\theta(a \mid u) f_\theta(a' \mid u), f_\theta(a \mid u) h_\theta(a' \mid u)\}, \end{aligned} \tag{13}$$

and

$$\left|F_{k\ell}^{\mathrm{RT}} - \Omega_{k\ell}\right| \leq \mathbb{E}_{\rho_\theta}[\delta_\theta(U)] \leq \sqrt{D_\theta}. \tag{14}$$

The final inequality additionally assumes $D_\theta < \infty$ on the same path-record coordinates.

Appendix B proves the comparison. The ρ_θ-averaged TV term directly measures the conditional path discrepancy; the $\sqrt{D_\theta}$ bound is a coarser trainable proxy because the full path KL also contains bridge-marginal mismatch.

The two terms therefore describe different limitations. If the intermediate configuration nearly determines the region, then $\Omega_{k\ell}$ is small even with exact conditional matching. For two regions, with $w_+ = \pi(A_+)$, $w_- = 1 - w_+$, and $m_+(u) = h_\theta(C_+(u) \mid u)$,

$$\Omega_{+-} = w_+ w_- - \mathrm{Var}_{\rho_\theta}[m_+(U)]. \tag{15}$$

The variance term vanishes when the region label is independent of U and reaches $w_+ w_-$ when the label is determined by U; conditional mismatch then controls the difference between accepted and ideal flow. The ideal kernel is a classical data-augmentation chain, with its spectral characterization in the appendix (Tanner & Wong, 1987; Liu et al., 1994). A related ideal data-augmentation representation and spectral-gap analysis has been developed for diffusion paths (Chen et al., 2026); the comparison above instead concerns the accepted chain when the conditional path distributions differ and the probability flow between regions of a prescribed partition of the target space.

3.3 THE FIRST STEP OF SECTOR RELAXATION

For the deliberately biased initialization $\mu_p = p\pi(\cdot \mid A_+) + (1-p)\pi(\cdot \mid A_-)$, define $p' = (\mu_p K_{\mathrm{RT}})(A_+)$. Stationarity gives equal flows in both directions, hence

$$p' - w_+ = \left(1 - \frac{F^{\mathrm{RT}}_{+-}}{w_+ w_-}\right)(p - w_+). \tag{16}$$

For a symmetric target the factor is $1 - 4F^{\mathrm{RT}}_{+-}$. This is an exact first-step relation for the stated within-sector conditionals. Later distributions need not retain those conditionals, so it does not imply geometric relaxation at all times or a global mixing-time bound.

4 EXPERIMENTS

The experiments test three parts of the construction separately. Many-well targets test whether recorded path weights recover analytic normalizers and mode probabilities. Lattice ϕ^4 isolates the round-trip mechanism by comparing Metropolis acceptance, ideal inter-sector flow, accepted inter-sector flow, and sector relaxation. Two-dimensional gauge theories compare the ESS from four proposal constructions sharing a structured reference and evaluation grid. Parameters are frozen before evaluation. Independent-path estimators and Markov-chain measurements are reported separately; uncertainty estimates and stored-data provenance are specified in the appendices.

4.1 MANY-WELL NORMALIZERS AND MODE PROBABILITIES

The many-well targets provide a direct correction test because their normalizers and unequal sign-mode probabilities are analytic. The four- and eight-dimensional members of the double-well family, DW4 and DW8, have four and sixteen sign modes, respectively; both use $q_0 = \mathcal{N}(0, I)$. Table 1 reports three independent evaluations of each selected frozen model at $p_{\mathrm{flip}} = 0.05$. All modes appear in every evaluation, and the mode error is computed from the same SNIS weights used for normalizer estimation.

Table 1: Independent-path many-well evaluation with a Gaussian reference and fixed deployment probability $p_{\mathrm{flip}} = 0.05$. Each row uses one trained model and three independent evaluation seeds, with N paths per seed. Entries are means $\pm$ sample standard deviations across evaluations. The mode error is computed by SNIS, not from a Markov chain.

Target	N	ESS/N	SNIS mode L_1	$\lvert \log \widehat{Z} - \log Z \rvert$
DW4	200,000	0.4146 ± 0.0038	0.00135 ± 0.00056	0.00052 ± 0.00043
DW8	120,000	0.1499 ± 0.0052	0.00672 ± 0.00285	0.00455 ± 0.00269

The normalized ESS computed from path weights is 0.4146 ± 0.0038 for DW4 and 0.1499 ± 0.0052 for DW8. The corresponding SNIS mode errors are 0.00135 ± 0.00056 and 0.00672 ± 0.00285, while $|\log \widehat{Z} - \log Z|$ is 0.00052 ± 0.00043 and 0.00455 ± 0.00269. These measurements test statistical correction rather than Markov-chain mixing: the recorded path weights recover the normalization and unequal mode masses without evaluating the endpoint proposal density. Earlier-model comparisons, frozen p_{flip} interventions, pathwise sign changes, and the archived DW2 trajectory are given in Appendix D.

4.2 ROUND-TRIP FLOW IN LATTICE ϕ^4

We use a periodic 8×8 lattice with $\lambda = 0.022$, $\kappa = 0.27$, and zero external field. The sectors are defined by the sign of $M = V^{-1} \sum_x \phi_x$. To isolate the round-trip mechanism from neural training, this experiment uses fixed Gaussian momentum laws and changes only a prescribed global sign-flip probability. Starting configurations are sampled from a stored HMC reference bank with equal numbers from both sectors. Each is pulled back through a fresh reverse path and used for one round-trip transition.

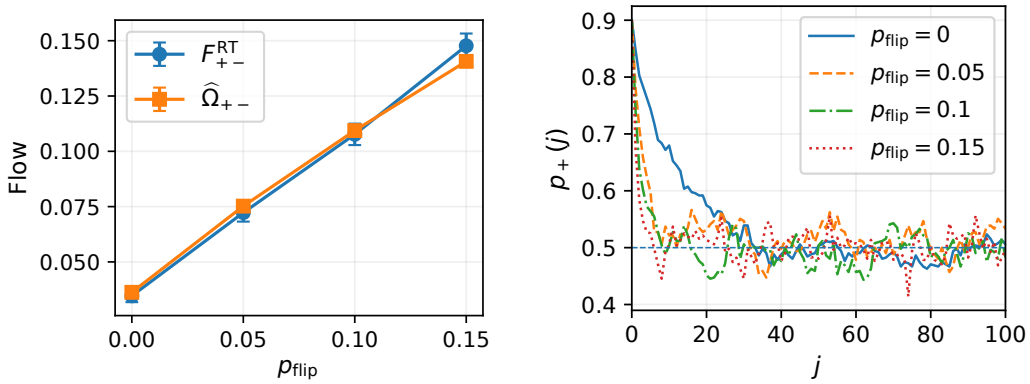

Figure 2: Round-trip flow and sector relaxation in zero-field ϕ^4 with fixed Gaussian momentum laws. Left: accepted stationary flow F^{RT}_{+-} and the nested SNIS estimate $\widehat{\Omega}_{+-}$ of the ideal flow versus the prescribed flip probability p_{flip}. Both flows are joint probabilities, estimated from 4096 reference-bank starts and 32 inner proposals per bridge. Bars show one-standard-error estimates across outer draws only. Mean acceptance is 0.956–0.959 throughout. Right: positive-sector probability $p_+(j)$ after j round-trip transitions, from 512 chains initialized with sector weight approximately 0.9. The dashed horizontal line is $1/2$.

The controlled intervention separates acceptance from transport. For $p_{\mathrm{flip}} = 0, 0.05, 0.10, 0.15$, the measured mean acceptances are $0.9575, 0.9561, 0.9592, 0.9561$, respectively. Over the same intervention, the accepted stationary $+ \to -$ flow increases from $F^{\mathrm{RT}}_{+-} = 0.0347$ to 0.1477. Thus an almost unchanged acceptance probability is compatible with more than a fourfold change in transported probability mass. The corresponding nested-SNIS estimates of the ideal flow are $\widehat{\Omega}_{+-} = 0.0363, 0.0752, 0.1094, 0.1407$ (Fig. 2, left).

The ideal-flow estimator uses 4096 reference-bank starts and 32 forward proposals per intermediate configuration. The displayed uncertainty includes outer-draw variation but not finite-inner-sample bias, so the comparison tests the dependence of the two flows on the intervention rather than exact numerical equality. Appendix E.1 gives the estimator and reference-bank limitations.

A separate experiment initializes 512 chains with positive-sector probability approximately 0.9, while drawing configurations within each sector from the same reference bank. Larger independently measured stationary flow is accompanied by faster recovery toward $1/2$ (Fig. 2, right). Equation (16) describes the first step under exact within-sector conditionals; later steps are measured empirically. The learned κ-scan and earlier finite-volume observable comparisons remain in Appendix E.

4.3 Reference overlap and learned gauge-field transport

The two-dimensional gauge study covers $U(1)$, $\mathrm{SU}(2)$, and $\mathrm{SU}(3)$ on $L = 4, 8, 16, 32$ and integer $\beta = 1, \ldots, 14$. The structured reference samples all but one plaquette from the corresponding single-plaquette Boltzmann distribution and reconstructs the remaining plaquette by torus closure; its normalized density and holonomy variables are given in Appendix F.4.

On the common 56-target grid per gauge group, analytic force plus a learned residual gives higher median ESS/N than analytic force alone or the direct neural path (Table 14, Appendix F.1). The conditional models are shared across the grid within each group. All four $U(1)$ proposals share a post-transport flux move; computational costs differ between proposals. The dependence of reference-distribution ESS on β and L is shown separately in Figs. 20 and 27.

Production path-SNIS estimates of rectangular Wilson loops on 32×32 lattices follow the finite-torus character-expansion references across the tested couplings (Fig. 3); the full observable and normalizer tables are in Appendix F.5.

Compact $U(1)$ also tests topological sampling. The charge Q is the sum of principal plaquette angles divided by 2π, and the topological susceptibility is $\chi_t = \langle Q^2 \rangle / L^2$. At $L = 32$, path-SNIS

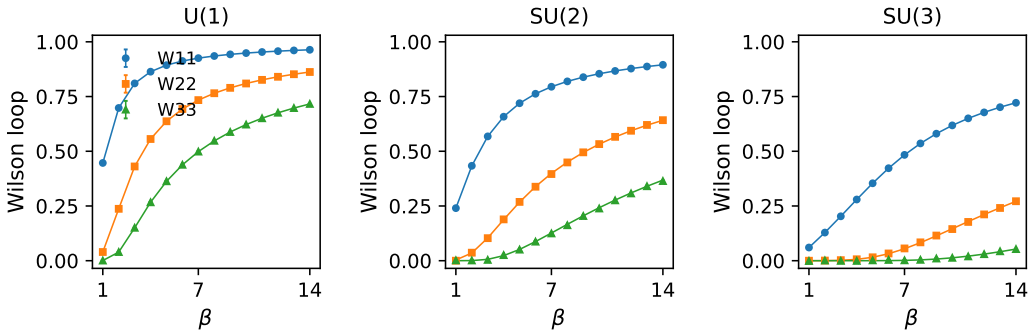

Figure 3: Production Wilson-loop estimates on 32×32 lattices for $U(1)$, $\mathrm{SU}(2)$, and $\mathrm{SU}(3)$. Curves are finite-torus character-expansion references; path-SNIS markers show the recorded one-standard-error estimates. These production pools are distinct from the smaller direct-neural evaluation pools in Table 14.

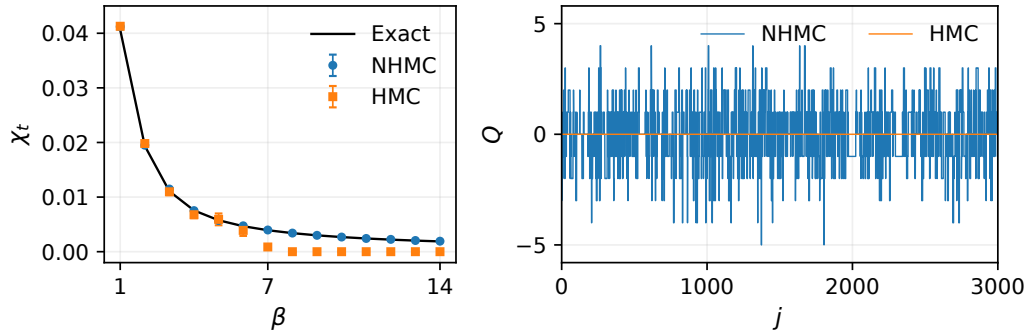

Figure 4: Topological sampling in compact $U(1)$ at $L = 32$. Left: topological susceptibility χ_t, with finite-volume exact references, NHMC path-SNIS estimates, and cold-start HMC estimates; bars show one standard error. The HMC estimates are zero at $\beta = 8, \ldots, 14$ because all 16 chains remain at $Q = 0$ during their 1024 saved states. Right: separate HMC and NHMC path-IMH charge histories at $\beta = 14$, showing all 3000 stored states of each trace. The NHMC chain uses no optional flux move. The index j counts saved Markov updates; computational costs per update differ.

estimates of χ_t follow the finite-volume references, while cold-start HMC estimates fall to zero at the largest couplings (Fig. 4, left). In separate runs at $\beta = 14$, the corrected NHMC chain visits multiple sectors without the optional flux move, while the displayed HMC chain remains at $Q = 0$ (Fig. 4, right). These runs use different computational budgets; chain lengths, costs, autocorrelations, and flux controls are reported in Appendix F.5.3. Separate frozen-depth measurements are given in Appendix F.2.

5 Discussion and Conclusion

NHMC defines tractable probability laws on stochastic Hamiltonian paths using conditional momentum sampling and reversible, volume-preserving dynamics. The complete forward and reverse path probabilities remain available even when the endpoint proposal density is implicit. Their ratio defines the recorded work used for training, self-normalized importance sampling, estimation of normalizing constants and free energies, and Metropolis correction.

Accepted inter-region flow depends on both the conditional region probabilities and the accuracy of the conditional path proposal. Conditional mismatch bounds the deviation from ideal flow, while the full path KL divergence provides a further bound. Even with exact conditional matching, inter-region flow remains small when the intermediate configuration nearly determines the endpoint region.

On many-well targets, the recorded path weights recover known normalizing constants and unequal mode probabilities. In lattice ϕ^4, the Metropolis acceptance changes little while the probability flow between sectors changes substantially, and the estimated flow under exact conditional matching changes in the same direction. In two-dimensional gauge theories, the structured reference improves overlap with the target distribution, while the analytic dynamics and learned components provide additional improvements in the tested systems. Corrected Wilson-loop estimates closely follow the corresponding finite-volume references, and compact $U(1)$ calculations also exhibit accepted transitions between topological sectors.

The theoretical results describe stationary probability flow and the first step of sector relaxation under the stated conditional initialization, while the later relaxation curves are direct measurements of the Markov chain.

AI Use Statement

Language-model tools assisted with mathematical exposition, plotting scripts, reference lookup, manuscript organization, language editing, and LaTeX preparation.

References

David Albandea, Pilar Hernández, Alberto Ramos, and Fernando Romero-López. Topological sampling through windings. *European Physical Journal C*, 81:873, 2021. doi: 10.1140/epjc/s10052-021-09677-6. Erratum: European Physical Journal C 83, 508 (2023), https://doi.org/10.1140/epjc/s10052-023-11683-9.

Michael Arbel, Alexander G. D. G. Matthews, and Arnaud Doucet. Annealed flow transport Monte Carlo. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 318–330. PMLR, 2021. doi: 10.48550/arXiv.2102.07501. URL https://proceedings.mlr.press/v139/arbel21a.html.

J. M. Aroca and Yu. Kubyshin. Calculation of wilson loops in two-dimensional yang–mills theories. *Annals of Physics*, 283(1):11–56, 2000. doi: 10.1006/aphy.2000.6044.

Filippo Ascolani, Gareth O. Roberts, and Giacomo Zanella. Scalability of metropolis-within-gibbs schemes for high-dimensional bayesian models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 88(4):1231–1250, 2026. doi: 10.1093/jrsssb/qkaf084.

Claudio Bonanno, Alessandro Nada, and Davide Vadacchino. Mitigating topological freezing using out-of-equilibrium simulations. *Journal of High Energy Physics*, 2024(4):126, 2024. doi: 10.1007/JHEP04(2024)126.

Claudio Bonanno, Andrea Bulgarelli, Elia Cellini, Alessandro Nada, Dario Panfalone, Davide Vadacchino, and Lorenzo Verzichelli. Scaling flow-based approaches for topology sampling in $SU(3)$ gauge theory. *Journal of High Energy Physics*, 2026(4):051, 2026. doi: 10.1007/JHEP04(2026)051.

Denis Boyda, Gurtej Kanwar, Sébastien Racanière, Danilo Jimenez Rezende, Michael S. Albergo, Kyle Cranmer, Daniel C. Hackett, and Phiala E. Shanahan. Sampling using $SU(N)$ gauge equivariant flows. *Physical Review D*, 103:074504, 2021. doi: 10.1103/PhysRevD.103.074504.

Andrea Bulgarelli, Elia Cellini, and Alessandro Nada. Scaling of stochastic normalizing flows in $SU(3)$ lattice gauge theory. *Physical Review D*, 111:074517, 2025. doi: 10.1103/PhysRevD.111.074517.

Michele Caselle, Gianluca Costagliola, Alessandro Nada, Marco Panero, and Arianna Toniato. Jarzynski's theorem for lattice gauge theory. *Physical Review D*, 94:034503, 2016. doi: 10.1103/PhysRevD.94.034503.

Michele Caselle, Alessandro Nada, and Marco Panero. QCD thermodynamics from lattice calculations with non-equilibrium methods: The $SU(3)$ equation of state. *Physical Review D*, 98:054513, 2018. doi: 10.1103/PhysRevD.98.054513.

Michele Caselle, Elia Cellini, Alessandro Nada, and Marco Panero. Stochastic normalizing flows as non-equilibrium transformations. *Journal of High Energy Physics*, 2022(7):015, 2022. doi: 10.1007/JHEP07(2022)015.

Han Chen, Sifan Liu, and Jun Yang. Markov chain monte carlo with diffusion paths. *arXiv preprint arXiv:2607.11631*, 2026. doi: 10.48550/arXiv.2607.11631. URL https://arxiv.org/abs/2607.11631.

Reuben Cohn-Gordon, Uroš Seljak, and Dries Sels. Counterdiabatic Hamiltonian Monte Carlo. *arXiv preprint arXiv:2602.21272*, 2026. doi: 10.48550/arXiv.2602.21272. URL https://arxiv.org/abs/2602.21272.

Michael Creutz. Gauge fixing, the transfer matrix, and confinement on a lattice. *Physical Review D*, 15(4):1128–1136, 1977. doi: 10.1103/PhysRevD.15.1128. URL https://doi.org/10.1103/PhysRevD.15.1128.

Gavin E. Crooks. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Physical Review E*, 60(3):2721–2726, 1999. doi: 10.1103/PhysRevE.60.2721. URL https://arxiv.org/abs/cond-mat/9901352.

Pierre Del Moral, Arnaud Doucet, and Ajay Jasra. Sequential monte carlo samplers. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 68(3):411–436, 2006. doi: 10.1111/j.1467-9868.2006.00553.x. URL https://doi.org/10.1111/j.1467-9868.2006.00553.x.

Simon Duane, A. D. Kennedy, Brian J. Pendleton, and Duncan Roweth. Hybrid Monte Carlo. *Physics Letters B*, 195(2):216–222, 1987. doi: 10.1016/0370-2693(87)91197-X. URL https://doi.org/10.1016/0370-2693(87)91197-X.

Sam Foreman, Xiao-Yong Jin, and James C. Osborn. Deep learning Hamiltonian Monte Carlo. In *ICLR 2021 Workshop on Deep Learning for Simulation (SimDL)*, 2021. doi: 10.48550/arXiv.2105.03418. URL https://arxiv.org/abs/2105.03418.

W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970. doi: 10.1093/biomet/57.1.97. URL https://doi.org/10.1093/biomet/57.1.97.

Mitch Hill. Exact global MCMC with denoising diffusion. *arXiv preprint arXiv:2609.00279*, 2026. doi: 10.48550/arXiv.2609.00279.

C. Jarzynski. Nonequilibrium equality for free energy differences. *Physical Review Letters*, 78(14):2690–2693, 1997. doi: 10.1103/PhysRevLett.78.2690. URL https://doi.org/10.1103/PhysRevLett.78.2690.

Leon Klein, Andreas Krämer, and Frank Noé. Equivariant flow matching. In *Advances in Neural Information Processing Systems*, volume 36, 2023. doi: 10.48550/arXiv.2306.15030. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/bc827452450356f9f558f4e4568d553b-Abstract-Conference.html.

Daniel Levy, Matthew D. Hoffman, and Jascha Sohl-Dickstein. Generalizing Hamiltonian Monte Carlo with neural networks. In *International Conference on Learning Representations*, 2018. doi: 10.48550/arXiv.1711.09268. URL https://arxiv.org/abs/1711.09268.

Jun S. Liu, Wing Hung Wong, and Augustine Kong. Covariance structure of the Gibbs sampler with applications to the comparisons of estimators and augmentation schemes. *Biometrika*, 81(1):27–40, 1994. doi: 10.1093/biomet/81.1.27. URL https://doi.org/10.1093/biomet/81.1.27.

Martin Lüscher. Topology of Lattice Gauge Fields. *Communications in Mathematical Physics*, 85:39–48, 1982. doi: 10.1007/BF02029132. URL https://doi.org/10.1007/BF02029132.

Martin Lüscher and Stefan Schaefer. Lattice QCD without topology barriers. *Journal of High Energy Physics*, 2011(7):036, 2011. doi: 10.1007/JHEP07(2011)036.

Kerrie L. Mengersen and Richard L. Tweedie. Rates of convergence of the Hastings and Metropolis algorithms. *The Annals of Statistics*, 24(1):101–121, 1996. doi: 10.1214/aos/1033066201. URL https://doi.org/10.1214/aos/1033066201.

Radford M. Neal. Sampling from multimodal distributions using tempered transitions. *Statistics and Computing*, 6(4):353–366, 1996. doi: 10.1007/BF00143556.

Radford M. Neal. Annealed importance sampling. *Statistics and Computing*, 11(2):125–139, 2001. doi: 10.1023/A:1008923215028. URL https://doi.org/10.1023/A:1008923215028.

Kirill Neklyudov, Max Welling, Evgenii Egorov, and Dmitry Vetrov. Involutive MCMC: A unifying framework. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 7273–7282. PMLR, 2020. URL https://proceedings.mlr.press/v119/neklyudov20a.html.

Kim A. Nicoli, Christopher J. Anders, Tobias Hartung, Karl Jansen, Pan Kessel, and Shinichi Nakajima. Detecting and mitigating mode-collapse for flow-based sampling of lattice field theories. *Physical Review D*, 108:114501, 2023. doi: 10.1103/PhysRevD.108.114501.

Jerome P. Nilmeier, Gavin E. Crooks, David D. L. Minh, and John D. Chodera. Nonequilibrium Candidate Monte Carlo is an efficient tool for equilibrium simulation. *Proceedings of the National Academy of Sciences*, 108(45):E1009–E1018, 2011. doi: 10.1073/pnas.1106094108. URL https://doi.org/10.1073/pnas.1106094108.

Peter Potaptchik, Cheuk-Kit Lee, and Michael S. Albergo. Tilt matching for scalable sampling and fine-tuning. *arXiv preprint arXiv:2512.21829*, 2025. doi: 10.48550/arXiv.2512.21829. URL https://arxiv.org/abs/2512.21829.

Qian Qin, Nianqiao Ju, and Guanyang Wang. Spectral gap bounds for reversible hybrid gibbs chains. *The Annals of Statistics*, 53(4):1613–1638, 2025. doi: 10.1214/25-AOS2516.

Jascha Sohl-Dickstein and Benjamin J. Culpepper. Hamiltonian annealed importance sampling for partition function estimation. *arXiv preprint arXiv:1205.1925*, 2012. doi: 10.48550/arXiv.1205.1925. URL https://arxiv.org/abs/1205.1925.

Martin A. Tanner and Wing Hung Wong. The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, 82(398):528–540, 1987. doi: 10.1080/01621459.1987.10478458. URL https://doi.org/10.1080/01621459.1987.10478458.

Michalis K. Titsias and Omiros Papaspiliopoulos. Auxiliary gradient-based sampling algorithms. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 80(4):749–767, 2018. doi: 10.1111/rssb.12269.

Hao Wu, Jonas Köhler, and Frank Noé. Stochastic normalizing flows. In *Advances in Neural Information Processing Systems*, volume 33, 2020. doi: 10.48550/arXiv.2002.06707. URL https://proceedings.neurips.cc/paper/2020/hash/41d80bfc327ef980528426fc810a6d7a-Abstract.html.

Qinsheng Zhang and Yongxin Chen. Path integral sampler: A stochastic control approach for sampling. In *International Conference on Learning Representations*, 2022. doi: 10.48550/arXiv.2111.15141. URL https://openreview.net/forum?id=_uCb2ynRu7Y.

A Path Measures and Correction Arguments

A.1 Forward and Reverse Path Laws

Throughout this appendix the configuration-space target is represented by an unnormalized density $\gamma(x) = \exp[-E(x)]$ with a finite positive normalizing constant Z, where $0 < Z = \int \gamma(x)\,\mu(dx) < \infty$. The learned parameter θ is fixed during evaluation. The arguments depend on the fixed proposal law used at evaluation time, not on the particular training objective. Endpoint corrections require a known endpoint proposal density and $\pi \ll q_\theta$; path corrections require a computable Radon–Nikodym weight between the fixed forward path law Q_θ and a reference reverse path law P_θ whose configuration endpoint marginal is π. Deterministic maps are invertible and either volume preserving or accompanied by their exact Jacobian determinants. For discrete moves, $s \mapsto s^\dagger$ is a bijection of the label space and satisfies $g_{s^\dagger} = g_s^{-1}$. Every stochastic choice appears through its forward and reverse transition probabilities.

The main construction is the forward–reverse path ratio. Endpoint SNIS and endpoint IMH are the corresponding special cases when the endpoint proposal density is available. Endpoint-SNIS and path-SNIS supply consistent weighted estimates through their weighted empirical measures. Endpoint-IMH and path-IMH define kernels that preserve their respective target laws, while finite-chain efficiency still depends on proposal overlap, acceptance rate, and autocorrelation. Round-trip NHMC-MH directly preserves the configuration-space target by Theorem 2.

A.2 Detailed Derivation of Theorem 1

The main theorem follows from the probability ratio between the explicit forward proposal path density and the reference reverse path density. All path densities below are taken with respect to the common path-coordinate measure induced by the deterministic maps. Equivalently, a forward path is parameterized by $(x_0, s_0, p_0, \dots, s_{T-1}, p_{T-1})$, with absent s_t variables omitted and all later configurations and output momenta obtained deterministically. Write the displayed path as

$$\Gamma = \left(x_0, \{s_t, \widetilde{x}_t, p_t, x_{t+1}, \bar{p}_{t+1}\}_{t=0}^{T-1}\right),$$

where $x_0 \sim q_0$, $s_t \sim r_{\theta,t}^F(\cdot \mid x_t)$, $\widetilde{x}_t = g_{s_t}(x_t)$, $p_t \sim q_{\theta,t}^F(\cdot \mid \widetilde{x}_t)$, and $(x_{t+1}, \bar{p}_{t+1}) = \Phi_{\theta,t}(\widetilde{x}_t, p_t)$. Here p_t is the sampled input momentum at stage t, while $\bar{p}_{t+1}$ is the deterministic map output. The optional g_{s_t} maps are invertible and volume preserving; the DW sign flip is an involution with unit absolute Jacobian determinant. The maps $\Phi_{\theta,t}$ are invertible and volume preserving for the leapfrog updates used in the reported NHMC paths:

Lemma 1 (Kick–drift–kick invertibility and reference-measure preservation)**.** *For the deterministic maps used in the reported runs, suppose the step size is state independent (it may depend on the stage and fixed external parameters). Then each kick–drift–kick stage is invertible and preserves the relevant position–momentum reference measure. With momentum reversal $\mathcal{R}(x,p) = (x,-p)$, the palindromic stage satisfies $\Phi_{\theta,t}^{-1} = \mathcal{R} \circ \Phi_{\theta,t} \circ \mathcal{R}$.*

Proof. For Euclidean configurations, a momentum half-step keeps x fixed and translates p by a function of x, while the position drift keeps p fixed and translates x by a function of p. These are triangular shears with unit Jacobian determinant. For compact gauge groups in the Hermitian momentum convention, the kick translates the Lie-algebra coordinate P, and the drift $U \mapsto \exp(i\varepsilon P)U$ is left multiplication by a fixed group element conditional on P. The kick preserves Lebesgue measure on the momentum coordinates and the drift preserves Haar measure on the links. Reversing the step sign and the momentum in the opposite palindromic order therefore gives the inverse in both settings. A genuinely state-dependent drift rate would require a separate Jacobian/reversibility analysis and is not part of the reported maps. □

For non-volume-preserving deterministic maps, the Jacobian terms enter the density ratio described in the deterministic-map discussion below. With the volume-preserving maps used here, the forward path density is

$$Q_\theta(d\Gamma) = q_0(x_0) \prod_{t=0}^{T-1} r_{\theta,t}^F(s_t \mid x_t) q_{\theta,t}^F(p_t \mid \widetilde{x}_t)\, d\lambda_\theta(\Gamma), \tag{17}$$

where $d\lambda_\theta$ denotes the common path-coordinate base measure induced by the deterministic maps.

Now define a reference reverse path law by first drawing the endpoint from the Boltzmann target, $x_T \sim \pi(x) = \gamma(x)/Z$. For a realized forward stage satisfying

$$\Phi_{\theta,t}(\widetilde{x}_t, p_t) = (x_{t+1}, \bar{p}_{t+1}), \tag{18}$$

time reversibility gives

$$\Phi_{\theta,t}(x_{t+1}, -\bar{p}_{t+1}) = (\widetilde{x}_t, -p_t). \tag{19}$$

Thus the reverse input momentum associated with forward stage t is $-\bar{p}_{t+1}$; applying $g_{s_t}^{-1}$ after the inverse leapfrog reconstructs x_t. The reference reverse density on the same path coordinates is

$$P_\theta(d\Gamma) = \frac{\gamma(x_T)}{Z} \prod_{t=0}^{T-1} r_{\theta,t}^R(s_t^\dagger \mid x_{t+1}) q_{\theta,t}^R(-\bar{p}_{t+1} \mid x_{t+1})\, d\lambda_\theta(\Gamma). \tag{20}$$

It is normalized because π, the reverse discrete laws, and the reverse momentum laws are normalized and the inverse deterministic reconstructions preserve volume. Since the construction first draws $x_T \sim \pi$ and then conditions all other path variables on this endpoint, integrating over the non-endpoint variables leaves endpoint marginal $\pi(dx_T)$.

The ratio of the two path densities is therefore

$$\log \frac{dQ_\theta}{dP_\theta}(\Gamma) = \log q_0(x_0) - \log \gamma(x_T) + \log Z \tag{21}$$

$$+ \sum_{t=0}^{T-1} \left[\log r_{\theta,t}^F(s_t \mid x_t) - \log r_{\theta,t}^R(s_t^\dagger \mid x_{t+1})\right]$$

$$+ \sum_{t=0}^{T-1} \left[\log q_{\theta,t}^F(p_t \mid \widetilde{x}_t) - \log q_{\theta,t}^R(-\bar{p}_{t+1} \mid x_{t+1})\right] \tag{22}$$

$$= W_\theta(\Gamma) + \log Z. \tag{23}$$

Equivalently,

$$\frac{dP_\theta}{dQ_\theta}(\Gamma) = \frac{\exp[-W_\theta(\Gamma)]}{Z}. \tag{24}$$

Thus $\widetilde{w}_\theta(\Gamma) = \exp[-W_\theta(\Gamma)]$ is an unnormalized path correction weight, and integrating the last display under Q_θ gives $\mathbb{E}_{Q_\theta}[\exp[-W_\theta]] = Z$. This completes the detailed derivation of Theorem 1.

A.3 Path Reweighting Identity

Let $Q_\theta(d\Gamma)$ be the learned forward path measure and let $P_\theta(d\Gamma)$ be the reverse reference path law constructed above. Define $\widetilde{w}_\theta(\Gamma) = \exp[-W_\theta(\Gamma)]$. Assume $0 < \mathbb{E}_{Q_\theta}[\widetilde{w}_\theta] < \infty$ and $\mathbb{E}_{Q_\theta}[\widetilde{w}_\theta |O(x_T)|] < \infty$. Since

$$P_\theta(d\Gamma) = \frac{\widetilde{w}_\theta(\Gamma)}{\mathbb{E}_{Q_\theta}[\widetilde{w}_\theta]} Q_\theta(d\Gamma), \tag{25}$$

then

$$\mathbb{E}_\pi[O] = \mathbb{E}_{P_\theta}[O(x_T)] \tag{26}$$

$$= \frac{\mathbb{E}_{Q_\theta}[\widetilde{w}_\theta(\Gamma) O(x_T)]}{\mathbb{E}_{Q_\theta}[\widetilde{w}_\theta(\Gamma)]}. \tag{27}$$

The empirical SNIS estimator follows by applying the strong law to numerator and denominator.

A.4 Detailed balance for path-IMH

Let $P_\theta(d\Gamma) \propto \widetilde{w}_\theta(\Gamma) Q_\theta(d\Gamma)$ with $0 < \mathbb{E}_{Q_\theta}[\widetilde{w}_\theta] < \infty$ and positive finite weights on the sampled support. Propose $\Gamma' \sim Q_\theta$. The independent path acceptance probability is

$$\alpha_\theta(\Gamma, \Gamma') = \min\left\{1, \frac{\widetilde{w}_\theta(\Gamma')}{\widetilde{w}_\theta(\Gamma)}\right\}. \tag{28}$$

The accepted transition measure is

$$P_\theta(d\Gamma)Q_\theta(d\Gamma')\alpha_\theta(\Gamma,\Gamma') \propto Q_\theta(d\Gamma)Q_\theta(d\Gamma')\min\{\widetilde{w}_\theta(\Gamma),\widetilde{w}_\theta(\Gamma')\}, \tag{29}$$

which is symmetric in Γ and Γ'. Self-loops handle rejections, so P_θ is invariant. Since the constructed P_θ has configuration endpoint marginal π, stationary x_T endpoints are Boltzmann distributed. This argument proves invariance. Convergence from an arbitrary initialization requires the usual irreducibility and aperiodicity conditions for the independent Metropolis kernel.

A.5 Round-Trip NHMC-MH Proof

Let X be configuration space with sigma-finite reference measure μ, and let A be the product space of all stochastic labels and momenta along a path, with sigma-finite product reference measure m. The forward and reverse auxiliary laws below are assumed to admit normalized densities with respect to these reference measures. A forward record is denoted by a and the associated reverse record by b.

For one stage, an optional discrete, measure-preserving map g_{t,s_t} is followed by a reversible, volume-preserving kick–drift map $\Phi_{\theta,t}$:

$$\widetilde{x}_t = g_{t,s_t}(x_t), \tag{30}$$

$$(x_{t+1},\bar{p}_{t+1}) = \Phi_{\theta,t}(\widetilde{x}_t,p_t), \tag{31}$$

$$p_t^\dagger = -\bar{p}_{t+1}. \tag{32}$$

The inverse label satisfies $g_{t,s_t^\dagger} = g_{t,s_t}^{-1}$. Since $\Phi_{\theta,t}^{-1} = \mathcal{R}\circ\Phi_{\theta,t}\circ\mathcal{R}$ for $\mathcal{R}(x,p) = (x,-p)$, the inverse stage is explicit:

$$(\widetilde{x}_t,-p_t) = \Phi_{\theta,t}(x_{t+1},p_t^\dagger), \tag{33}$$

$$x_t = g_{t,s_t}^{-1}(\widetilde{x}_t). \tag{34}$$

Every factor in this stage map is invertible and preserves its reference measure. Composition over the stages therefore gives a bimeasurable, measure-preserving bijection

$$\Psi_\theta : (u,a) \longmapsto (x,b) \tag{35}$$

from $(\mathsf{X}\times\mathsf{A},\mu\otimes m)$ to itself. This is the coordinate statement needed below; a non-volume-preserving implementation would instead require its exact Jacobian in the acceptance ratio.

Let $f_\theta(a \mid u)$ and $r_\theta(b \mid x)$ be the normalized forward and reverse auxiliary densities. Given the current x, the transition draws

$$b\sim r_\theta(\cdot\mid x), \qquad (u,a) = \Psi_\theta^{-1}(x,b), \qquad a'\sim f_\theta(\cdot\mid u), \qquad (y,b') = \Psi_\theta(u,a'). \tag{36}$$

Figure 5 summarizes this round-trip construction and the single acceptance decision made after both path maps have been evaluated.

For $z = (x,b,a')$, write $(u(x,b),a(x,b)) = \Psi_\theta^{-1}(x,b)$ and define the augmented density, with respect to $\nu = \mu\otimes m\otimes m$, by

$$\xi_\theta(z) = \pi(x)r_\theta(b\mid x)f_\theta(a'\mid u(x,b)). \tag{37}$$

Successive integration over a' and b shows that ξ_θ is normalized and that its x-marginal is π. No assumption on the induced bridge distribution is needed.

Define the path-swap map

$$T_\theta(x,b,a') = (y,b',a) = (\Psi_\theta\times I)\circ\mathrm{Swap}\circ(\Psi_\theta^{-1}\times I)(x,b,a'), \tag{38}$$

where $\mathrm{Swap}(u,a,a') = (u,a',a)$. This conjugation immediately gives $T_\theta^2 = I$. Moreover, Ψ_θ, its inverse, and the swap all preserve their respective product reference measures, so $(T_\theta)_\#\nu = \nu$.

For $z\sim\xi_\theta$, the denominator $\xi_\theta(z)$ is positive almost surely. The direct Metropolis ratio is therefore defined ξ_θ-almost everywhere; if $\xi_\theta(T_\theta z) = 0$, the acceptance probability is defined to be zero. If the measures with densities $\xi_\theta(z)$ and $\xi_\theta(T_\theta z)$ are mutually absolutely continuous, the displayed densities are finite almost everywhere on their common support, and $q_0(u)$ is finite and positive on the induced bridge support, then the work-difference representation below is finite.

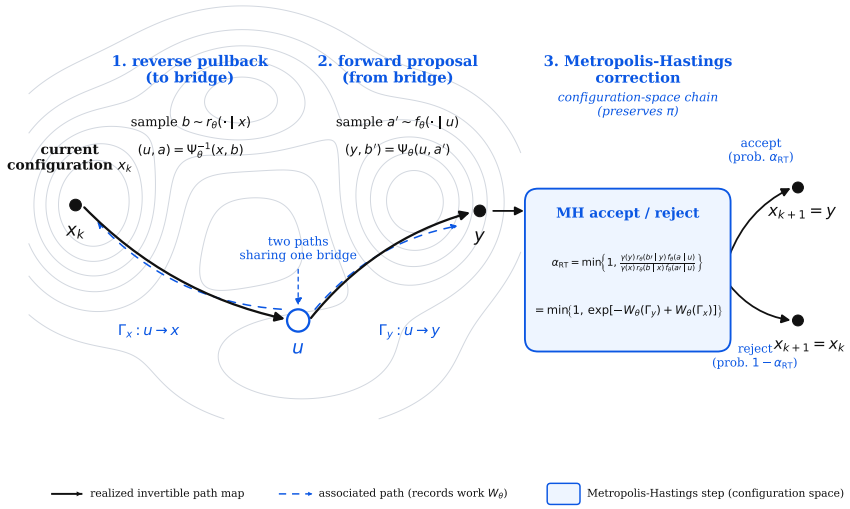

Figure 5: Round-trip NHMC-MH transition. Starting from x_k, the algorithm traverses the forward-oriented path $\Gamma_x : u \to x_k$ in reverse to recover u, then follows $\Gamma_y : u \to y$ to the candidate. One Metropolis decision accepts y or retains x_k. The curves are realized invertible path maps conditional on the sampled auxiliary records; their forward orientations define the two work values.

Bijectivity gives $\Psi_\theta^{-1}(y, b') = (u, a')$. Hence

$$\frac{\xi_\theta(T_\theta z)}{\xi_\theta(z)} = \frac{\pi(y) r_\theta(b' \mid y) f_\theta(a \mid u)}{\pi(x) r_\theta(b \mid x) f_\theta(a' \mid u)} \tag{39}$$

$$= \frac{\gamma(y) r_\theta(b' \mid y) f_\theta(a \mid u)}{\gamma(x) r_\theta(b \mid x) f_\theta(a' \mid u)}. \tag{40}$$

The unknown target normalizer cancels. With the Metropolis probability $\alpha_\theta(z) = 1 \wedge \xi_\theta(T_\theta z)/\xi_\theta(z)$, the accepted transition measure has density

$$\xi_\theta(z)\alpha_\theta(z) = \min\{\xi_\theta(z), \xi_\theta(T_\theta z)\}. \tag{41}$$

For measurable $A, B \subseteq \mathsf{X}$, the accepted transition measure from A to B is therefore

$$J(A, B) = \int \mathbf{1}_A(x(z)) \mathbf{1}_B(x(T_\theta z)) \min\{\xi_\theta(z), \xi_\theta(T_\theta z)\} \, \nu(\mathrm{d}z). \tag{42}$$

The change of variables $w = T_\theta z$, together with $T_\theta^2 = I$ and $(T_\theta)_\# \nu = \nu$, gives $J(A, B) = J(B, A)$. Rejections remain at the current configuration and are symmetric on the diagonal. Thus the projected kernel satisfies detailed balance with respect to π, proving Theorem 2.

Finally, under the mutual absolute continuity condition, define for the two forward-oriented realized paths

$$W_x = \log q_0(u) - \log \gamma(x) + \log f_\theta(a \mid u) - \log r_\theta(b \mid x), \tag{43}$$

$$W_y = \log q_0(u) - \log \gamma(y) + \log f_\theta(a' \mid u) - \log r_\theta(b' \mid y). \tag{44}$$

Because both paths share the same u, subtracting these expressions cancels $\log q_0(u)$ and turns Eq. (40) into

$$\alpha_\theta = 1 \wedge \exp[-W_y + W_x]. \tag{45}$$

This equality explains why independent path-IMH and round-trip NHMC-MH have similar-looking work-difference formulas even though they act on different state spaces.

A.6 Work, Free Energy, and Path-Space Kullback–Leibler Divergence

The path ratio gives the nonequilibrium work identities (Jarzynski, 1997; Crooks, 1999). With $\Delta F = -\log Z$ and endpoint marginal $q_{\theta,T}$ of Q_θ,

$$D_{\mathrm{KL}}(Q_\theta \| P_\theta) = \mathbb{E}_{Q_\theta}[W_\theta] - \Delta F, \qquad D_{\mathrm{KL}}(q_{\theta,T} \| \pi) \le D_{\mathrm{KL}}(Q_\theta \| P_\theta). \tag{46}$$

The equality follows by taking the expectation of $\log(dQ_\theta/dP_\theta) = W_\theta - \Delta F$; the inequality follows by marginalizing the path. Since ΔF is fixed by the target and base convention, minimizing mean work minimizes this upper bound on endpoint KL.

For independent paths $\Gamma_i \sim Q_\theta$, the identity $\mathbb{E}_{Q_\theta}[e^{-W_\theta}] = Z$ gives

$$\widehat{Z} = \frac{1}{N}\sum_{i=1}^{N} e^{-W_i}, \qquad \widehat{\Delta F} = -\log \widehat{Z}. \tag{47}$$

Under the support and integrability conditions above, $\widehat{Z}$ is unbiased and converges almost surely to $Z > 0$. The free-energy estimate converges almost surely to ΔF and generally has finite-sample bias. Multiplying all weights by a common positive constant leaves SNIS and path-IMH unchanged; estimating Z uses the unscaled weights e^{-W_i}.

A.7 Jacobian and Volume Preservation

For a deterministic map used in a path proposal, unit Jacobian determinant gives

$$\left|\det \frac{\partial z_T}{\partial z_0}\right| = 1 \tag{48}$$

and contributes no additional density term. If $z_{k+1} = T_{k,\theta}(z_k)$ and $J_{k,\theta} = |\det \partial T_{k,\theta}(z_k)/\partial z_k|$, a deterministic work convention is

$$W_\theta(\Gamma) = \sum_{k=0}^{T-1} \left[H_{k+1}(z_{k+1}) - H_k(z_k) - \log J_{k,\theta}\right]. \tag{49}$$

This is a generic deterministic-map convention. In the NHMC runs reported here, the deterministic kick–drift maps are volume preserving, so $J_{k,\theta} = 1$ and the Jacobian term vanishes; the stochastic momentum refreshments contribute instead through the forward-minus-reverse log-density terms in Eq. (5).

A.8 Endpoint Corrections

When the endpoint proposal density q_θ is available and $\pi \ll q_\theta$, set $w_\theta(x) = \gamma(x)/q_\theta(x)$. The independence Metropolis acceptance is $\alpha_\theta(x, y) = 1 \wedge w_\theta(y)/w_\theta(x)$ on the positive-weight support (Hastings, 1970; Mengersen & Tweedie, 1996). Detailed balance follows from

$$\pi(x)q_\theta(y)\alpha_\theta(x, y) = \frac{q_\theta(x)q_\theta(y)}{Z}\min\{w_\theta(x), w_\theta(y)\}, \tag{50}$$

which is symmetric in x, y; rejection contributes a diagonal self-loop. For independent draws $X_i \sim q_\theta$ and π-integrable O,

$$\mathbb{E}_{q_\theta}[w_\theta] = Z, \qquad \mathbb{E}_{q_\theta}[w_\theta O] = Z\mathbb{E}_\pi[O]. \tag{51}$$

The strong law therefore gives the same SNIS and normalizer consistency results as for paths: the sample mean of w_θ is unbiased for Z, while its logarithm is generally biased at finite sample size and converges almost surely to $\log Z$ once the sample mean is positive. A normalized auxiliary law $\nu(p)$ also gives $\int \pi(x)\nu(p)\,dp = \pi(x)$, so marginalizing a stationary extended-state kernel recovers the target configuration law. All these statements use fixed proposal parameters, as specified at the start of this appendix; online adaptation requires additional conditions.

A.9 Symmetry-Structured Proposal Densities

For a finite group G of differentiable, volume-preserving transformations,

$$q_\theta^G(x) = \frac{1}{|G|} \sum_{g \in G} q_\theta(g^{-1}x) \tag{52}$$

is normalized by a change of variables in each term. If $\pi \ll q_\theta^G$, endpoint correction applies with weight $\gamma(x)/q_\theta^G(x)$, including when the target breaks the group symmetry. The tilted DW target retains its unequal target mode masses.

A.10 Weighted Empirical Estimates and Effective Sample Size

For nonnegative weights with $0 < \sum_i w_i < \infty$, let $\bar{w}_i = w_i / \sum_j w_j$ and $\widehat{\mu}_N^w = \sum_i \bar{w}_i \delta_{x_i}$, where δ_x is the unit point mass at x. Endpoint weights use $w_i = \gamma(x_i)/q_\theta(x_i)$; path weights use $w_i = e^{-W_i}$ and $x_i = x_T^{(i)}$. Under the support and first-moment conditions above,

$$\int O \, d\widehat{\mu}_N^w = \sum_i \bar{w}_i O(x_i) \xrightarrow{\text{a.s.}} \mathbb{E}_\pi[O]. \tag{53}$$

Mode probabilities and histogram bins follow by taking O to be the corresponding indicator. Resampled visualizations draw from $\widehat{\mu}_N^w$, the finite-sample empirical approximation to π. The weight-based effective sample size is

$$\mathrm{ESS}_{\mathrm{SNIS}} = \frac{1}{\sum_i \bar{w}_i^2}, \qquad \frac{\mathrm{ESS}_{\mathrm{SNIS}}}{N} = \frac{(\sum_i w_i)^2}{N \sum_i w_i^2}. \tag{54}$$

The effective sample size $\mathrm{ESS}_{\mathrm{SNIS}}$ lies in $[1, N]$ and equals N when all weights are equal. Both $\mathrm{ESS}_{\mathrm{SNIS}}$ and $\mathrm{ESS}_{\mathrm{SNIS}}/N$ are invariant under common positive rescaling. These quantities describe weight concentration; the chain effective sample size $N/(2\tau_{\mathrm{int}})$ uses the integrated autocorrelation time τ_{int}.

B Round-Trip Transition Analysis

This appendix studies the round-trip kernel of Section 3.1 beyond target reversibility. Forward–backward nonequilibrium path moves and whole-path Metropolis corrections appear in tempered transitions, NCMC, and recent diffusion-path samplers (Neal, 1996; Nilmeier et al., 2011; Chen et al., 2026; Hill, 2026). The question considered here is how a correct forward–backward kernel moves between prescribed regions of configuration space.

The analysis uses the conditional distribution of a path record given the intermediate configuration. In the ideal limit it reduces to the classical data-augmentation identities of Tanner & Wong (1987); Liu et al. (1994); conditional independence-Metropolis and hybrid-Gibbs results provide the corresponding comparison tools away from that limit (Mengersen & Tweedie, 1996; Qin et al., 2025; Ascolani et al., 2026). We derive the stationary transition probability between specified regions and then bound its deviation when the conditional path laws differ.

B.1 Common coordinates for the intermediate configuration

Fix the trained parameters θ throughout this appendix. We use the same auxiliary-variable representation as Appendix A.5. Let X be the configuration space with a sigma-finite reference measure μ, and let A be the complete auxiliary-record space with a sigma-finite reference measure m; both are standard Borel. In the NHMC paths the bridge and endpoint use the same configuration space and the forward and reverse records use isomorphic record spaces, so we write a single pair (X, A). A more general formulation may use distinct but measurably isomorphic bridge/endpoint and forward/reverse record spaces.

The realized reversible path construction induces a bimeasurable bijection

$$\Psi_\theta : (u, a) \longleftrightarrow (x, b) \tag{55}$$

that preserves the product reference measure $\mu \otimes m$. If a non-volume-preserving path map were used, the corresponding Radon–Nikodym/Jacobian factor must be retained.

Let Q_0 be the normalized bridge/source law with density q_0 with respect to μ. Let $f_\theta(\cdot \mid u)$ and $r_\theta(\cdot \mid x)$ be normalized forward- and reverse-record probability kernels. Whenever density notation is used below, these kernels are assumed absolutely continuous with respect to m, and the same symbols denote their densities. Standard-Borel assumptions guarantee regular conditional probabilities; they do not by themselves imply the existence of these densities. Define the forward augmented measure in bridge coordinates by

$$Q_\theta(du, da) := Q_0(du) f_\theta(da \mid u). \tag{56}$$

Let

$$\mathsf{E}_\theta(u, a) := \mathrm{pr}_{\mathsf{X}}\, \Psi_\theta(u, a) \tag{57}$$

be the endpoint map and denote its forward marginal by $q_{\theta,T} = (\mathsf{E}_\theta)_\# Q_\theta$.

The target reverse augmented measure is

$$\mathsf{P}_\theta^{XB}(dx, db) := \pi(dx) r_\theta(db \mid x). \tag{58}$$

Pull it back to the same (u, a) coordinates as (56):

$$P_\theta := (\Psi_\theta^{-1})_\# \mathsf{P}_\theta^{XB}. \tag{59}$$

By construction,

$$(\mathsf{E}_\theta)_\# P_\theta = \pi. \tag{60}$$

When π has a density with respect to μ and the reverse kernel has a density with respect to m, measure preservation gives the $\mu \otimes m$ density

$$\overline{p}_\theta(u, a) = \pi(\mathsf{E}_\theta(u, a)) r_\theta(b(u, a) \mid \mathsf{E}_\theta(u, a)). \tag{61}$$

The measure-level definitions remain primary; density formulas are used only under these absolute-continuity assumptions. Whenever a finite forward KL divergence is used, assume $Q_\theta \ll P_\theta$ and define

$$D_\theta := D_{\mathrm{KL}}(Q_\theta \| P_\theta). \tag{62}$$

Under the stronger common-support assumptions of Theorem 1, this is the same path-space KL as in Section 3, because the coordinate encoding is bijective. In particular,

$$D_\theta = \mathbb{E}_{Q_\theta}[W_\theta] - \Delta F, \qquad \Delta F = -\log Z. \tag{63}$$

B.2 PATH KL CONTROLS ENDPOINT AND REGION MASSES

Let $\{A_k\}_{k=1}^K$ be a measurable partition of configuration space up to null sets, and define

$$p_k := q_{\theta,T}(A_k), \qquad w_k := \pi(A_k). \tag{64}$$

Proposition 1 (Endpoint and region-mass contraction)**.** *Under the assumptions above,*

$$D_{\mathrm{KL}}(q_{\theta,T}\|\pi) \le D_\theta, \tag{65}$$

$$D_{\mathrm{KL}}(\boldsymbol{p}\|\boldsymbol{w}) \le D_\theta, \tag{66}$$

$$\|\boldsymbol{p}-\boldsymbol{w}\|_1 \le \sqrt{2D_\theta}, \tag{67}$$

where $\boldsymbol{p} = (p_1, \dots, p_K)$ and $\boldsymbol{w} = (w_1, \dots, w_K)$. For each individual region,

$$\mathrm{kl}(p_k\|w_k) \le D_\theta. \tag{68}$$

Consequently, if $p_k = 0$ while $0 < w_k < 1$, then

$$D_\theta \ge -\log(1-w_k). \tag{69}$$

Proof. The endpoint map E_θ is measurable and its pushforwards are $q_{\theta,T}$ and π under Q_θ and P_θ, respectively. Data processing for relative entropy therefore gives (65). Applying data processing once more to the measurable region label $M(x) = k$ for $x \in A_k$ gives (66). Pinsker's inequality, with $d_{\mathrm{TV}}(P,Q) = \frac{1}{2}\|P-Q\|_1$, gives (67). Applying data processing to the binary indicator $\mathbf{1}_{A_k}$ gives (68). Finally, $\mathrm{kl}(0\|w_k) = -\log(1-w_k)$, proving (69). □

Scope of the bound. For a rare region, $-\log(1-w_k) = w_k + O(w_k^2)$. Thus a small forward KL controls aggregate probability-mass error but is not a uniform support guarantee for arbitrarily rare regions. More importantly, even $D_\theta = 0$ does not imply rapid Markov-chain mixing; dependence on the intermediate configuration enters separately below.

B.3 TARGET-INDUCED CONDITIONAL LAW AND INDEPENDENCE METROPOLIS–HASTINGS

Because the spaces are standard Borel, P_θ admits a regular conditional disintegration

$$P_\theta(du, da) = \rho_\theta(du) h_\theta(da \mid u), \tag{70}$$

where ρ_θ is the target-induced bridge marginal and $h_\theta(\cdot \mid u)$ is the target-induced conditional probability kernel for the forward record at fixed bridge. When densities with respect to μ and m exist, we use $\rho_\theta(u)$ and $h_\theta(a \mid u)$ for those densities. In this shorthand,

$$Q_\theta = q_0 f_\theta, \qquad P_\theta = \rho_\theta h_\theta. \tag{71}$$

Lemma 2 (Chain-rule decomposition of the path KL)**.** *If $D_\theta < \infty$, then $Q_0 \ll \rho_\theta$ and $f_\theta(\cdot \mid u) \ll h_\theta(\cdot \mid u)$ for Q_0-almost every u. Moreover,*

$$D_\theta = D_{0,\theta} + D_{1,\theta}, \tag{72}$$

where

$$D_{0,\theta} := D_{\mathrm{KL}}(Q_0\|\rho_\theta), \tag{73}$$

$$D_{1,\theta} := \int Q_0(du)\, D_{\mathrm{KL}}(f_\theta(\cdot \mid u)\|h_\theta(\cdot \mid u))\,. \tag{74}$$

Proof. Finite D_θ implies $Q_\theta \ll P_\theta$. Marginalization gives $Q_0 \ll \rho_\theta$, and disintegration gives conditional absolute continuity for Q_0-almost every bridge. On this set the Radon–Nikodym derivative factorizes as

$$\frac{dQ_\theta}{dP_\theta}(u,a) = \frac{dQ_0}{d\rho_\theta}(u)\frac{df_{\theta,u}}{dh_{\theta,u}}(a). \tag{75}$$

Taking logarithms and integrating with respect to $Q_0(du)f_\theta(da \mid u)$ proves (72). □

The next proposition identifies the conditional Metropolis kernel hidden inside the round-trip construction. This identification is standard auxiliary-variable/Metropolis-within-Gibbs reasoning rather than a new general MCMC principle; related auxiliary constructions are discussed by Titsias & Papaspiliopoulos (2018). For the density ratio below, assume additionally that $f_\theta(\cdot \mid u)$ and $h_\theta(\cdot \mid u)$ admit densities with respect to m for ρ_θ-almost every u; zero-density cases use the usual Metropolis conventions.

Proposition 2 (Independence Metropolis–Hastings at fixed intermediate configuration)**.** *For ρ_θ-almost every bridge u, the round-trip record update is an independence Metropolis–Hastings update with target $h_\theta(\cdot \mid u)$ and proposal $f_\theta(\cdot \mid u)$. Its acceptance probability is*

$$\alpha_u(a, a') = 1 \wedge \frac{h_\theta(a' \mid u) f_\theta(a \mid u)}{h_\theta(a \mid u) f_\theta(a' \mid u)}. \tag{76}$$

Proof. At a fixed bridge, Eqs. 61 and 70 give

$$\pi(\mathsf{E}_\theta(u, a)) r_\theta(b(u, a) \mid \mathsf{E}_\theta(u, a)) = \rho_\theta(u) h_\theta(a \mid u) \tag{77}$$

for almost every (u, a) in density notation. The same identity holds for the proposal record a'. Substituting both identities into the direct round-trip ratio (10) cancels the common $\rho_\theta(u)$ factor and yields (76), which is exactly the standard independence-MH ratio. □

Let $R_{\theta,u}$ denote this record-space MH kernel:

$$\begin{aligned} R_{\theta,u}(a, da') = {} & f_\theta(da' \mid u) \alpha_u(a, a') \\ & + \left[1 - \int f_\theta(dc \mid u) \alpha_u(a, c)\right] \delta_a(da'). \end{aligned} \tag{78}$$

It is reversible and invariant with respect to $h_\theta(\cdot \mid u)$. This conditional invariance does not imply irreducibility: missing conditional proposal support can leave the chain reducible even though the Metropolis identity is valid on the communicating class.

B.4 Conditional mean acceptance

At stationarity and at fixed u, the old and proposed records have joint law $h_\theta(da \mid u) f_\theta(da' \mid u)$. Multiplying this density by the Metropolis acceptance function and integrating gives

$$\bar{\alpha}(u) = \iint \min\{h_\theta(a \mid u) f_\theta(a' \mid u), f_\theta(a \mid u) h_\theta(a' \mid u)\} \, m(da) m(da') \tag{79}$$

$$= 1 - d_{\mathrm{TV}}(h_{\theta,u} \otimes f_{\theta,u}, f_{\theta,u} \otimes h_{\theta,u}). \tag{80}$$

The second equality follows from $\int \min(p, q) = 1 - d_{\mathrm{TV}}(p, q)$ for probability densities. Marginalization and the triangle inequality give

$$\delta_\theta(u) \le 1 - \bar{\alpha}(u) \le 2\delta_\theta(u). \tag{81}$$

This identity concerns all accepted proposals, rather than any particular pair of endpoint regions.

B.5 Ideal conditional matching and the data-augmentation limit

The condition relevant to the round-trip inference kernel is

$$f_\theta(\cdot \mid u) = h_\theta(\cdot \mid u) \quad \text{for } \rho_\theta\text{-almost every } u. \tag{82}$$

This is weaker than requiring the full path KL D_θ to vanish, because q_0 cancels from the round-trip acceptance ratio.

Define the target-induced joint law of endpoint and bridge by

$$J_\theta(dx, du) := (\mathsf{E}_\theta, U)_\# P_\theta. \tag{83}$$

Its marginals are π and ρ_θ. Let

$$\kappa_\theta(du \mid x) := J_\theta(du \mid x), \qquad \nu_\theta(dx \mid u) := J_\theta(dx \mid u) \tag{84}$$

be regular conditional laws. Since $X = \mathsf{E}_\theta(U, A)$ under the pullback target law,

$$\nu_\theta(\cdot \mid u) = (\mathsf{E}_\theta(u, \cdot))_\# h_\theta(\cdot \mid u). \tag{85}$$

Theorem 4 (Ideal round trip equals a DA chain)**.** *Under* (82)*, every conditional record proposal is accepted almost surely and the configuration-space kernel is*

$$K_\star(x, dy) = \int \kappa_\theta(du \mid x)\nu_\theta(dy \mid u). \tag{86}$$

Thus, at stationarity, one ideal round trip is the two-step conditional update

$$X_t \longrightarrow U_t \sim J_\theta(du \mid X_t) \longrightarrow X_{t+1} \sim J_\theta(dx \mid U_t), \tag{87}$$

which is the classical two-component DA chain (Tanner & Wong, 1987).

Proof. Under (82), the ratio in (76) equals one wherever the target conditional density is positive, hence $R_{\theta,u}(a, da') = h_\theta(da' \mid u)$, independently of the current record a. Starting from $X = x$ under the target, the reverse draw and pullback produce (U, A) according to the conditional law of (U, A) under P_θ given $X = x$, whose U-marginal is $\kappa_\theta(du \mid x)$. The refreshed record $A' \sim h_\theta(\cdot \mid u)$ maps through E_θ to $Y \sim \nu_\theta(\cdot \mid u)$ by (85). Integrating over u gives (86). □

Define centered Hilbert spaces $L^2_0(\pi)$ and $L^2_0(\rho_\theta)$ and the conditional-expectation operator

$$(C_\theta g)(u) := \mathbb{E}[g(X) \mid U = u]. \tag{88}$$

Its adjoint is $(C^*_\theta \varphi)(x) = \mathbb{E}[\varphi(U) \mid X = x]$. The DA Markov operator is

$$K_\star = C^*_\theta C_\theta. \tag{89}$$

Define the maximal correlation under J_θ by

$$\gamma_{\mathrm{mc}} := \sup_{\substack{g \in L^2_0(\pi),\ \|g\|_2 = 1 \\ \varphi \in L^2_0(\rho_\theta),\ \|\varphi\|_2 = 1}} \mathbb{E}[g(X)\varphi(U)]. \tag{90}$$

The classical DA operator identity gives

$$\|C_\theta\| = \gamma_{\mathrm{mc}}, \qquad \|K_\star\|_{L^2_0(\pi)} = \gamma^2_{\mathrm{mc}}. \tag{91}$$

Since $C^*_\theta C_\theta$ is positive and self-adjoint, its Poincaré (right) spectral gap is therefore

$$\boxed{\mathrm{Gap}(K_\star) = 1 - \gamma^2_{\mathrm{mc}}.} \tag{92}$$

This is the standard DA convergence characterization of Liu et al. (1994); here it applies because Theorem 4 identifies the ideal NHMC round trip with a DA kernel.

B.6 Inter-region transitions in the ideal limit

For the partition $\{A_k\}$, define the target conditional region probability at a fixed bridge,

$$m_k(u) := \nu_\theta(A_k \mid u) = \mathbb{P}(X \in A_k \mid U = u), \tag{93}$$

and the matrix

$$\Omega_{k\ell} := \int \rho_\theta(du) m_k(u) m_\ell(u). \tag{94}$$

Since $w_k = \int \rho_\theta(du) m_k(u)$, $\Omega_{k\ell}$ is the joint probability of regions A_k and A_ℓ for two conditionally independent endpoints drawn from the same intermediate configuration.

Proposition 3 (Ideal inter-region flow)**.** *At stationarity, the ideal kernel* (86) *has inter-region flow*

$$F^\star_{k\ell} := \int_{A_k} \pi(dx) K_\star(x, A_\ell) = \Omega_{k\ell}. \tag{95}$$

Hence $F^\star_{k\ell} = F^\star_{\ell k}$ and $\sum_\ell F^\star_{k\ell} = w_k$.

Proof. Under the stationary DA construction, conditional on $U = u$, the two successive endpoints are conditionally independent and both distributed as $\nu_\theta(\cdot \mid u)$. Therefore

$$F^\star_{k\ell} = \mathbb{P}(X_t \in A_k, X_{t+1} \in A_\ell) \tag{96}$$

$$= \int \rho_\theta(du) m_k(u) m_\ell(u) = \Omega_{k\ell}. \tag{97}$$

Symmetry is immediate, and summing over ℓ uses $\sum_\ell m_\ell(u) = 1$. □

For two regions A and A^c, write $w = \pi(A)$ and $m(u) = \mathbb{P}(X \in A \mid U = u)$. Then

$$\boxed{F^\star_{A,A^c} = w(1-w) - \mathrm{Var}_{\rho_\theta}[m(U)].} \tag{98}$$

Indeed, $\mathbb{E}_{\rho_\theta}[m] = w$ and $\mathbb{E}[m(1-m)] = w - \mathbb{E}[m^2]$. The same conditional-independence argument gives

$$\mathrm{Cov}(\mathbf{1}_A(X_t), \mathbf{1}_A(X_{t+1})) = \mathrm{Var}_{\rho_\theta}[m(U)], \tag{99}$$

which is the region-indicator specialization of the classical DA covariance identity (Liu et al., 1994).

Equation (98) shows how dependence on the intermediate configuration limits transitions between the two regions. If the region label is measurable from the bridge, then $m(U) \in \{0, 1\}$ almost surely and $F^\star_{A,A^c} = 0$: perfect conditional matching and unit acceptance can still coexist with complete region trapping. At the opposite extreme, if the region label is independent of U, then $m(U) = w$ almost surely and $F^\star_{A,A^c} = w(1-w)$; one ideal round trip refreshes the region label to its equilibrium distribution.

For $w_k > 0$ one may define the stationary one-step conditional matrix $T^\star_{k\ell} := \Omega_{k\ell}/w_k$. In general the region-label process $M(X_t)$ need not itself be Markov, so powers of $T^\star$ need not reproduce multi-step region probabilities without an additional lumpability condition. The full configuration-space gap statement above therefore follows from the operator factorization (89), not from eigenvalues of this coarse one-step matrix.

B.7 Finite conditional mismatch and accepted inter-region flow

We now allow the conditional forward path law to differ from the target-induced conditional law. The proof uses a common subdensity and total-variation distance, followed by Pinsker's inequality to connect the conditional mismatch to the recorded path KL. The resulting quantity is the accepted stationary joint transition probability between specified regions, which we compare with the value obtained when the conditional path distributions match exactly. Assume in this subsection that both conditional laws have densities with respect to the common record reference measure m, and define

$$\delta_\theta(u) := d_{\mathrm{TV}}(f_\theta(\cdot \mid u), h_\theta(\cdot \mid u)) = \frac{1}{2}\int |f_\theta(a \mid u) - h_\theta(a \mid u)|\, m(da). \tag{100}$$

Let

$$C_k(u) := \{a : \mathsf{E}_\theta(u, a) \in A_k\}, \tag{101}$$

so that $h_\theta(C_k(u) \mid u) = m_k(u)$.

The ideal flow $\Omega_{k\ell}$ uses h_θ for both records. By contrast, before the Metropolis test the learned proposal has joint flow

$$G_{k\ell} = \mathbb{E}_{\rho_\theta}[m_k(U) n_\ell(U)], \qquad n_\ell(u) = f_\theta(C_\ell(u) \mid u). \tag{102}$$

In general $G_{k\ell} \neq \Omega_{k\ell}$. The accepted flow satisfies $F^{\mathrm{RT}}_{k\ell} \leq G_{k\ell}$.

Theorem 5 (Exact stationary round-trip flow)**.** *For $k \neq \ell$, the stationary round-trip flow is*

$$\begin{aligned} F^{\mathrm{RT}}_{k\ell} &:= \int_{A_k} \pi(dx) K_{\mathrm{RT}}(x, A_\ell) \\ &= \int \rho_\theta(du) \int_{C_k(u)} m(da) \int_{C_\ell(u)} m(da') \\ &\qquad \times \min\{h_\theta(a \mid u) f_\theta(a' \mid u), f_\theta(a \mid u) h_\theta(a' \mid u)\}\,. \end{aligned} \tag{103}$$

Proof. At stationarity, the reverse draw and pullback give $(U, A) \sim \rho_\theta(du) h_\theta(da \mid u)$. Conditional on (u, a), the new record a' is drawn from $f_\theta(\cdot \mid u)$ and accepted according to (76). Hence the accepted pair density is

$$\begin{aligned} &h_\theta(a \mid u) f_\theta(a' \mid u) \alpha_u(a, a') \\ &\qquad = \min\{h_\theta(a \mid u) f_\theta(a' \mid u), f_\theta(a \mid u) h_\theta(a' \mid u)\}\,. \end{aligned} \tag{104}$$

For $k \neq \ell$ a rejection cannot contribute to a cross-region transition. Restricting $a \in C_k(u)$ and $a' \in C_\ell(u)$ and integrating gives (103). □

Theorem 6 (Two-sided comparison of accepted and ideal flow)**.** *For every $k \neq \ell$,*

$$F_{k\ell}^{\mathrm{RT}} \geq \int \rho_\theta(du)[m_k(u)m_\ell(u) - \delta_\theta(u)]_+ \tag{105}$$

$$\geq \Omega_{k\ell} - \mathbb{E}_{\rho_\theta}[\delta_\theta(U)]. \tag{106}$$

In addition,

$$|F_{k\ell}^{\mathrm{RT}} - \Omega_{k\ell}| \leq \mathbb{E}_{\rho_\theta}[\delta_\theta(U)]. \tag{107}$$

Proof. Fix u and define the common subdensity

$$s_u(a) := \min\{f_\theta(a \mid u), h_\theta(a \mid u)\}. \tag{108}$$

Both products inside the minimum in (103) are at least $s_u(a)s_u(a')$. Define

$$\ell_j(u) := \int_{C_j(u)} s_u(a)m(da), \qquad e_j(u) := m_j(u) - \ell_j(u) \geq 0. \tag{109}$$

Because $C_k(u)$ and $C_\ell(u)$ are disjoint,

$$e_k(u) + e_\ell(u) \leq \int (h_\theta - s_u)\, dm = \delta_\theta(u). \tag{110}$$

The fixed-u cross-flow is therefore at least

$$\ell_k\ell_\ell = (m_k - e_k)(m_\ell - e_\ell) \tag{111}$$

$$\geq m_k m_\ell - (e_k + e_\ell) \tag{112}$$

$$\geq m_k m_\ell - \delta_\theta(u). \tag{113}$$

It is nonnegative, which gives (105); since $[z]_+ \geq z$, integration gives (106). For the upper bound, write $\phi_{k\ell}(u)$ for the fixed-u integral in (103). The event form of total variation gives $f_\theta(C_\ell(u) \mid u) \leq m_\ell(u) + \delta_\theta(u)$, hence

$$\phi_{k\ell}(u) \leq m_k(u) f_\theta(C_\ell(u) \mid u) \tag{114}$$

$$\leq m_k(u)m_\ell(u) + m_k(u)\delta_\theta(u) \tag{115}$$

$$\leq m_k(u)m_\ell(u) + \delta_\theta(u). \tag{116}$$

Combining both conditional inequalities and integrating with respect to ρ_θ proves (107). □

The previous result uses a ρ_θ-average, whereas the training KL uses q_0-averages. The following lemma connects them.

Lemma 3 (Path KL controls the conditional TV discrepancy)**.** *With $D_{0,\theta}$ and $D_{1,\theta}$ from Lemma 2,*

$$\mathbb{E}_{\rho_\theta}[\delta_\theta(U)] \leq \sqrt{\frac{D_{1,\theta}}{2}} + \sqrt{\frac{D_{0,\theta}}{2}} \leq \sqrt{D_\theta}. \tag{117}$$

Proof. Conditional Pinsker and Jensen's inequality give

$$\mathbb{E}_{Q_0}[\delta_\theta(U)] \leq \sqrt{\frac{1}{2}\mathbb{E}_{Q_0} D_{\mathrm{KL}}(f_\theta(\cdot \mid U) \| h_\theta(\cdot \mid U))} = \sqrt{\frac{D_{1,\theta}}{2}}. \tag{118}$$

Since $0 \leq \delta_\theta \leq 1$,

$$|\mathbb{E}_{\rho_\theta}\delta_\theta - \mathbb{E}_{Q_0}\delta_\theta| \leq d_{\mathrm{TV}}(\rho_\theta, Q_0) \leq \sqrt{\frac{D_{0,\theta}}{2}}, \tag{119}$$

where the final inequality is Pinsker. Combining the bounds gives the first inequality in (117). The second follows from $(\sqrt{D_{0,\theta}} + \sqrt{D_{1,\theta}})/\sqrt{2} \leq \sqrt{D_{0,\theta} + D_{1,\theta}}$ and Lemma 2. □

Corollary 1 (Finite path-KL inter-region flow)**.** *For every $k \neq \ell$,*

$$\boxed{|F_{k\ell}^{\mathrm{RT}} - \Omega_{k\ell}| \leq \sqrt{D_\theta}.} \tag{120}$$

For two regions A and A^c,

$$\boxed{F_{A,A^c}^{\mathrm{RT}} \geq w(1-w) - \mathrm{Var}_{\rho_\theta}[m(U)] - \sqrt{D_\theta}.} \tag{121}$$

Thus

$$\mathrm{Var}_{\rho_\theta}[m(U)] + \sqrt{D_\theta} < w(1-w) \tag{122}$$

is a sufficient condition for strictly positive accepted cross-region flow.

Proof. Combine Theorem 6 with Lemma 3; the two-region identity follows from (98). □

The first bound in (106) is the direct round-trip quantity. The $\sqrt{D_\theta}$ corollary is deliberately coarser: the full path KL contains both bridge-marginal mismatch $D_{0,\theta}$ and conditional mismatch $D_{1,\theta}$, whereas q_0 cancels from the fixed-bridge Metropolis ratio. Thus D_θ is a trainable sufficient proxy, not in general a tight predictor of round-trip mixing.

This is stronger than detailed balance alone. Detailed balance enforces $F^{\mathrm{RT}}_{k\ell} = F^{\mathrm{RT}}_{\ell k}$, but both quantities may vanish. Corollary 1 gives an explicit sufficient condition under which the conditional overlap exceeds the path-law discrepancy and therefore guarantees nonzero accepted probability flow between the two regions.

B.8 A stronger Poincaré-gap bound under uniform conditional domination

The finite-KL result above is an average, region-specific bound. A full configuration-space spectral comparison requires a stronger uniform condition.

Uniform conditional domination. Assume there exists $c_{\mathrm{dom}} \in (0, 1]$ such that, for ρ_θ-almost every u,

$$f_\theta(a \mid u) \geq c_{\mathrm{dom}} h_\theta(a \mid u) \quad \text{for } m\text{-almost every } a. \tag{123}$$

This is the usual sufficient condition for uniform ergodicity of an independence sampler, imposed here conditionally at each bridge (Mengersen & Tweedie, 1996). The comparison below is a conditional-minorization specialization of the broader hybrid-Gibbs/MwG comparison theme (Qin et al., 2025; Ascolani et al., 2026); it is retained as a sufficient condition, not as the main novelty claim.

Lemma 4 (Conditional minorization)**.** *Under* (123)*,*

$$R_{\theta,u}(a, da') \geq c_{\mathrm{dom}} h_\theta(da' \mid u) \tag{124}$$

for ρ_θ-almost every u and $h_\theta(\cdot \mid u)$-almost every current record a.

Proof. Write $w_u(a) = h_\theta(a \mid u)/f_\theta(a \mid u)$ on the support of the target conditional. The domination assumption gives $w_u(a) \leq 1/c_{\mathrm{dom}}$. The accepted proposal density equals

$$f_\theta(a' \mid u) \min\left\{1, \frac{w_u(a')}{w_u(a)}\right\}. \tag{125}$$

If $w_u(a') \leq w_u(a)$, this is $h_\theta(a' \mid u)/w_u(a) \geq c_{\mathrm{dom}} h_\theta(a' \mid u)$. If $w_u(a') > w_u(a)$, it is $f_\theta(a' \mid u) \geq c_{\mathrm{dom}} h_\theta(a' \mid u)$. Thus the accepted part alone dominates $c_{\mathrm{dom}} h_\theta(da' \mid u)$; the rejection atom is nonnegative. □

Let $\eta_\theta(du, da \mid x)$ be the target conditional law of (U, A) given $X = x$ under P_θ. The actual and ideal kernels can be written as

$$K_{\mathrm{RT}}(x, G) = \int \eta_\theta(du, da \mid x) R_{\theta,u}\big(a, \mathsf{E}_\theta(u, \cdot)^{-1}(G)\big), \tag{126}$$

$$K_\star(x, G) = \int \eta_\theta(du, da \mid x) h_\theta\big(\mathsf{E}_\theta(u, \cdot)^{-1}(G) \mid u\big). \tag{127}$$

Lemma 4 therefore gives the measure domination

$$K_{\mathrm{RT}}(x, \cdot) \geq c_{\mathrm{dom}} K_\star(x, \cdot) \quad \text{for } \pi\text{-almost every } x. \tag{128}$$

For a π-reversible kernel K, define its Dirichlet form

$$\mathcal{E}_K(g, g) := \frac{1}{2} \int \pi(dx) K(x, dy)[g(y) - g(x)]^2 \tag{129}$$

and Poincaré (right) spectral gap

$$\mathrm{Gap}(K) := \inf_{\mathrm{Var}_\pi(g) > 0} \frac{\mathcal{E}_K(g, g)}{\mathrm{Var}_\pi(g)}. \tag{130}$$

Proposition 4 (Round-trip Poincaré-gap lower bound)**.** *Under* (123)*,*

$$\boxed{\mathrm{Gap}(K_{\mathrm{RT}}) \geq c_{\mathrm{dom}}\,\mathrm{Gap}(K_\star) = c_{\mathrm{dom}}(1-\gamma_{\mathrm{mc}}^2).} \tag{131}$$

Moreover, for $k \neq \ell$*,*

$$F_{k\ell}^{\mathrm{RT}} \geq c_{\mathrm{dom}}\Omega_{k\ell}. \tag{132}$$

Proof. Theorem 2 makes K_{RT} π-reversible. Since the Dirichlet-form integrand is nonnegative, (128) implies

$$\mathcal{E}_{K_{\mathrm{RT}}}(g,g) \geq c_{\mathrm{dom}}\mathcal{E}_{K_\star}(g,g) \tag{133}$$

for every square-integrable g. Dividing by $\mathrm{Var}_\pi(g)$ and taking the infimum gives $\mathrm{Gap}(K_{\mathrm{RT}}) \geq c_{\mathrm{dom}}\,\mathrm{Gap}(K_\star)$. Equation (92) gives the final equality. Integrating (128) over $x \in A_k$ and evaluating the measure on A_ℓ gives (132) using Proposition 3. □

The domination condition is deliberately stronger than a small average path KL: a KL divergence can be small while $f_\theta(a \mid u)/h_\theta(a \mid u)$ becomes arbitrarily small on a thin set. Therefore Corollary 1 and Proposition 4 are complementary. The former is an average, region specific result for differing conditional path laws; the latter is a stronger sufficient condition for a full configuration-space Poincaré gap.

B.9 The first step of sector relaxation

Let A_+, A_- partition configuration space and let $w_+ = \pi(A_+) \in (0,1)$, $w_- = 1 - w_+$. For any π-stationary kernel K, define $F = \int_{A_+} \pi(dx)K(x, A_-)$. Stationarity gives

$$w_+ = \int_{A_+} \pi(dx)K(x,A_+) + \int_{A_-} \pi(dx)K(x,A_+) = w_+ - F + F_{-+}, \tag{134}$$

so $F_{-+} = F$. For $\mu_p = p\pi(\cdot \mid A_+) + (1-p)\pi(\cdot \mid A_-)$, direct integration yields

$$p' = (\mu_p K)(A_+) = \frac{p}{w_+}(w_+ - F) + \frac{1-p}{w_-}F, \tag{135}$$

$$p' - w_+ = (p - w_+)\left(1 - \frac{F}{w_+ w_-}\right). \tag{136}$$

This proves (16). The factor can be negative; no general monotonicity claim is made. Subsequent distributions need not retain target conditionals within sectors, so this formula cannot generally be iterated. At any time, $|\mu_j(A_+) - w_+| \leq d_{\mathrm{TV}}(\mu_j, \pi)$: sector imbalance is a lower bound on full-distribution error, not a certificate of convergence when it vanishes.

B.10 Scope of the transition bounds

The results above imply the following hierarchy:

$$D_\theta \text{ small} \Longrightarrow \text{aggregate raw endpoint region masses are close,} \tag{137}$$

$$f_\theta(\cdot \mid u) \approx h_\theta(\cdot \mid u) \Longrightarrow \text{small conditional Metropolis mismatch,} \tag{138}$$

$$\Omega_{k\ell} \text{ large and mismatch small} \Longrightarrow \text{non-negligible accepted inter-region flow,} \tag{139}$$

$$f_\theta \geq c_{\mathrm{dom}}h_\theta,\ \gamma_{\mathrm{mc}} < 1 \Longrightarrow \mathrm{Gap}(K_{\mathrm{RT}}) \geq c_{\mathrm{dom}}(1-\gamma_{\mathrm{mc}}^2) > 0. \tag{140}$$

The bounds also show that the following implications require additional assumptions:

$$D_\theta \text{ small} \not\Rightarrow \text{fast mixing}, \qquad \alpha_{\mathrm{RT}} \text{ high} \not\Rightarrow \text{inter-region mixing.} \tag{141}$$

For example, in an ideal DA construction for which the intermediate configuration determines the region label exactly, $f_\theta = h_\theta$ and every proposal is accepted, but (98) gives zero cross-region flow.

The regions A_k may in particular be topological sectors. The preceding results then give sufficient conditions for inter-sector transport at a fixed lattice target. Extending the statement uniformly toward

the continuum would require additional control of the corresponding family of kernels. For a family indexed by lattice spacing a, a uniform conclusion from Proposition 4 would require, for example,

$$\inf_{a \le a_0} c_{\mathrm{dom}}(a) > 0, \qquad \sup_{a \le a_0} \gamma_{\mathrm{mc}}(a) < 1. \tag{142}$$

These conditions are not implied by the average-work objective alone. The result therefore characterizes finite-volume inter-region transport, while its behavior in a thermodynamic or continuum limit must be established separately.

C Numerical Setup

C.1 Notation and reported statistical quantities

Table 2: Notation used in the construction and experiments. Symbols with a system-specific role are defined locally.

Symbol	Meaning
$\gamma = e^{-E}$, Z, $\pi = \gamma/Z$	Target density, normalizer, normalized target
q_0, Q_θ, P_θ	Reference configuration law, forward and reverse path laws
t, T, n_{LF}	Path stage, number of momentum draws, substeps per stage
j	Saved Markov-transition index or lag
N; L, V; N_c	Evaluation paths; lattice side and volume; color rank
p_t, $\bar{p}_{t+1}$	Drawn input momentum and deterministic output momentum
W_θ, $\widetilde{w} = e^{-W_\theta}$, $\bar{w}$	Work, unnormalized and normalized weights
u, a, x, b; Ψ_θ	Bridge, forward record, endpoint, reverse record; path bijection
ρ_θ, h_θ, f_θ	Bridge marginal, target-induced and proposed conditional record laws
$\Omega_{k\ell}$, $G_{k\ell}$, $F^{\mathrm{RT}}_{k\ell}$	Ideal, proposed, and accepted stationary joint flows
δ_θ, D_θ	Conditional total variation and forward path KL
λ, λ_{var}	Scalar-field coupling and work-variance coefficient
γ_{mc}, c_{dom}	Maximal correlation and conditional-domination constant

The bare work W_θ is distinct from a size-indexed Wilson loop $W_{r\times s}[U]$. Topological charge $Q[U]$ is a configuration observable, whereas Q_θ is a path probability law. We reserve d_{TV} for half the L_1 distance between probability measures.

C.2 Target Definitions

All numerical studies use unnormalized target densities $\gamma(x) = \exp[-E(x)]$. For the double-well family, the target is

$$\log \gamma_d(x) = \sum_{j=0}^{d/2-1} \left[-x_{2j}^4 + 6x_{2j}^2 + 0.5x_{2j} - \frac{1}{2}x_{2j+1}^2 \right]. \tag{143}$$

The Gaussian coordinate in each pair factorizes, while the even coordinate is a tilted double well. The base law is $q_0 = \mathcal{N}(0, I)$ for the neural NHMC path runs. Analytic normalizers and exact target samples are obtained from the one-dimensional factorization; sign modes are defined by the signs of the double-well coordinates.

For the full-depth LJ13 benchmark, configurations contain 13 particles in three dimensions with their center of mass removed. At temperature one, the target energy is

$$E_{\mathrm{LJ13}}(x) = \sum_{i\neq j} \left(r_{ij}^{-12} - 2r_{ij}^{-6} \right) + \frac{1}{2}\sum_i \|x_i - \bar{x}\|^2. \tag{144}$$

The ordered-pair convention in the first term matches the standard energy-factor-one benchmark. Endpoint energies use this target exactly. The proposal dynamics floor r_{ij}^2 at 0.49 only inside the force field to avoid numerical overflow for near-collisions; this reversible proposal field does not change the endpoint target density or the recorded-work correction.

For the two-dimensional scalar ϕ^4 benchmark we use the finite-volume lattice action convention

$$S[\phi] = \sum_x \left[-2\kappa \sum_{\mu=1}^{2} \phi_x \phi_{x+\hat{\mu}} + (1-2\lambda)\phi_x^2 + \lambda\phi_x^4 \right], \tag{145}$$

with periodic boundary conditions, $\lambda = 0.022$, and the κ values listed in the main and appendix tables. The magnetization is $M = V^{-1}\sum_x \phi_x$ and the ordering observable shown is $|M|$. We use

$$\chi = V\left(\langle M^2 \rangle - \langle |M| \rangle^2 \right), \qquad U_B = 1 - \frac{\langle M^4 \rangle}{3\langle M^2 \rangle^2}. \tag{146}$$

The same definitions are applied to the raw proposal, weighted NHMC, and HMC reference samples.

For compact $U(1)$ we parameterize links by angles $U_{x,\mu} = \exp(i\theta_{x,\mu})$ on a periodic two-dimensional lattice. The paper-level Wilson convention is

$$S_W[U] = \beta \sum_p (1 - \cos\theta_p). \tag{147}$$

For $\mathrm{SU}(N_c)$ the same convention is $S_W = \beta \sum_p [1 - N_c^{-1} \operatorname{ReTr} U_p]$. Some calculations use the reduced action $\widetilde{S}_W = -(\beta/N_c) \sum_p \operatorname{ReTr} U_p$, which differs only by the known additive constant βN_p. The two conventions define the same normalized target and fixed-β observables, but their partition functions and absolute free energies differ by this known constant. Absolute free energies use the stated convention, including this constant whenever the two conventions are compared. Plaquette observables use $\cos\theta_p$. The geometric topological charge is computed from wrapped plaquette angles,

$$Q[U] = \frac{1}{2\pi} \sum_p \mathrm{Arg}\{\exp(i\theta_p)\}, \tag{148}$$

and rounded sectors are used only for finite-chain trajectory summaries. This wrapped-plaquette convention follows the geometric lattice topological-charge construction (Lüscher, 1982). Finite-volume plaquette and square Wilson-loop references are evaluated from the character expansion $Z_V \propto \sum_{n\in\mathbb{Z}} I_n(\beta)^V$, with an area-$A$ loop replacing I_n^V by $I_n^{V-A} I_{n+1}^A$ in the numerator. The finite-volume character-expansion framework follows the standard two-dimensional lattice Yang–Mills construction (Aroca & Kubyshin, 2000).

For two-dimensional $\mathrm{SU}(2)$ the links are unit quaternions. The integration reference measure is product Haar measure, and the earlier pilot uses the normalized Haar source distribution (density one with respect to that reference measure). The action is $S[U] = \beta V\big(1 - P[U]\big)$, where P is the mean plaquette. The single-plaquette reference is $I_2(\beta)/I_1(\beta)$. The proposal uses a learned-force leapfrog kernel with five leapfrog steps per stage. Independent path-SNIS starts from Haar; the round-trip chains reported in Table 26 are initialized by a short physics HMC run from Haar, unlike the $U(1)$ pilot, which used a single forward NHMC path and no HMC warm start. The Haar volume rows in Appendix F.8 keep the same action and kernel, but initialize every NHMC estimator from Haar plus one forward path. Two-dimensional $\mathrm{SU}(3)$ uses the same $S = \beta V(1-P)$ with $P = \operatorname{Re}\operatorname{Tr} U_p/3$ and the same normalized Haar source convention. The plaquette reference at $\beta = 5$ is the single-plaquette Weyl integral 0.35395. The force and momentum-mean networks are the gauge-equivariant CNNs described in that appendix; HMC in those tables uses $n_{\mathrm{lf}} = 10$ and $\varepsilon = 0.08$.

The systematic two-dimensional gauge production study in Appendix F.5 uses the same Wilson targets but a different source distribution from the earlier Haar-reference experiments. For $U(1)$, the normalized source is a punctured plaquette prior over $V - 1$ independent angles, with the last plaquette restored by periodic closure in the target action. For $\mathrm{SU}(2)$ and $\mathrm{SU}(3)$, the source uses a finite-volume tree/plaquette representation with both torus holonomies retained. These structured priors are computational devices; all reported weights use the full finite-volume Wilson target. The frozen production paths use 64 fresh-momentum stages and two leapfrog substeps per stage. The learned components are separate forward/reverse momentum means, learned widths, gauge-equivariant residual kicks, and a learned positive drift scale. In production this drift rate depends only on stage time, β, and fixed lattice metadata, not on the current gauge field or momentum; hence it does not introduce an unrecorded state-dependent Jacobian. The $U(1)$ production path also records a post-transport global flux move and its reverse probability.

The four-dimensional $\mathrm{SU}(3)$ extension in Appendix H.3 uses Wilson and tree-level Symanzik pure-gauge actions. Its present results are small-volume structural and overlap tests, not part of the systematic two-dimensional production count.

Autocorrelation and scaling convention. For a normalized autocorrelation $\rho_O(j)$ at saved-transition lag j,

$$\tau_{\mathrm{int}}(O) = \frac{1}{2} + \sum_{j=1}^{j_{\max}} \rho_O(j). \tag{149}$$

Table 3: Settings for the revised main DW rows and their earlier-MLP control. T counts momentum-refresh stages and n_{LF} the leapfrog substeps in each stage. All models use width 128, batch size 4096, and a learning-rate schedule from 2×10^{-3} to 10^{-5}. The displayed error bars concern three evaluation seeds of one trained model, not three training seeds.

Model	T	n_{LF}	Updates	Parameters	Continuous configuration
DW4 earlier MLP	16	4	800	78,882	Fixed target force
DW4 updated block	16	4	800	282,650	Annealing floor 0.05
DW8 updated block	20	5	2,500	285,870	Analytic score input

For the topology chains this is estimated on one continuous chain with the Wolff/Γ-method automatic window at $S = 2$. Independent chains are not concatenated for the primary autocorrelation estimate; a constant Q^2 trace is reported as frozen rather than assigned a finite autocorrelation time. In two-dimensional compact $U(1)$, $\beta \propto a^{-2}$, so fixed physical area L^2a^2 corresponds to fixed L^2/β. The scaling fit therefore uses a separate line with $L^2/\beta \simeq 80$ and $\log \tau = \log A + b \log \beta$, weighted by $1/\sigma^2_{\log \tau}$ with $\sigma_{\log \tau} = \delta\tau/\tau$.

C.3 Training and deployment settings

Each evaluated model is frozen. The new DW4 and DW8 models use separate forward and reverse conditional momentum laws around a reversible kick–drift composition. The mean-work objective is supplemented by $\lambda_{\mathrm{var}} \operatorname{Var}(W)$, with $\lambda_{\mathrm{var}} = 0.05$. This is a path-KL objective with variance regularization, not a data-likelihood loss or an endpoint-only KL. The block architecture couples the repeated coordinate pairs through shared features; "force-conditioned" denotes analytic score information in the continuous networks.

The stored DW models were trained with separate, state-independent Bernoulli stage logits initialized at zero, corresponding to probability $1/2$. Those logits did not share parameters with the continuous networks. The historical training differentiated the loss at fixed sampled labels, without the complete likelihood-ratio term for the label sampling law. It therefore does not provide an unbiased joint gradient for all continuous and discrete parameters. For these state-independent separate logits, continuous parameters do not determine the label probabilities; the missing label-score term concerns the logit directions, not the continuous-parameter derivative at the current label schedule.

The revised main evaluations retain the trained continuous parameters and replace the historical logits by the prescribed deployment probability $p_{\mathrm{flip}} = 0.05$. They are not retrainings with that probability fixed from initialization. Appendix D.2 measures this intervention and reports the historical schedules separately. All actual forward and reverse probabilities still enter the frozen path weight.

The main ϕ^4 mechanism experiment instead has fixed Gaussian momentum laws throughout and a prescribed sign-flip probability. It has no neural checkpoint and makes no claim about a neural training gain. The earlier learned κ-conditional scalar scan and its historical discrete-move training are reported separately. For any model with state-dependent label probabilities, a label-score contribution may also involve continuous parameters through the preceding state; the DW argument above cannot be applied without checking that dependence.

A corrected local reverse helper uses the same reverse-stage probability index that is recorded by the forward step. The archived DW chain helper already uses that index. These code changes are supplied separately and are not described as having produced the historical evaluations. In the constant-probability intervention the stage-index difference cancels.

The earlier protocols are retained in Table 4 so that the archived figures and chain comparisons remain distinguishable from the revised independent-path table.

C.4 Statistical Quantities

For the revised independent-path tables, define

$$\widehat{p}_k^{\mathrm{SNIS}} = \sum_i \bar{w}_i \mathbf{1}\{x_T^{(i)} \in A_k\}, \qquad L_1^{\mathrm{SNIS}} = \sum_k |\widehat{p}_k^{\mathrm{SNIS}} - w_k|. \tag{150}$$

Table 4: Historical training and evaluation settings. The earlier DW4/DW8 rows and learned scalar scan are not the updated models in Table 1 or the fixed-Gaussian round-trip mechanism experiment.

Result	Target and path	Network	Optimization	Evaluation
DW4 earlier	$d = 4$; 16×4 LF	MLP, width 128	batch 4096; 800 epochs; $2{\times}10^{-3} \to 10^{-5}$	200000 paths; path-SNIS/path-IMH
DW4 round trip	$d = 4$; 32×5 LF	MLP, width 128	batch 4096; 1200 epochs; $2{\times}10^{-3} \to 10^{-5}$	three checkpoints; matched-path comparison
DW8 earlier	$d = 8$; 20×5 LF	MLP, width 128	batch 4096; 1000 epochs; $2{\times}10^{-3} \to 10^{-5}$	120000 paths; path-SNIS/path-IMH and matched-budget round trip
ϕ^4	8×8; 16×5 LF	CNN, channels 32/32	batch 64; 20000 epochs; $10^{-3} \to 10^{-5}$	102400 paths/κ; 5000-transition HMC/IMH; 10000-transition round trip
LJ13 full depth	13 particles; 56×10 LF	force-equivariant, width 128	batch 320; 2600 epochs; $5{\times}10^{-4} \to 10^{-5}$	50000 paths; raw/path-SNIS W_2
LJ13 ablation	13 particles; 56×10 LF	same architecture	batch 320; 1300 epochs; $5{\times}10^{-4} \to 10^{-5}$	one seed; 100000 paths/setting
2D gauge production	$U(1)$, $\mathrm{SU}(2)$, $\mathrm{SU}(3)$; $L = 4, 8, 16, 32$; $\beta = 1{:}14$; 64×2 LF	learned F/R momenta, widths, analytic and learned kicks, drift	trained mainly with 16-stage paths; frozen before production	2×8192 paths/target for $U(1)$; 2×1024 for $\mathrm{SU}(2)$/$\mathrm{SU}(3)$; Wilson loops and $U(1)$ topology
4D $\mathrm{SU}(3)$ extension	$L = 4$, $\beta = 1{:}14$; $L = 8$, $\beta = 6$; $K = 2, n = 1$ main path	scalar, wide, multi-operator, covariance variants	short training across β values; transfer to $L = 8$	small-volume SNIS/RT estimates; Wilson and Symanzik
$U(1)$ pilot	8×8; 20×5 LF	CNN, channels 48/48	batch 16; 30000 epochs; $10^{-3} \to 5{\times}10^{-5}$	16 chains; 2000/5000 RT transitions at $\beta = 4/8$; 400/1000 burn-in
$\mathrm{SU}(2)$ learned	6×6; $(10 \text{ or } 20) \times 5$ LF	CNN, channels 48/48	batch 16; 8000 epochs; $10^{-3} \to 5{\times}10^{-5}$	8192 Haar paths; 8 RT chains $\times 300$ after 60
Haar $\mathrm{SU}(2)$	$L = 4, 6, 8$; 10×5 LF	gauge-equiv. CNN, 48/48, 243k	batch 16; 8000 epochs; $10^{-3} \to 5{\times}10^{-5}$	8192 SNIS; 8 IMH $\times 2000$; 8 RT $\times 1500$; HMC $n_{\mathrm{lf}} = 10$
Haar $\mathrm{SU}(3)$	$L = 4, 6$; 10×5 LF	gauge-equiv. CNN, 48/48, 252k	batch 8; 8000 epochs; $10^{-3} \to 5{\times}10^{-5}$	same protocol as Haar $\mathrm{SU}(2)$

For a saved Markov chain, the corresponding mode error uses unweighted post-burn frequencies instead. We do not compare these as though they were the same finite-sample estimator. Historical chain tables explicitly identify that second definition.

The new frozen-ablation acceptance diagnostic draws a current index from the weighted empirical path pool and a proposal index uniformly from that pool. Its conditional expectation is

$$\widehat{\alpha}_{\mathrm{pool}} = \sum_i \bar{w}_i \frac{1}{N} \sum_j \min\{1, e^{W_i - W_j}\}. \tag{151}$$

We estimate this sum by sampling random pairs from the empirical pool. It measures stationary acceptance under that pool's weighted distribution. Three repeated evaluations use distinct path seeds; comparisons between flip probabilities within a seed are paired. Uncertainties for these frozen DW evaluations use sample standard deviation with divisor $R - 1$ for $R = 3$ evaluations.

For weighted SNIS estimates with unnormalized weights w_i, the normalized effective sample size is

$$\mathrm{ESS}/N = \frac{(\sum_i w_i)^2}{N \sum_i w_i^2}. \tag{152}$$

This quantity is invariant to multiplying all weights by a positive constant and is separate from autocorrelation-based Markov-chain effective counts. The log-weight spread is the sample standard deviation of the log weights used for the weighted estimator. Free-energy or log-normalizer error is computed from $\log \widehat{Z} = \log(N^{-1} \sum_i w_i)$ against the exact or named reference for that result. IMH chains are summarized by acceptance and integrated autocorrelation time. For multimodal targets we record observed mode coverage and mode-mass error. For LJ13, energy-W_2 is the one-dimensional 2-Wasserstein distance between target-energy distributions. Geometric W_2 first uses a Hungarian assignment for particle permutations and Kabsch alignment for rigid rotations, then computes the 2-Wasserstein distance between aligned point clouds. It is estimated with 400 generated and 400 reference configurations over three subsample seeds. Pair-distance W_2 is computed from the pooled distribution of all 78 interparticle distances. Raw and path-SNIS values use the same 50000-path proposal pool; the reference contains 10000 configurations from Klein et al. (2023). For the ϕ^4 single-chain comparison, standard errors of linear trace averages use the autocorrelation-based effective count. Susceptibility and Binder-cumulant uncertainties use contiguous-block delete-one jackknife estimates; the high-statistics HMC reference is analyzed with the same observable definitions.

For any weighted histogram, mode-mass, or action-overlap panel, the plotted finite-sample object is the weighted empirical measure $\widehat{\mu}_N^w = \sum_i \bar{w}_i \delta_{x_i}$ with $\bar{w}_i = w_i / \sum_j w_j$. If a panel displays an SNIS-resampled visual set, the resampling is from this empirical measure. Weighted histograms and resampled panels are therefore visualizations of the same finite-sample weighted measure.

For the rotated DW8 action-overlap comparison, the overlap is the common histogram mass between the reference action distribution and the weighted NHMC action distribution. With normalized histogram masses $\widehat{p}_b$ and $\widehat{q}_b$ over common bins, this is

$$\mathrm{overlap}(p, q) = \sum_b \min\{\widehat{p}_b, \widehat{q}_b\}. \tag{153}$$

For the compact $U(1)$ round-trip pilot, uncertainties are standard errors across the 16 chains. Topological sectors are assigned by rounding the measured charge to the nearest integer after burn-in, and only the aggregate number of adjacent sector changes is reported. These traces are not used to estimate topological susceptibility or autocorrelation. The maximum inverse-map, momentum-reversal, and work-ratio residuals are evaluated directly from the recorded paths.

C.5 Lattice ϕ^4 Settings

The learned finite-volume scan uses the settings in Table 5; these differ from the fixed-Gaussian mechanism experiment in Appendix E.1.

Table 5: ϕ^4 settings for the earlier learned finite-volume observable comparison.

Component	Setting	Values	Reported quantity
Target family	2D scalar lattice ϕ^4	$\lambda = 0.022$; κ varied or fixed	Boltzmann observables and importance weights
8×8 scan	NHMC proposal evaluation	$\kappa = 0.20, \ldots, 0.30$; $N = 102400$ proposals per κ	raw and SNIS-weighted $\langle\lvert M\rvert\rangle$, χ, Binder cumulant, and ESS
8×8 reference	HMC grid	60 chains; 10000 steps; burn-in 2000; save every step; $\tau = 1.0$; $\epsilon = 0.03$	high-statistics reference $\langle\lvert M\rvert\rangle$, χ, and Binder cumulant
Observable panel	NHMC and HMC comparison	raw proposal, SNIS-corrected estimate, HMC reference, and ESS	main ϕ^4 benchmark panel

D MANY-WELL RESULTS

D.1 INDEPENDENT-PATH EVALUATIONS OF FROZEN MODELS

Table 1 reports independent-path evaluations of frozen models. The DW4 and DW8 models use, respectively, an annealed block architecture and a force-conditioned block architecture. The DW8 continuous model receives analytic score information. Training settings and the treatment of discrete-label gradients are in Appendix C.3.

Table 6: Frozen sign-flip interventions. All entries are means $\pm$ sample standard deviations across three evaluation seeds of one trained model. The acceptance column is an empirical stationary path-IMH acceptance estimate from the weighted path pool, not a measured post-burn chain acceptance.

Model	Flip	ESS/N	SNIS mode L_1	$\lvert\Delta \log Z\rvert$	Est. acc.
DW2 block	historical	0.587 ± 0.008	0.0003 ± 0.0002	0.0003 ± 0.0000	0.568 ± 0.002
DW2 block	0	0.590 ± 0.012	0.0005 ± 0.0003	0.0014 ± 0.0013	0.586 ± 0.001
DW2 block	0.05	0.561 ± 0.030	0.0005 ± 0.0004	0.0014 ± 0.0020	0.574 ± 0.001
DW4 MLP	historical	0.182 ± 0.011	0.0018 ± 0.0004	0.0033 ± 0.0024	0.299 ± 0.002
DW4 MLP	0	0.171 ± 0.025	0.0026 ± 0.0015	0.0022 ± 0.0016	0.297 ± 0.001
DW4 MLP	0.05	0.181 ± 0.010	0.0020 ± 0.0009	0.0036 ± 0.0020	0.299 ± 0.000
DW4 block	historical	0.407 ± 0.004	0.0012 ± 0.0005	0.0008 ± 0.0005	0.428 ± 0.001
DW4 block	0	0.437 ± 0.017	0.0028 ± 0.0017	0.0034 ± 0.0013	0.457 ± 0.000
DW4 block	0.05	0.415 ± 0.004	0.0013 ± 0.0006	0.0005 ± 0.0004	0.434 ± 0.001
DW8 block	historical	0.142 ± 0.016	0.0070 ± 0.0018	0.0023 ± 0.0015	0.255 ± 0.001
DW8 block	0	0.134 ± 0.022	0.0107 ± 0.0038	0.0084 ± 0.0056	0.249 ± 0.002
DW8 block	0.05	0.150 ± 0.005	0.0067 ± 0.0028	0.0046 ± 0.0027	0.252 ± 0.001

All three evaluation seeds are retained for each condition. Means and sample standard deviations are computed over those three evaluations; they do not estimate variability across independently trained models. Model selection used held-out ESS. The choice of a common $p_{\text{flip}} = 0.05$ deployment setting was made after the frozen-intervention comparison, and is therefore an empirical choice rather than a preregistered optimal probability.

D.2 PRESCRIBED SIGN-FLIP PROBABILITIES

A sign-flip label is sampled at each stage and its forward and reverse probabilities are included in the work. The historical models parameterize the stage logits separately from the continuous networks. At evaluation, replacing all logits by a constant changes this discrete law while leaving all continuous parameters fixed. No new continuous-network training is used in this intervention. For $p_{\text{flip}} = 0$, the move is removed, with its log-probability contribution identically zero; the implementation need not evaluate $\log 0$.

Figures 6 and 7 report the full prescribed-probability range. The historical schedule is plotted as a separate categorical point, not represented as a constant probability when its stage values differ. Across these evaluations, setting a constant probability retains complete observed sign-mode coverage. This is a statement about frozen deployment, not a comparison of training from scratch with and without discrete moves.

D.3 SECTOR CHANGES ALONG A CONTINUOUS PATH

With discrete flips disabled, define the paired-path Hamming distance

$$H_{\text{sign}}(\Gamma) = \sum_{j=0}^{d/2-1} \mathbf{1}\{\operatorname{sign} x_{0,2j} \neq \operatorname{sign} x_{T,2j}\}. \tag{154}$$

Figure 8 uses 49,152 paths per model, pooled from three independent evaluations of 16,384 paths. The probabilities of at least one changed sign are 0.4324, 0.5320, and 0.8378 for updated DW4, the

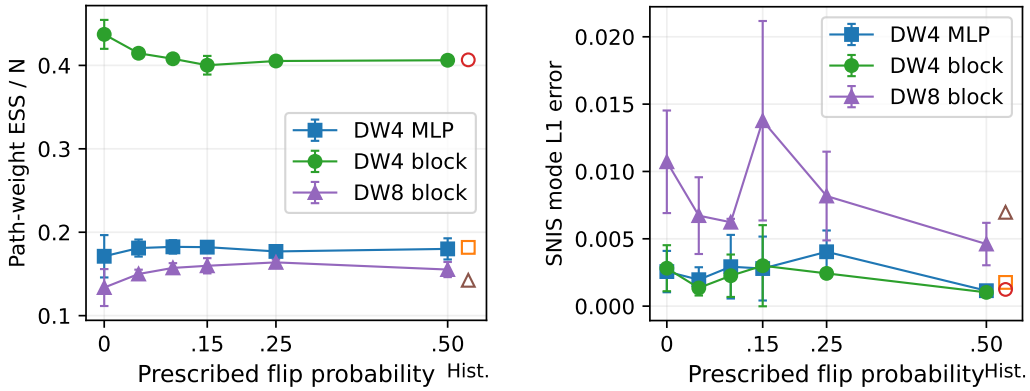

Figure 6: Frozen sign-flip interventions for the updated DW4 and DW8 models and the earlier DW4 MLP. Left: independent-path ESS/N. Right: SNIS mode probability error. Bars are sample standard deviations across three evaluation seeds. "Hist." denotes each model's historical stage schedule.

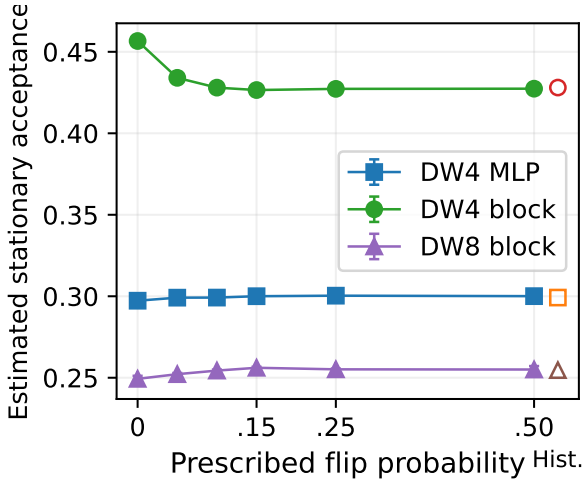

Figure 7: Empirical stationary path-IMH acceptance estimated from each frozen weighted path pool. These are acceptance-function estimates, not post-burn acceptances from newly run Markov chains. Error bars have the same three-evaluation definition as Fig. 6.

DW4 MLP, and updated DW8. Because initial and final signs belong to the same path, these changes cannot be explained by the initial reference covering different modes across different samples. They do not, however, establish that the historical sign moves had no effect during training.

D.4 Earlier path and chain evaluations

The learned-momentum Hamiltonian DW run uses the many-well density

$$\log \gamma_d(x) = \sum_{j=0}^{d/2-1} \left[-x_{2j}^4 + 6x_{2j}^2 + 0.5x_{2j} - \frac{1}{2}x_{2j+1}^2 \right], \tag{155}$$

with analytic $\log Z$ and exact target sampling available by the one-dimensional factorization. The neural NHMC proposal starts from $q_0 = \mathcal{N}(0, I)$, samples learned momenta, and applies leapfrog updates with the target force. With forward and reverse learned momentum laws $q_{\theta,t}^F$ and $q_{\theta,t}^R$, and

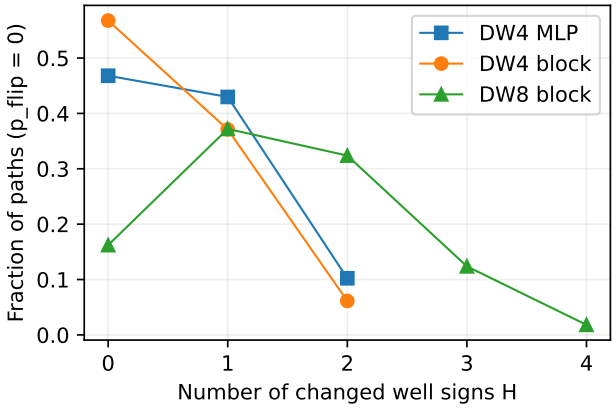

Figure 8: Paired initial-to-final sign changes with the discrete move disabled. The histogram measures changes caused by the continuous Hamiltonian path. It is distinct from mode coverage over independent paths and from sector transitions in a Metropolis chain.

with $\bar{p}_{t+1}$ denoting the output momentum of the Hamiltonian map at stage t, the recorded work is

$$\begin{aligned} W_\theta(\Gamma) = \; & \log q_0(x_0) - \log \gamma(x_T) \\ & + \sum_{t=0}^{T-1} \big[\log r^F_{\theta,t}(s_t \mid x_t) - \log r^R_{\theta,t}(s_t^\dagger \mid x_{t+1}) \\ & \qquad + \log q^F_{\theta,t}(p_t \mid \widetilde{x}_t) - \log q^R_{\theta,t}(-\bar{p}_{t+1} \mid x_{t+1})\big]. \end{aligned} \tag{156}$$

and the correction weight is proportional to $\exp[-W_\theta(\Gamma)]$. A stochastic sign flip with the historical stage schedule is included in the recorded path. The move is an involution, but its stage-indexed forward and reverse log probabilities are retained explicitly in the recorded work above.

These earlier selected evaluations use their original training and deployment settings. They are separate from the fixed-probability independent-path results in Table 1.

Separate DW4 and DW8 checkpoint ensembles support the configuration-space round-trip kernel of Theorem 2. Table 7 uses three independently trained checkpoints per target and fixes the total number of evaluated NHMC paths within each comparison. Round-trip NHMC-MH improves acceptance and autocorrelation per transition on both targets. After accounting for its two paths per transition, it is less force-efficient on DW4 and approximately matched to independent path-IMH on DW8.

Table 7: DW4 and DW8 independent path-IMH and round-trip NHMC-MH across three trained checkpoints per target. These checkpoint ensembles are separate from the selected main evaluations; the DW4 ensemble uses 32 proposal stages and five leapfrog steps per stage. Values are means $\pm$ sample standard deviations. Within each target, both kernels use the same total number of evaluated NHMC paths; a round-trip transition uses two paths. Chains start from endpoints generated by a Gaussian base draw and one forward NHMC path. For every checkpoint, 32 chains retain 2000 path-IMH transitions after discarding 400, or 1000 round-trip transitions after discarding 200; both protocols therefore evaluate the same total number of NHMC paths. Mode masses use all retained endpoints. The reported $\tau_{\mathrm{int}}(x_0)$ is the median across chains, using an FFT autocorrelation estimate with a self-consistent window and a maximum lag of one quarter of the retained chain length. The force-cost column multiplies this value by the number of target-force evaluations per transition.

Target	Kernel	Paths / transition	Acceptance	$\tau_{\mathrm{int}}(x_0)$	Path-cost τ_{int}	Force-cost τ_{int}	Mode L_1
DW4	Independent path-IMH	1	0.3278 ± 0.0025	2.19 ± 0.09	2.19 ± 0.09	421 ± 18	0.0050 ± 0.0016
DW4	Round-trip NHMC-MH	2	0.3537 ± 0.0034	1.74 ± 0.05	3.49 ± 0.10	670 ± 19	0.0109 ± 0.0055
DW8	Independent path-IMH	1	0.1262 ± 0.0011	9.08 ± 0.86	9.08 ± 0.86	1090 ± 103	0.0480 ± 0.0107
DW8	Round-trip NHMC-MH	2	0.1587 ± 0.0045	4.46 ± 0.25	8.91 ± 0.51	1070 ± 61	0.0410 ± 0.0089

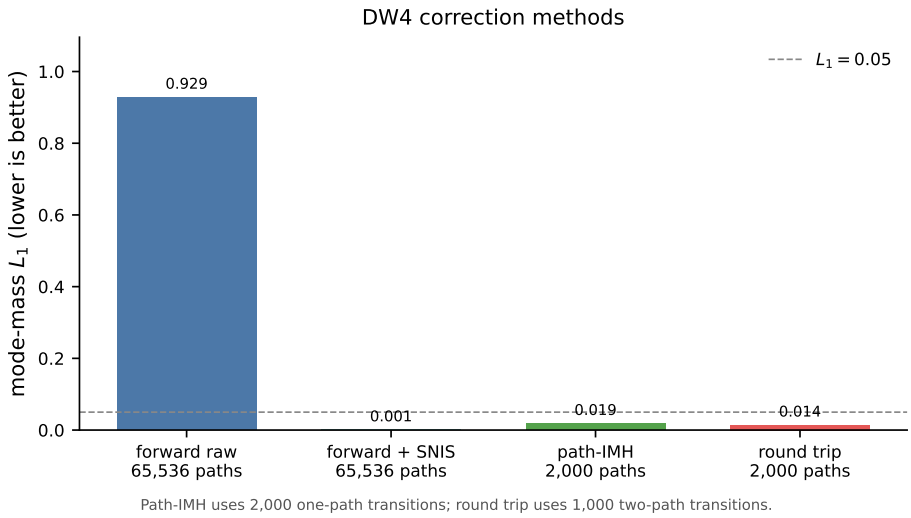

Figure 9: DW4 mode-mass error under four statistically valid uses of the NHMC proposal. Raw and path-SNIS estimates share 65536 forward paths. Independent path-IMH uses 2000 one-path transitions; round-trip NHMC-MH uses 1000 two-path transitions, so the two chains have the same total path budget. The forward-path work is not reused as an importance weight after a round-trip transition.

Table 8: Bidirectional path bounds on analytically normalized double-well targets. The path evidence lower bound is $-\mathbb{E}_{Q_\theta}[W_\theta]$, and the path evidence upper bound is $-\mathbb{E}_{P_\theta}[W_\theta]$. Entries are means and standard errors over three trained checkpoints, using 50000 forward and 50000 reverse paths per checkpoint.

Target	path ELBO	exact $\log Z$	path EUBO
DW4	19.410 ± 0.010	20.587	21.579 ± 0.011
DW8	38.668 ± 0.010	41.174	43.385 ± 0.013

The inequalities in Table 8 follow directly from

$$-\mathbb{E}_{Q_\theta}[W_\theta] \leq \log Z \leq -\mathbb{E}_{P_\theta}[W_\theta], \tag{157}$$

because the two gaps are respectively $\mathrm{KL}(Q_\theta \| P_\theta)$ and $\mathrm{KL}(P_\theta \| Q_\theta)$. Both analytically known normalizers lie inside the measured bidirectional path bounds.

The DW2 trajectory in Fig. 10 is retained as an illustration of the earlier path construction. It uses its historical sign-flip schedule and is not a visualization of the fixed-$p_{\mathrm{flip}} = 0.05$ evaluations in the main table. The coordinate labelled x_0 in this image is the first component of the configuration, not the initial configuration of a path.

D.5 Rotated DW8 Endpoint-Density Ablation

The native many-well target factorizes across coordinate pairs, so a factorized proposal can be unusually favorable. To remove that coordinate advantage, we rotate the DW8 target by a fixed orthogonal map R,

$$\gamma_R(x) = \gamma(xR), \tag{158}$$

which preserves the analytic log normalizer and target sampler but destroys coordinate-wise factorization in the observed coordinates. We then train a global affine proposal map on top of the many-well base proposal with a rotation curriculum. This endpoint-density ablation asks whether proposal training restores concentrated correction weights on a coupled version of the same Boltzmann target.

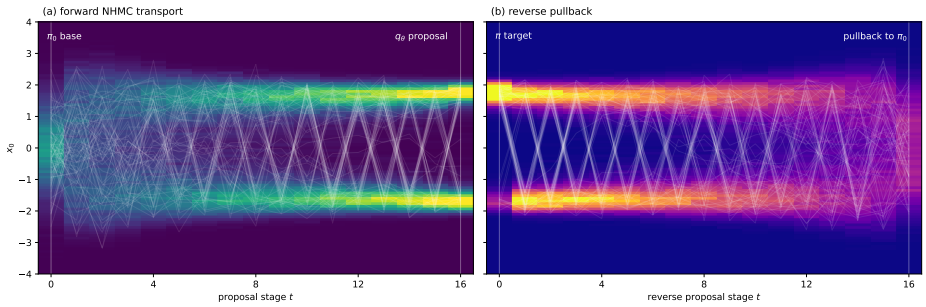

Figure 10: Illustrative DW2 forward paths and reverse pullback from the earlier model. The horizontal coordinate is path stage (not Markov time); the vertical x_0 printed in this figure denotes the first configuration component. The historical sign schedule is retained here. This illustration is not an evaluation of the updated fixed-$p_{\text{flip}} = 0.05$ models in Table 1.

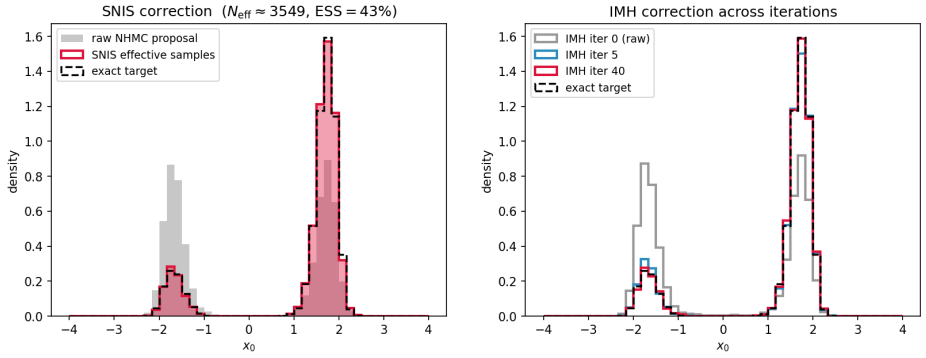

Figure 11: DW2 correction on the asymmetric two-well target. SNIS and IMH correction recover the intentionally asymmetric target masses; the right mode is heavier because the target factor includes the linear tilt $0.5x$.

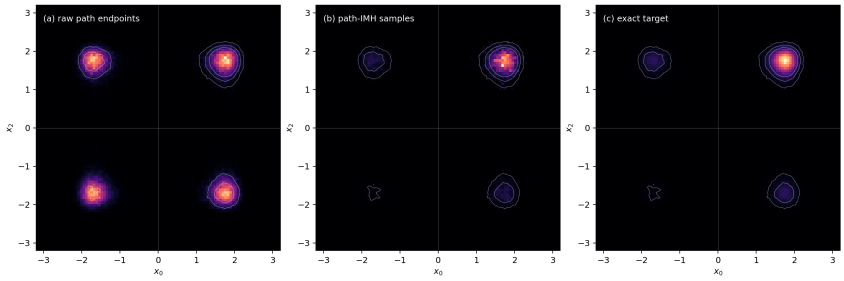

Figure 12: DW4 endpoints and corrected mode probabilities. The learned proposal covers all four sign modes; path-IMH changes the finite-sample mode frequencies toward the unequal analytic target probabilities. White contours show the target density.

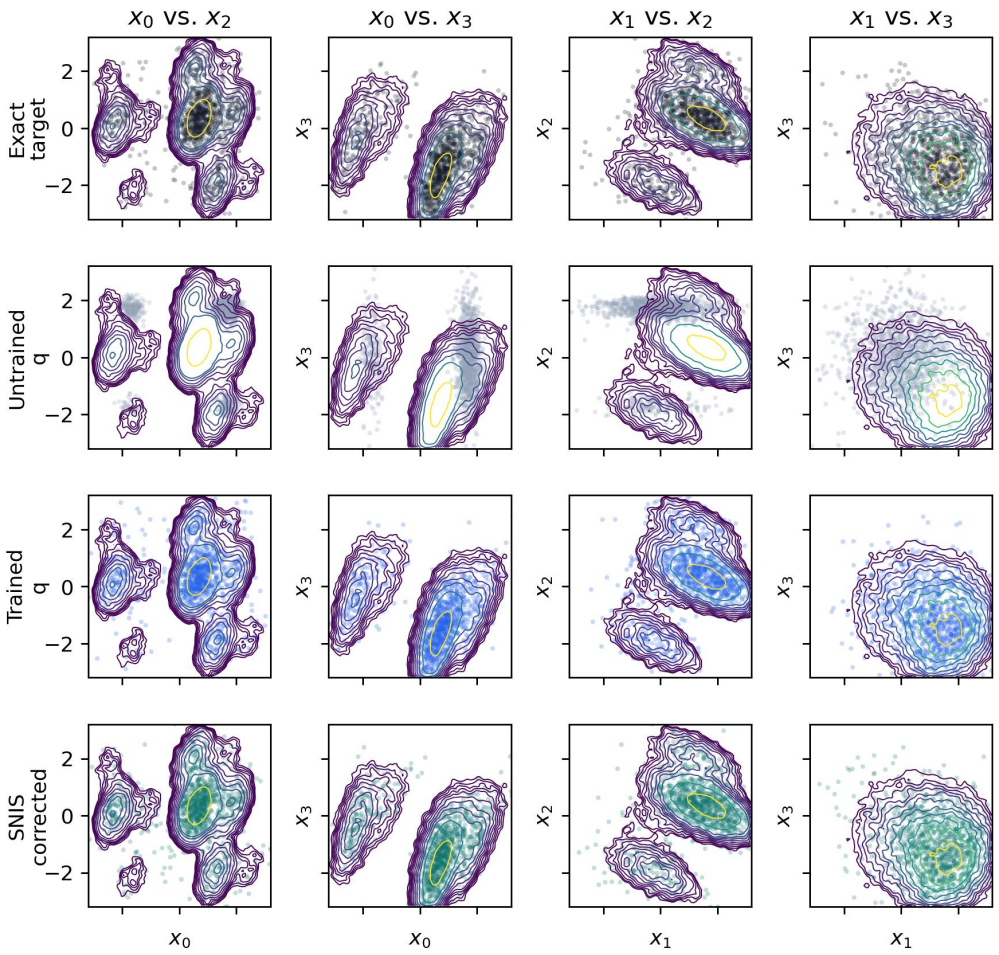

Figure 13: Rotated DW8 pairwise marginal coverage for the endpoint-density ablation. Contours are empirical pairwise target marginals from the analytic target sampler. The untrained factorized proposal is misaligned in the rotated coordinates; the trained proposal and its SNIS-resampled visual set track the target marginals under the same correction weights.

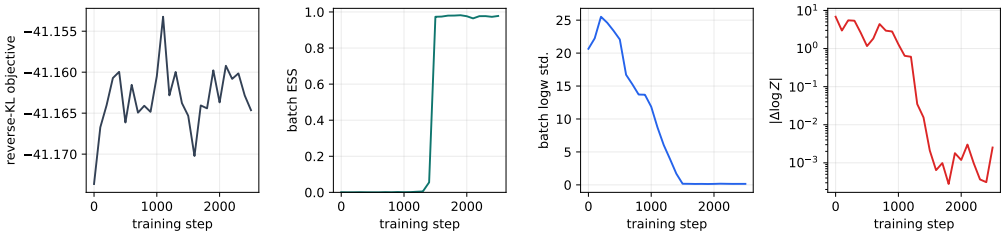

Figure 14: Rotated DW8 endpoint-density training trajectory. The curriculum reaches the final rotated target at step 1500; after training the batch ESS stabilizes near one and the log-weight spread collapses.

D.6 Action Overlap for Rotated DW8

The rotated-DW8 action distributions use one fixed $N = 50000$ evaluation. The factorized proposal provides the pre-training baseline and the trained affine proposal the learned map. The target action is $-\log \gamma_R(x)$. Training changes the action overlap from 1.46% to 99.56%, collapses the log-weight spread from 19.97 to 0.157, and gives 96.67% ESS and 91.58% stationary IMH acceptance. The SNIS-resampled visual set has 99.38% action overlap and small mode-mass error.

Table 9: Unified rotated-DW8 endpoint-density evaluation with $N = 50{,}000$. The untrained q_0 row is a before-training control. Action overlap is the common histogram mass between proposal and exact target actions; the SNIS-resampled row visualizes the weighted empirical measure.

Sample set	action overlap	mode L_1	ESS	logw std	$\lvert\Delta \log Z\rvert$	IMH acc.
untrained q_0	1.46%	1.413	0.02%	19.966	5.3	$< 0.01\%$
trained q_θ	99.56%	0.011	96.67%	0.157	0.000279	91.58%
SNIS corrected	99.38%	0.016	–	–	–	–

E Scalar-Field Round-Trip Experiments

E.1 Reference-bank starts and one-step measurements

The stationary-flow and sector-relaxation experiments use the same zero-field action, $L = 8$, $\lambda = 0.022$, and $\kappa = 0.27$. Their round-trip paths have $T = 20$ stages, five leapfrog substeps per stage, and step size 0.135. Both momentum laws are fixed Gaussians. No learned momentum model is used in these two experiments.

The stored HMC reference bank is generated from 128 chains initialized with both signs. Each chain uses 2000 discarded trajectories, 2000 production trajectories, storage every 20 trajectories, HMC step size 0.05, and ten leapfrog substeps. The stored pool contains 6461 positive- and 6339 negative-sector configurations. The one-step diagnostic samples 4096 starting configurations with equal numbers from these two empirical sector pools. This stratification imposes the known zero-field sector masses; it does not by itself prove equilibration within each sector.

For exact $X \sim \pi$ and a fresh reverse record $B \sim r_\theta(\cdot \mid X)$, Eq. (11) gives $(U, A) \sim \rho_\theta h_\theta$, hence $U \sim \rho_\theta$. The experiment approximates those exact starts using the finite HMC bank. A new reverse record is drawn for each outer sample, and one round trip from the same starting configuration estimates $F^{\mathrm{RT}}_{+-} = \Pr_\pi(X \in A_+, X' \in A_-)$. This joint probability is not divided by the positive-sector mass.

Table 10: One-step round-trip measurements at $L = 8$, $\kappa = 0.27$, $\lambda = 0.022$, $h = 0$, with fixed Gaussian momentum laws. There are 4096 outer draws and $K = 32$ inner proposals per bridge. Errors are one-standard-error estimates across outer draws, conditional on the stored HMC bank; they do not include the inner-SNIS bias or the uncertainty of the reference bank.

p_{flip}	$\bar{\alpha}_{\mathrm{RT}}$	F^{RT}_{+-}	$\widehat{\Omega}_{+-}$
0.00	0.9575 ± 0.0032	0.0347 ± 0.0029	0.0363 ± 0.0011
0.05	0.9561 ± 0.0032	0.0723 ± 0.0040	0.0752 ± 0.0010
0.10	0.9592 ± 0.0031	0.1077 ± 0.0048	0.1094 ± 0.0009
0.15	0.9561 ± 0.0032	0.1477 ± 0.0055	0.1407 ± 0.0008

E.2 Nested estimation of the ideal flow

For each fixed bridge u, draw K independent forward records $a_j \sim f_\theta(\cdot \mid u)$. An unnormalized conditional density is $\widetilde{h}_u(a) = \gamma(\mathsf{E}_\theta(u, a)) r_\theta(b(u, a) \mid \mathsf{E}_\theta(u, a))$. Thus

$$v_j = \frac{\widetilde{h}_u(a_j)}{f_\theta(a_j \mid u)}, \qquad \widehat{m}_+(u) = \frac{\sum_j v_j \mathbf{1}_{C_+(u)}(a_j)}{\sum_j v_j}, \qquad \widehat{\Omega}_{+-} = \frac{1}{N_{\mathrm{out}}} \sum_i \widehat{m}_+(U_i)[1 - \widehat{m}_+(U_i)]. \tag{159}$$

Under conditional support and finite first-moment assumptions, inner SNIS converges as $K \to \infty$. With independent exact outer draws, the resulting nested estimator is consistent in a suitable joint or iterated limit. Increasing only N_{out} at fixed K does not remove inner bias.

Even when $f_u = h_u$, let $\widehat{m} = K^{-1} \sum_{j=1}^K I_j$, with $I_j \mid u$ independent Bernoulli variables of mean $m = m_+(u)$. Then

$$\mathbb{E}[\widehat{m} \mid u] = m, \qquad \mathrm{Var}(\widehat{m} \mid u) = m(1 - m)/K, \tag{160}$$

$$\mathbb{E}[\widehat{m}(1 - \widehat{m}) \mid u] = m - m^2 - \mathrm{Var}(\widehat{m} \mid u) = (1 - K^{-1}) m(1 - m). \tag{161}$$

At the reported $K = 32$ this is a 3.125% downward bias in that equal-weight case. General SNIS has additional ratio bias, so multiplying the estimate by $32/31$ is not an exact correction. The error bars in Table 10 concern variation across outer draws, conditional on the stored bank; they omit this inner bias and reference-bank uncertainty. No larger-K comparison is included in the present data. Accordingly, the claim is the observed joint increase of estimated ideal and accepted flow, not precise equality or a numerical saturation of the bound.

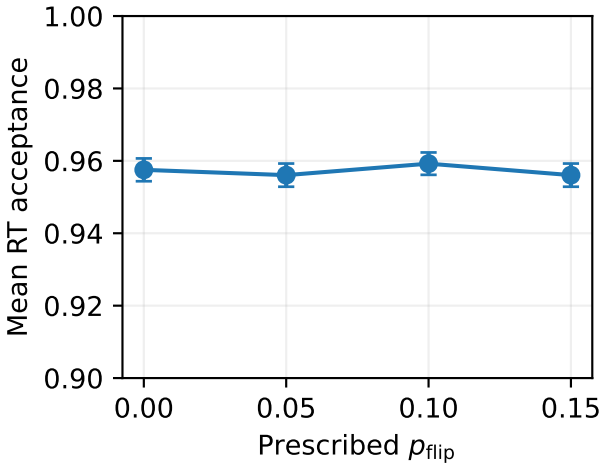

Figure 15: Mean round-trip acceptance for the same reference-bank starts as Table 10. The vertical scale is fixed to $[0.9, 1]$. The small variation in acceptance is separate from the increase in the accepted inter-sector joint flow.

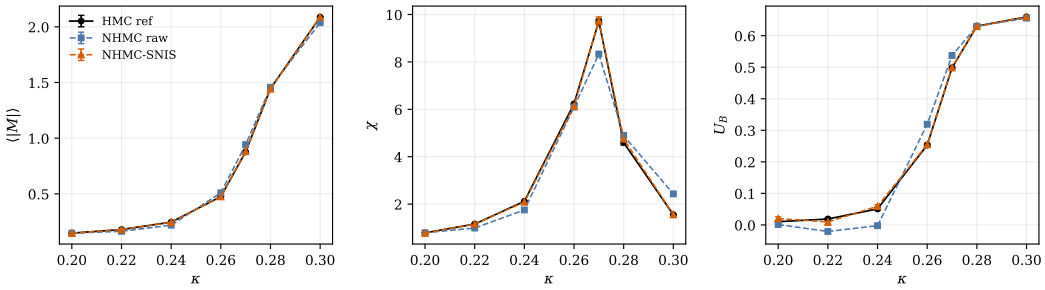

Figure 16: Earlier learned 8×8 scalar-field scan. Raw and path-SNIS observables are compared with an independent HMC reference using the same susceptibility and Binder-cumulant definitions. This learned model is not the fixed-Gaussian model in the stationary-flow diagnostic.

E.3 Sector-weight relaxation

The relaxation experiment initializes 512 chains with positive-sector probability approximately 0.9, using the stored within-sector samples. It then applies 100 round-trip transitions with fixed p_{flip}. The population paths in Fig. 2 are not stationary flow measurements. They test the finite-time consequence of biased sector weights, while Appendix B.9 establishes the exact first-step relation for target within-sector conditionals. The reference-bank approximation and finite-chain sampling remain part of this empirical test.

E.4 Earlier learned finite-volume scan

Table 11 summarizes the two-dimensional 8×8 κ scan and the smaller 4×4 correction regime.

Table 11: Finite-volume ϕ^4 NHMC-SNIS estimates.

Lattice	Setting	Main quantities	ESS fraction
8×8	κ scan, $N = 102400$	weighted magnetization, susceptibility, and Binder cumulant	0.054–0.599
4×4	same κ range	small-volume correction regime	0.36–0.83

The zero-field scalar action is invariant under the global sign flip $\phi_x \mapsto -\phi_x$. In finite volume this means the signed magnetization $M = V^{-1} \sum_x \phi_x$ cancels under sector-balanced target sampling, giving $\langle M \rangle = 0$. The main text therefore uses $|M|$, M^2, susceptibility, and Binder-cumulant quantities. Table 12 gives sign-sector balance for the 8×8 scan. It pairs the proposal sign fraction with the SNIS-corrected observable estimates and ESS, quantifying finite-budget sector coverage.

Table 12: Z_2 sign-sector summary for representative points in the ϕ^4 8×8 scan. The column $\widehat{p}_+$ is the SNIS-weighted positive-sign fraction; corrected observables and ESS use the same SNIS weights.

κ	N	$\widehat{p}_+$	$\lvert\widehat{p}_+ - 1/2\rvert$	SNIS $\langle\lvert M\rvert\rangle$	SNIS χ	ESS/N
0.20	102400	0.5009	9.0×10^{-4}	0.147	0.777	0.599
0.27	102400	0.5038	3.8×10^{-3}	0.881	9.734	0.054
0.30	102400	0.4917	8.3×10^{-3}	2.087	1.554	0.065

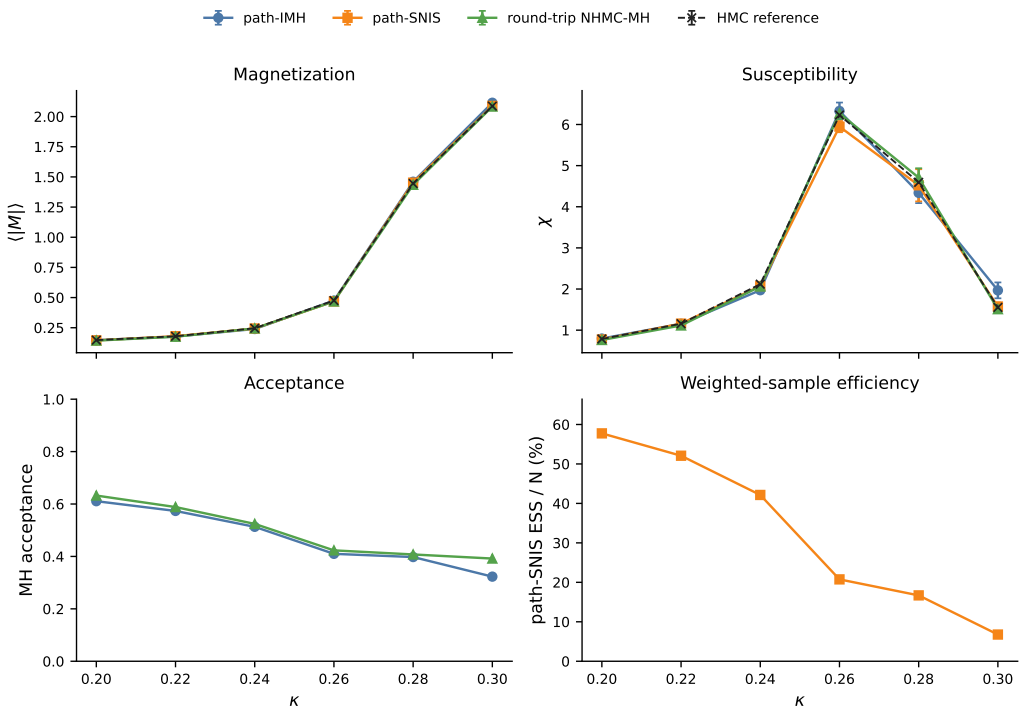

Figure 17: Three correction modes from the same κ-conditional 8×8 ϕ^4 proposal. Path-SNIS uses 12000 forward paths per coupling; path-IMH and round-trip NHMC-MH each report 10000 post-burn states from 11000 transitions. Error bars use weighted estimates for path-SNIS and trace-aware estimates for the two chains. The HMC curve is an independent high-statistics reference. The lower panels show the acceptance of the Markov corrections and the path-SNIS ESS.

The 8×8 $|M|$ autocorrelation comparison uses one HMC chain, one independent NHMC path-IMH chain, and one round-trip NHMC-MH chain at $\kappa = 0.2705$. The HMC and path-IMH chains each have 5000 saved states and discard the first 200. The round-trip chain is initialized by drawing from the Gaussian base and running one forward NHMC path; it then runs for 10000 transitions and discards the first 1000. The comparison uses the first 4800 post-burn values from each chain with the same dot-normalized autocorrelation estimator. The integrated autocorrelation times per saved transition are 19.95, 8.94, and 2.57, respectively. Because a round-trip transition uses two paths, its path-cost-normalized value is 5.15. Using all 9000 post-burn round-trip states gives $\langle|M|\rangle = 0.9066 \pm 0.0100$, $\chi = 9.98 \pm 0.32$, and $U_B = 0.5099 \pm 0.0060$. The corresponding 198000-state HMC reference

values, computed with the same observable definitions, are 0.9068 ± 0.0068, 9.71 ± 0.12, and 0.5097 ± 0.0036. This 198000-state reference is a separate long single-chain HMC run from the 60-chain grid used in Fig. 16. Table 13 gives sector and autocorrelation summaries for all three matched traces.

Table 13: ϕ^4 8×8 matched single-chain summaries at $\kappa = 0.2705$. Each row uses 4800 post-burn states. Acceptance is measured per method-specific transition and is not a force-cost-matched comparison.

Chain	Acceptance	p_+	Sector changes	$\tau_{\mathrm{int}}(\lvert M\rvert)$
HMC	98.50%	0.507	80	19.95
Independent path-IMH	36.67%	0.502	839	8.94
Round-trip NHMC-MH	38.92%	0.465	782	2.57

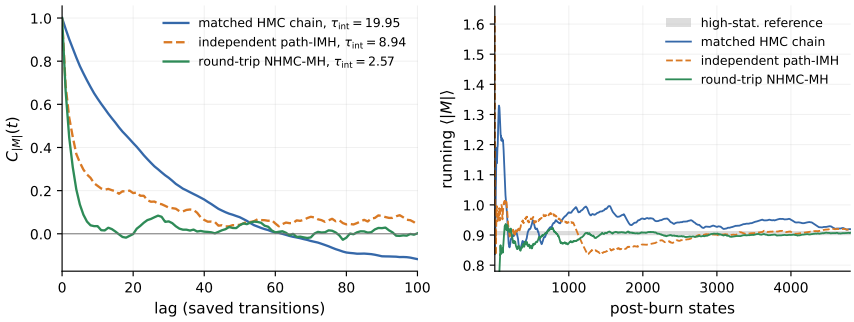

Figure 18: ϕ^4 8×8 single-chain comparison for $|M|$ at $\kappa = 0.2705$. HMC, independent path-IMH, and round-trip NHMC-MH each use 4800 post-burn states. The round-trip chain begins from an NHMC forward endpoint; the HMC chain is used only as an independent reference. The left panel compares autocorrelation per saved transition; the right panel shows running means against the high-statistics HMC reference. A round-trip transition evaluates two NHMC paths.

E.5 Controlled mode weights and sector transport in ϕ^4

We separate mode weights from the rate of sector communication in two controlled calculations at $\kappa = 0.27$. First, a source with $p_+^{q_0} \simeq 0.90$ gives raw endpoint fractions 0.766 and 0.791 at $L = 8$ and 16. Path-SNIS gives 0.483 and 0.472, while round-trip MH gives $0.497(22)$ and $0.482(22)$, close to the symmetric target value $1/2$.

Second, at $L = 8$ we add $h = 0.01$. An independent high-statistics reference gives $p_+^{\mathrm{ref}} = 0.7562206(2480)$. For $p_{\mathrm{flip}} = 0, 0.1, 0.5$, the round-trip chains contain 640000 post-burn states in total and give $p_+ = 0.7566(28), 0.7538(13), 0.7545(7)$. Over the same sequence, $\tau_{\mathrm{int}}(\mathrm{sign}\, M) = 8.88(22), 2.853(43), 1.056(10)$ and the accepted $+ \leftrightarrow -$ flow increases. The $|M|$ autocorrelation stays between 1.68 and 1.95. Thus the target sector weight remains consistent while the proposal-dependent rate of communication between sectors changes.

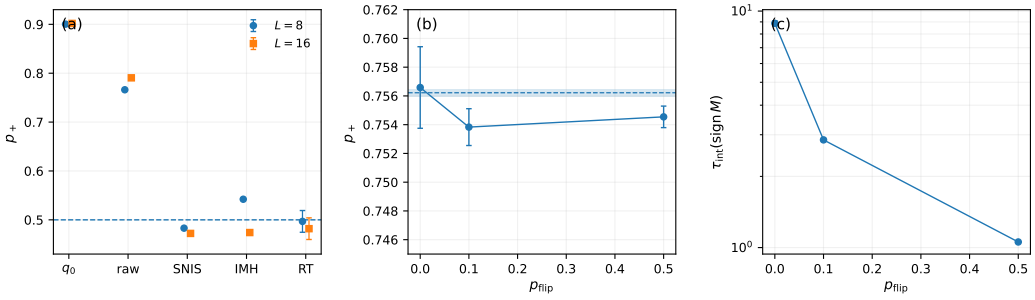

Figure 19: Round-trip MH in the ϕ^4 sector tests. (a) Source, raw-endpoint, and separate corrected estimates for a deliberately biased source; the dashed line is the symmetric target value. The categories are comparisons, not sequential updates. (b) Positive-sector probability for the asymmetric $h = 0.01$ target; the band is the independent reference. (c) Integrated autocorrelation of $\mathrm{sign}\, M$ for the same round-trip chains.

F GAUGE-FIELD RESULTS AND REFERENCE DISTRIBUTIONS

The production scan and direct-neural comparison use structured reference distributions. The path-depth and full-link experiments use separate models with a Haar reference; their sampling settings are given below.

F.1 ANALYTIC-FORCE AND DIRECT-NEURAL COMPARISON

Table 14 compares four proposals on 168 targets, with 56 evaluation points per group. Each direct model is trained for 512 optimizer updates and evaluated conditionally over the grid. It has learned forward/reverse momentum laws, width, a direct kick, and a positive state-independent drift rate, without analytic Wilson force calls or a precomputed force basis. The other dynamic configurations retain the analytic force. All proposals use the same source and paired evaluation seeds. Path lengths and evaluation-pool sizes are given in the table caption.

Table 14: Gauge-field path comparison. Entries are median ESS/N over 56 evaluation targets per group. Three conditional direct models are evaluated, not 168 independently trained models. All dynamic arms use $T = 64$ and two leapfrog substeps per stage; evaluation pools contain 1024 paths for $U(1)$ and 512 for each non-Abelian group. The $U(1)$ arms, including the no-Hamiltonian-path arm, share a recorded post-transport flux move.

Group	No Hamiltonian path	Analytic force	Analytic + learned	Direct NN
$U(1)$	0.157	0.236	0.411	0.247
$\mathrm{SU}(2)$	0.110	0.514	0.728	0.222
$\mathrm{SU}(3)$	0.059	0.642	0.823	0.181

The direct model improves on the no-Hamiltonian-path arm at 52/56, 56/56, and 50/56 targets in $U(1)$, $\mathrm{SU}(2)$, and $\mathrm{SU}(3)$, respectively, while the analytic-plus-learned ESS is larger at every evaluated target. Each group uses one trained model per learned proposal, evaluated over finite path pools. Training and wall-clock costs are not matched.

For the non-Abelian direct models, learning uses maximal-tree reduced coordinates. This removes local gauge redundancy but leaves simultaneous global conjugation. The lower efficiency of the direct model therefore cannot simply be attributed to learning local gauge redundancy, nor does a reduced-coordinate MLP automatically become a fully covariant field model. The covariance construction is stated separately in Appendix G.

The accompanying data contain direct-path weights, direct-model parameters, and per-target baseline summaries. The three historical residual-model parameter sets and complete raw baseline path pools are unavailable in this release. The smaller experiments with a Haar reference below use separate models and evaluation pools.

F.2 PATH DEPTH FOR MODELS TRAINED WITH A HAAR REFERENCE

Table 15: Path-weight ESS/N versus depth for models trained with a Haar reference at $L = 4$, $\beta = 8$, with $N = 1024$ paths per entry. Each group uses one model trained with 16 momentum-refresh stages. All displayed evaluations use four leapfrog substeps per stage. The production residual models in Table 14 are trained separately.

Group	$T = 16$	32	64	128
$U(1)$	0.0049	0.0172	0.0268	0.0100
$\mathrm{SU}(2)$	0.0075	0.0156	0.0484	0.0351
$\mathrm{SU}(3)$	0.0248	0.0275	0.0934	0.1132

The four depths within each group use the same trained parameters. These models use a Haar reference, were trained at 16 fresh-momentum stages, and are separate from the 64-stage, two-substep production models. Increasing T changes the number of momentum draws, not just deterministic

integration resolution. The step and momentum-mean factors vary with T, so each depth defines a different stochastic path law. The path scale is proportional to $\sqrt{\tau/T}$, where τ is held fixed within this comparison. Table 15 reports the resulting weight ESS.

F.3 Structured reference definition

For the two-dimensional gauge experiments we use a tractable structured reference chosen to improve overlap; the path correction itself does not depend on this choice. Consider a periodic two-dimensional $L \times L$ lattice, where L is the number of sites in each direction, $V = L^2$ is the number of plaquettes, and β is the Wilson coupling. With normalized Haar reference measure, write $e_p = (\beta/N_c)\,\mathrm{ReTr}\,P_p$ for $\mathrm{SU}(N_c)$ and $e_p = \beta\cos\theta_p$ for $U(1)$. We use the reduced Wilson action $S = -\sum_p e_p$; the conventional action adds the known constant βV.

A maximal-tree gauge choice leaves $V-1$ independent plaquette matrices and two holonomies A, B around the torus (Creutz, 1977; Boyda et al., 2021). The remaining plaquette P_d is reconstructed from the ordered product and the holonomy commutator $ABA^\dagger B^\dagger$. The tree removes a constant gauge-volume factor, while a simultaneous global conjugation remains. We sample the independent plaquettes from

$$q_d(P_{p\neq d}, A, B) = \prod_{p\neq d} \frac{e^{e_p}}{z_1(\beta)}, \qquad z_1(\beta) = \int d_H P\, e^{(\beta/N_c)\,\mathrm{ReTr}\,P}, \tag{162}$$

with independent Haar holonomies. For $U(1)$, $z_1 = I_0(\beta)$ and the last angle is fixed by periodic closure. All V plaquettes enter the target. Consequently direct importance weighting requires only

$$\log\frac{\gamma(U)}{q_d(U)} = e_d + (V-1)\log z_1(\beta), \tag{163}$$

instead of the full sum $\sum_p e_p$ required by a Haar source. At fixed β, the fluctuating correction in (163) is bounded independently of V.

The distinguished plaquette breaks manifest translation symmetry. A uniform translation mixture, $q_{\mathrm{sym}}(U) = V^{-1}\sum_d q_d(U)$, removes this choice; its normalized density and translation properties are derived in Appendix F.4. The production scan uses the fixed structured reference. The mixture is evaluated by direct importance sampling without learned transport.

F.4 Reference normalization and translation mixture

F.4.1 Reference measure and source normalization

We use normalized Haar measure on the independent group variables. A rooted maximal-tree gauge choice removes $V-1$ pure-gauge variables. The remaining coordinates contain $V-1$ independent, suitably transported plaquettes and two torus holonomies. They retain a simultaneous global conjugation, which leaves every traced Wilson loop invariant. Reconstructing periodic links and sampling the omitted gauge transformations from normalized Haar measure gives a full-link density with the same gauge-invariant expectations. Statements about $q_d(U)$ in Section F.3 refer to this lift, or equivalently to its density in the reduced reference measure. They do not assign a full-dimensional density to a fixed gauge slice.

For a distinguished plaquette d, define

$$c_\beta = z_1(\beta)^{-(V-1)}, \qquad q_d(U) = c_\beta \exp\left(\sum_p e_p(U) - e_d(U)\right). \tag{164}$$

Each independent plaquette integral contributes z_1 and each holonomy integral contributes one, so q_d is normalized. The reconstructed plaquette is absent from this source action but present in the target. For $\mathrm{SU}(N_c)$ it includes the noncommuting ordered plaquette product and $ABA^\dagger B^\dagger$; setting that commutator to the identity would select a different finite-volume distribution. Traced observables are evaluated on the reconstructed periodic links.

The reduced target partition function is $Z = \int d_H U\ \exp(\sum_p e_p)$, and

$$\frac{\gamma(U)}{q_d(U)} = c_\beta^{-1} e^{e_d(U)}. \tag{165}$$

For fixed β, the range of the nonconstant term in its logarithm is independent of the lattice volume. This statement concerns the direct source-to-target ratio. A learned path adds its recorded forward/reverse probability terms and the endpoint energy change.

F.4.2 Translation mixture and density evaluation

Let T_a translate links by $a \in \mathbb{Z}_L^2$. Haar measure and $\sum_p e_p$ are invariant. Translating a source sample moves the distinguished plaquette, so uniform translation produces

$$q_{\mathrm{sym}}(U) = \frac{1}{V}\sum_d q_d(U) = c_\beta\, \gamma(U) A_V(U), \qquad A_V(U) = \frac{1}{V}\sum_p e^{-e_p(U)}. \tag{166}$$

This is the finite-group density average used to symmetrize proposals (Boyda et al., 2021). In the present source its sum is available from the plaquette values, with no repeated neural evaluation. It is invariant under all lattice translations and has direct log weight

$$\log w_{\mathrm{sym}}(U) = -\log c_\beta - \log A_V(U). \tag{167}$$

A random translation alone changes the source law; the corresponding density for position-dependent importance sampling is the mixture density above.

For a fixed U, translating the configuration while keeping the coordinate choice fixed gives

$$\log q_d(T_a U) = \log c_\beta + \sum_p e_p - e_{d-a}, \tag{168}$$

$$\log q_d(T_a U) - \frac{1}{V}\sum_b \log q_d(T_b U) = -(e_{d-a} - \bar{e}), \tag{169}$$

$$\mathrm{Var}_a[\log q_d(T_a U)] = \mathrm{Var}_p[e_p]. \tag{170}$$

Thus the source's translation fluctuation is obtained directly from a plaquette field. We also evaluate this fluctuation by translating and returning to the same tree gauge.

Under a translation-invariant ensemble, put $Y_p = e^{-e_p}$. If $\sup_V \sum_r |\operatorname{Cov}(Y_0, Y_r)| < \infty$ at fixed coupling, then

$$\mathrm{Var}(A_V) = \frac{1}{V^2}\sum_{p,r} \operatorname{Cov}(Y_p, Y_r) \le \frac{C}{V}. \tag{171}$$

Since $e^{-\beta} \le A_V \le e^{\beta}$, the same bound controls fluctuations of $\log A_V$ up to a coupling-dependent constant. This explains the observed improvement with L at fixed β when the local average concentrates; it is not a coupling-uniform continuum bound.

F.4.3 Weight moments under the translation mixture

Let $F(U)$ be invariant under translations. Every translated component has the same expectation of F, hence

$$\mathbb{E}_{q_{\mathrm{sym}}}[F] = \frac{1}{V}\sum_a \mathbb{E}_{q_d}[F(T_a U)] = \mathbb{E}_{q_d}[F(U)]. \tag{172}$$

Both w_{sym} and w_{sym}^2 have this property. We therefore estimate their moments using samples from the fixed reference distribution. The reported finite-sample weight ESS fraction is

$$\frac{\widehat{\mathrm{ESS}}}{n} = \frac{(\sum_i w_i)^2}{n\sum_i w_i^2}. \tag{173}$$

The same argument applies to any translation-invariant loop average, but not to the loop at a specified origin. Position maps in this experiment use the fixed-source weight e^{e_d}, and are labeled accordingly.

F.4.4 Numerical protocol and translation identities

The source comparison covers $U(1)$, $\mathrm{SU}(2)$, and $\mathrm{SU}(3)$, $L = 4, 8, 16, 32$, and $\beta = 1, \ldots, 14$, with 4096 draws per source and parameter combination. The Haar draws use source coupling zero, while the structured source uses the target coupling. The random seed is 20260912. Table 16 compares the reconstructed and analytic densities at $L = 8$, $\beta = 6$ over 16 configurations per group and all translations.

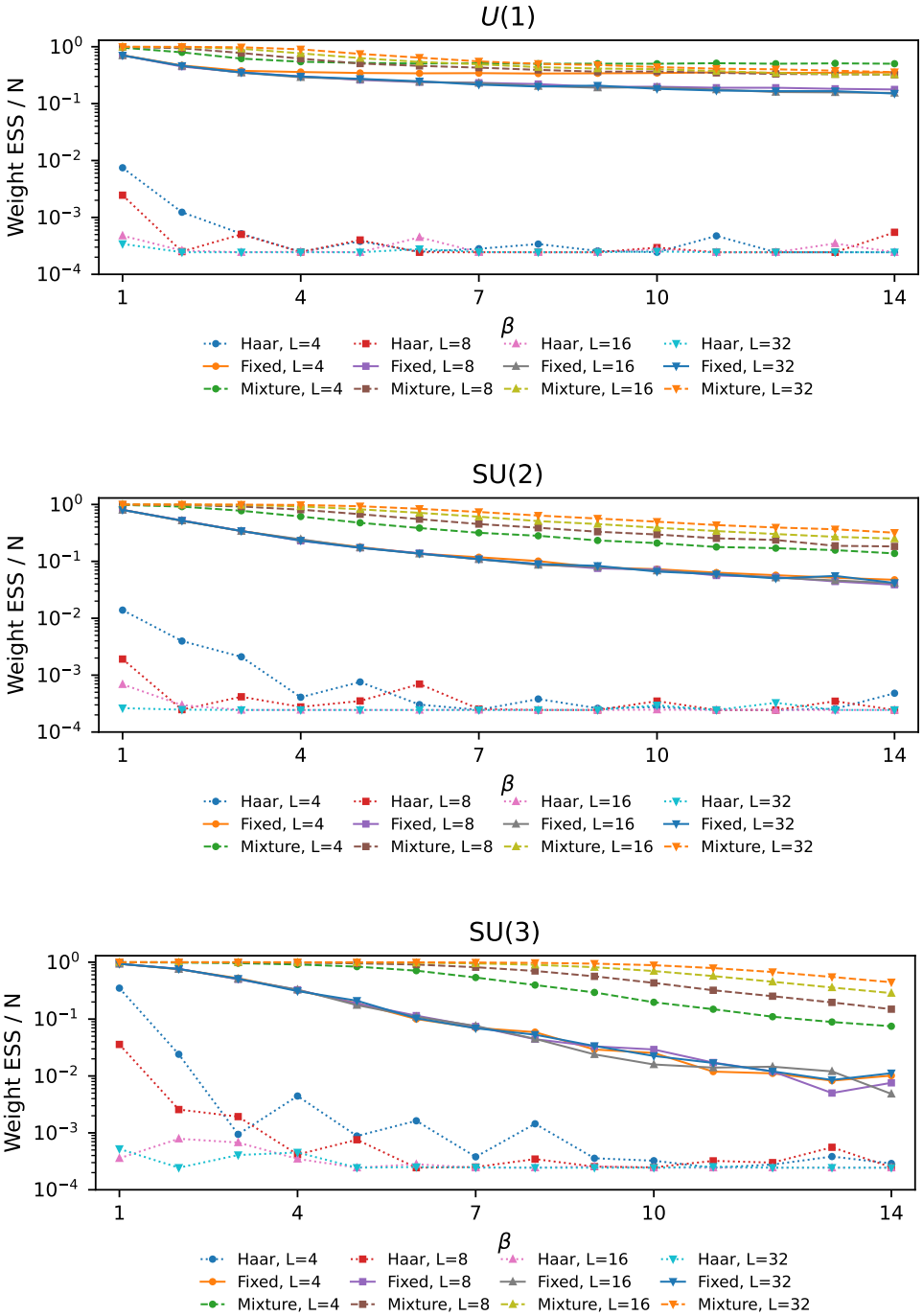

Figure 20: Direct importance-sampling weight ESS over $\beta = 1, \ldots, 14$ and $L = 4, 8, 16, 32$. Each curve corresponds to a reference distribution and volume. Haar pools close to the finite-sample floor do not resolve population ESS.

Table 16: Maximum differences between reconstructed and analytic log densities, and between translated and original actions, at $L = 8, \beta = 6$. The last column compares the variances in (170).

Group	$\lvert \log q_{\text{numerical}} - \log q_{\text{formula}} \rvert$	$\lvert S(T_a U) - S(U) \rvert$	Variance residual
$U(1)$	1.71×10^{-13}	1.14×10^{-13}	3.93×10^{-14}
SU(2)	4.95×10^{-12}	5.68×10^{-14}	2.54×10^{-13}
SU(3)	5.23×10^{-12}	5.68×10^{-14}	3.22×10^{-13}

Table 17: Selected source-only SU(3) weight ESS fractions. Each row uses 4096 draws. The Haar pool is dominated by one or two weights in the listed cases and is marked as unresolved.

L	β	Haar	Fixed source	Translation mixture
4	6	0.0016	0.1002	0.7097
8	6	unresolved	0.1150	0.9102
16	6	unresolved	0.1097	0.9766
32	6	unresolved	0.1033	0.9941
32	10	unresolved	0.0225	0.8856
32	14	unresolved	0.0112	0.4440

F.4.5 Position-resolved plaquette estimates

The spatial test uses 4096 configurations at $L = 8, 16$ and $\beta = 2, 6, 10, 14$. On each full plaquette map we compute

$$D_W = W_{11}(d) - \frac{1}{V-1} \sum_{p \neq d} W_{11}(p). \tag{174}$$

The target expectation is zero. Raw and fixed-source SNIS estimates use the same configuration pool. Bootstrap resampling uses 1000 replicates, with the whole configuration as the resampling unit; lattice origins are not treated as independent observations. The largest raw spatial effect is localized

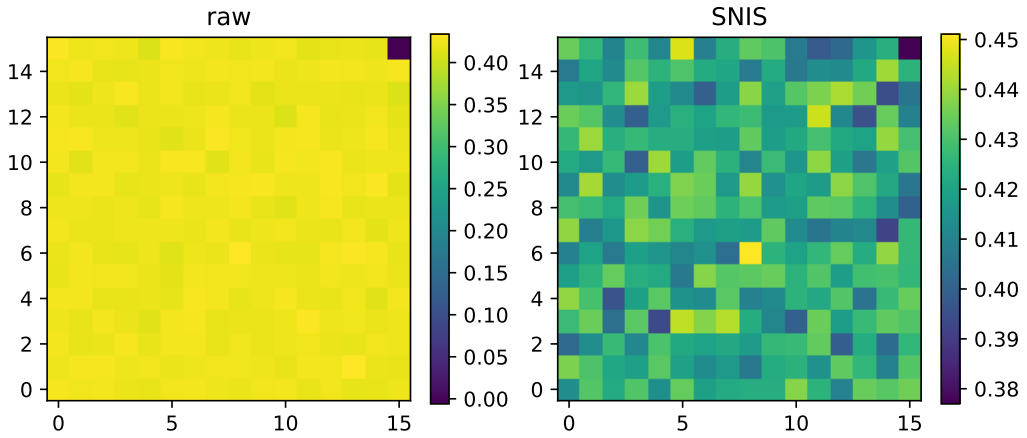

Figure 21: Position-resolved plaquette means for the fixed structured SU(3) source at $L = 16, \beta = 6$. The source has a distinguished plaquette; direct SNIS reduces its large raw displacement. The residual D_W/SE is about -3.61 in this finite pool. Samples are drawn directly from the fixed reference distribution.

at the reconstructed plaquette. Reweighting substantially reduces it. At the more concentrated SU(3) weights, residuals remain: D_W/SE is approximately -3.03 at $L = 8, \beta = 14$ and -3.61 at $L = 16, \beta = 6$. The normalized mixture and its translation symmetry follow from (166); the spatial measurements above use the fixed source.

F.5 Two-dimensional production results

F.5.1 Production protocol

The production scan uses normalized structured finite-volume priors and a longer NHMC path than the earlier calculation with a Haar reference in Appendix F.8. Both target the same periodic Wilson theories. The full target action and recorded path weight include the global closure constraint omitted from the independent source factors.

For compact $U(1)$, write the $V-1$ independent plaquette angles as $a_1, \ldots, a_{V-1}$ and impose periodic closure through $a_V = \mathrm{wrap}(-\sum_{i=1}^{V-1} a_i)$. The normalized source law is

$$q_0(a) = \prod_{i=1}^{V-1} \frac{e^{\beta \cos a_i}}{I_0(\beta)}, \tag{175}$$

while the target contains all V plaquettes. The learned path has separate forward and reverse momentum means, a learned momentum width, a position-dependent residual kick, and a learned positive drift scale. The production path uses 64 fresh-momentum stages with two leapfrog substeps per stage. Each production path includes a global flux move after the continuous transport, with its forward and reverse probabilities in the same path ratio. The autocorrelation study below repeats the frozen 64-stage proposal with this move disabled, providing a direct no-flux control.

For $\mathrm{SU}(2)$ and $\mathrm{SU}(3)$, the source law uses a two-dimensional tree/plaquette representation with the two torus holonomies retained. With $N_c = 2$ or 3, the $V-1$ independent plaquette variables have normalized source law

$$q_0(P_1, \ldots, P_{V-1}, A, B) = \prod_{i=1}^{V-1} \frac{\exp[(\beta/N_c)\,\mathrm{ReTr}\,P_i]}{z_1^{(N_c)}(\beta)}, \tag{176}$$

where $z_1^{(N_c)}(\beta) = \int_{\mathrm{SU}(N_c)} d_H P \, \exp[(\beta/N_c)\,\mathrm{ReTr}\,P]$, and A and B are independent normalized-Haar torus holonomies whose Haar-density factors are implicit. The missing plaquette is reconstructed with the holonomy commutator $ABA^\dagger B^\dagger$ rather than set to the identity, and the full Wilson target is used in the recorded work. The learned force and momentum means are gauge equivariant, while the momentum widths are gauge invariant. The production path again uses 64 fresh-momentum stages with two deterministic substeps per stage. Its learned positive drift rate depends only on stage time, β, and fixed lattice metadata, and is state-independent. No topology observable analogous to the $U(1)$ first-Chern geometric charge is assigned to these two-dimensional periodic $\mathrm{SU}(N_c)$ scans.

The structured production calculation is formulated on these reduced plaquette/holonomy coordinates. Gauge fixing removes only a target-independent Haar gauge-volume factor, and gauge-invariant observables are evaluated after lifting the reduced variables back to periodic links. The production estimator is path-SNIS on this reduced finite-volume target.

Table 18: Systematic 2D gauge NHMC protocol. The production study differs from the earlier Haar-base pilot: it uses normalized structured finite-volume priors and longer recorded paths.

Group	prior/base	learned path	production path	discrete/global move	evaluation
$U(1)$	punctured finite-volume plaquette prior	separate forward/reverse momentum means, learned width, residual kick, learned positive drift	64 fresh-momentum stages, 2 leapfrog substeps/stage	recorded post-transport global flux move	2×8192 independent paths/target; six Wilson loops, Q, Q^2, χ_t
$\mathrm{SU}(2)$	finite-volume tree/plaquette prior with torus holonomies	separate forward/reverse momentum means, learned width, residual gauge-equivariant kick, learned drift	64 stages, 2 substeps/stage	none in headline production	2×1024 paths/target; six Wilson loops
$\mathrm{SU}(3)$	finite-volume tree/plaquette prior with torus holonomies	same NHMC components in a gauge-equivariant $\mathrm{SU}(3)$ parameterization	64 stages, 2 substeps/stage	none in headline production	2×1024 paths/target; six Wilson loops

The finite-volume references are computed by converged character expansions. For $U(1)$ this reduces to integer Bessel-function sums. For $\mathrm{SU}(2)$ and $\mathrm{SU}(3)$ the references use finite-torus representation sums and Wilson-loop insertion coefficients, following the standard two-dimensional character-expansion construction (Aroca & Kubyshin, 2000).

Table 19: Earlier production component comparison: median path-weight ESS/N over 56 targets per group. These production pools and component configurations differ from the direct-neural comparison in Table 14; their summary values are not substituted into that table.

Group	Physical force	Learned residual	Learned kick only
SU(2)	0.514	0.738	0.749
SU(3)	0.630	0.819	0.797

F.5.2 Wilson-loop estimates and statistical uncertainty

Table 20: Per-volume summary of the systematic 2D NHMC gauge scan. "Resolved" counts Wilson-loop reference signals larger than three times the reported NHMC standard error; it is a precision diagnostic distinct from statistical consistency.

Group	L	min. ESS	median ESS/N	max Wilson $\lvert z\rvert$	resolved Wilson	χ_t max $\lvert z\rvert$
$U(1)$	4	4406.7	0.276	1.541	83/84	2.732
$U(1)$	8	2755.6	0.291	2.764	83/84	1.680
$U(1)$	16	4264.9	0.507	2.388	83/84	1.409
$U(1)$	32	5117.2	0.563	1.507	84/84	3.015
SU(2)	4	1238.4	0.742	2.438	79/84	–
SU(2)	8	1115.0	0.729	2.535	78/84	–
SU(2)	16	1141.7	0.732	2.183	79/84	–
SU(2)	32	1100.9	0.734	2.135	79/84	–
SU(3)	4	1172.3	0.801	3.151	66/84	–
SU(3)	8	1176.1	0.818	2.205	66/84	–
SU(3)	16	1210.3	0.827	2.687	67/84	–
SU(3)	32	1243.6	0.820	3.015	67/84	–

Across the fixed production grid there are 1008 Wilson-loop comparisons. The reported three-standard-error intervals contain the finite-volume reference in 1006 cases. The two Wilson exceptions are SU(3) at $(L, \beta) = (4, 1)$ for W_{13} and at $(32, 6)$ for W_{33}; the $U(1)$ topological susceptibility has one analogous exception at $(32, 3)$. The table also lists signal resolution, since several small-β, large-area non-Abelian loops lie below three times the NHMC standard error.

Partition functions. The same production weights give $\widehat{Z} = N^{-1} \sum_i e^{-W_i}$. We compare this estimate with the finite-torus partition function in the reduced action convention of Section F.3. Table 21 summarizes the estimates for all 168 targets. The standardized residual is $z_Z = (\widehat{Z}/Z - 1)/\mathrm{SE}(\widehat{Z}/Z)$; all values lie between -3 and 3. The corresponding free-energy estimate is $\widehat{F} = -\log \widehat{Z}$.

Table 21: Normalizer estimates from the same two-dimensional production path pools. Each group has 56 targets. The final column is the largest standardized deviation of $\widehat{Z}/Z$ from one; free-energy errors use the dimensionless quantity $F = -\log Z$.

Group	Targets	$\max \lvert\widehat{F} - F\rvert$	$\max \lvert z_Z\rvert$
$U(1)$	56	0.0305	2.481
SU(2)	56	0.0367	2.186
SU(3)	56	0.0376	2.467

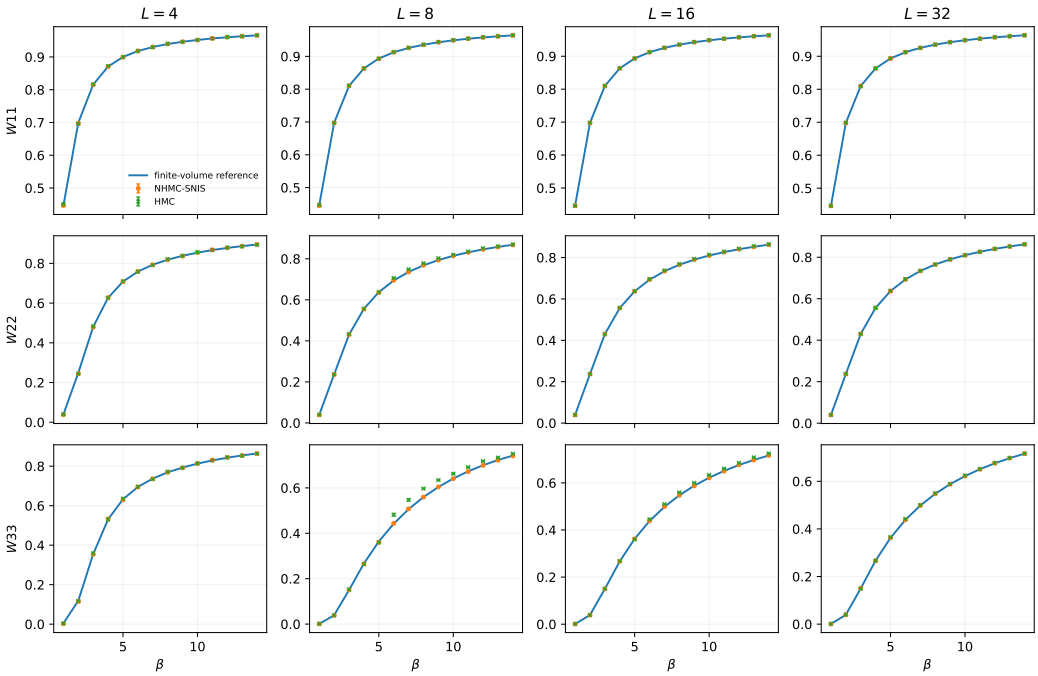

Figure 22: Complete two-dimensional compact $U(1)$ Wilson-loop scan. Rows show W_{11}, W_{22}, and W_{33} and columns show $L = 4, 8, 16, 32$. Curves are the finite-volume character-expansion references, circles are NHMC path-SNIS estimates, and crosses are independent finite HMC measurements.

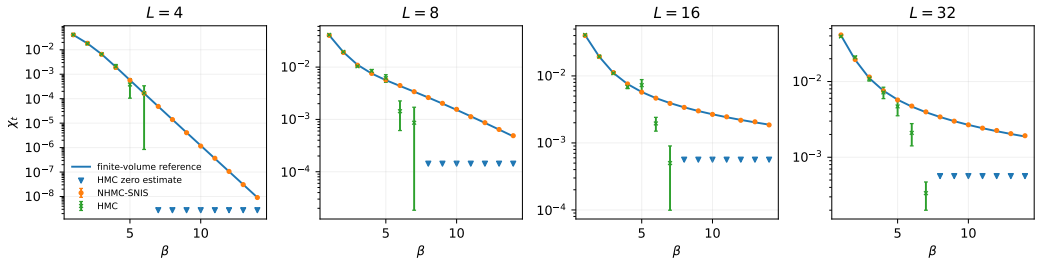

Figure 23: Complete compact $U(1)$ topological-susceptibility scan. Curves are the finite-volume references, circles are NHMC path-SNIS estimates, crosses are nonzero HMC estimates, and downward triangles mark finite HMC histories with empirical $\chi_t = 0$ after no nonzero topological charge was observed.

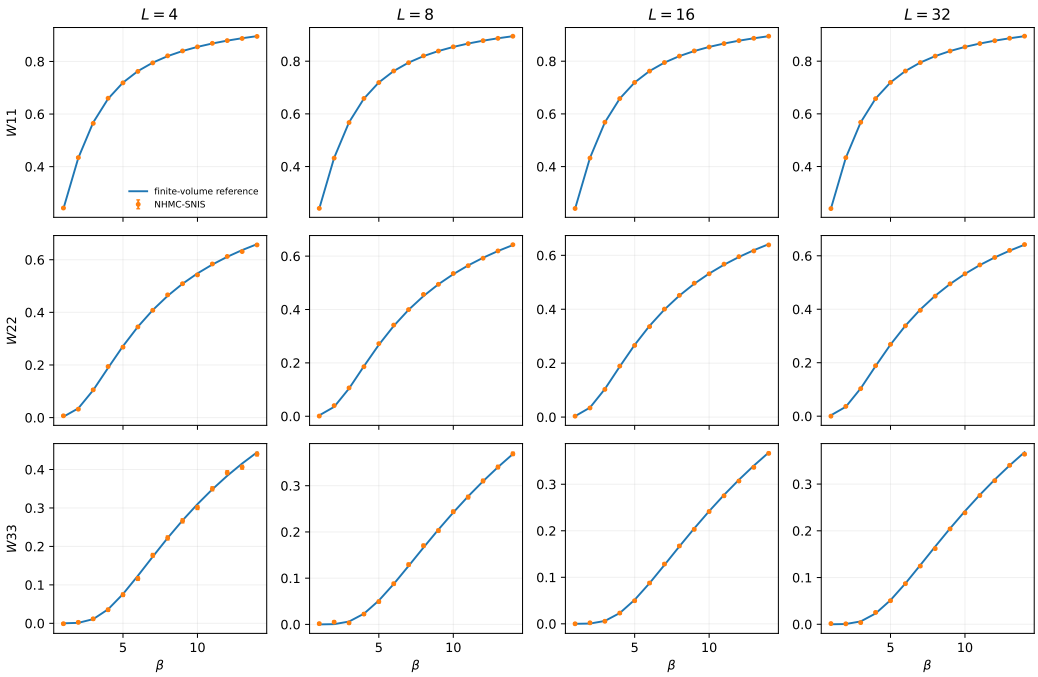

Figure 24: Systematic two-dimensional $\mathrm{SU}(2)$ Wilson-loop scan. Rows show W_{11}, W_{22}, and W_{33} and columns show the four volumes. Curves are finite-volume character-expansion references; points are NHMC path-SNIS production estimates.

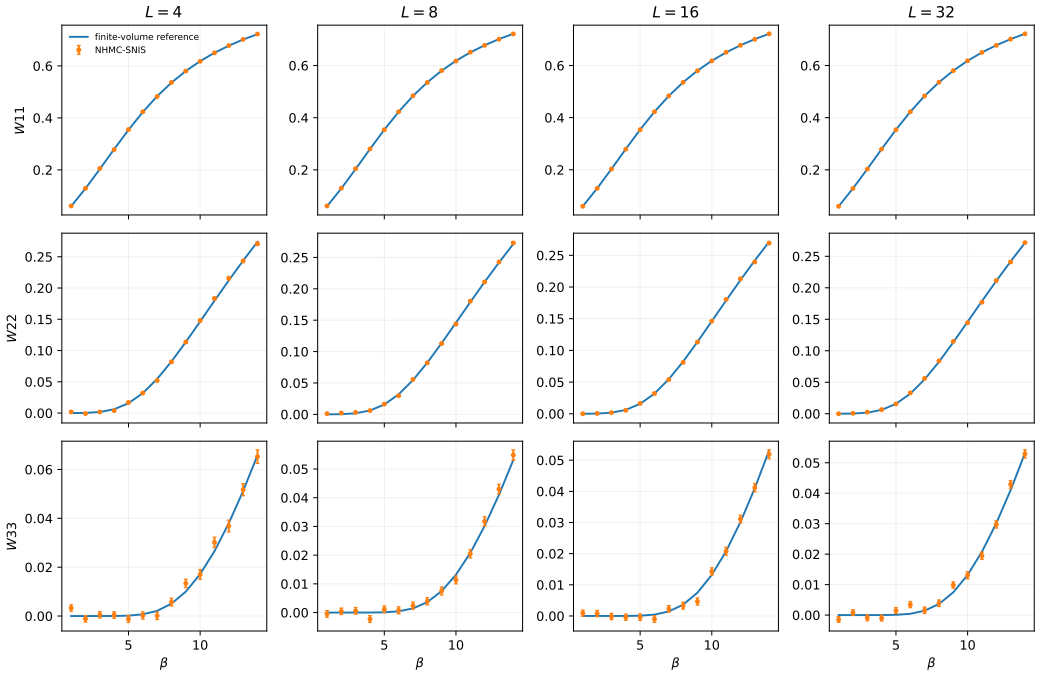

Figure 25: Systematic two-dimensional $\mathrm{SU}(3)$ Wilson-loop scan, with the same organization as Fig. 24. Very small weak-coupling loop signals are retained even when they are below the finite production sample's relative-resolution threshold.

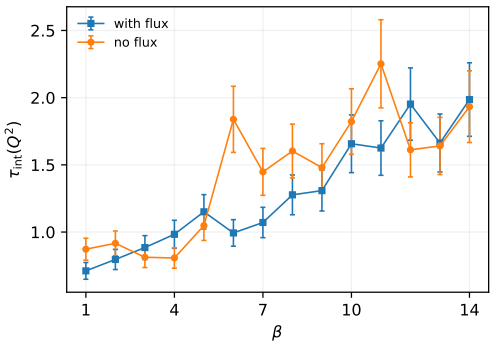

Figure 26: $L = 32$ path-IMH autocorrelation with and without the optional flux proposal, using the same residual checkpoint and 64-stage path.

F.5.3 Topological autocorrelation and flux control

To isolate the role of the optional flux proposal, the control uses the same residual checkpoint and 64-stage, two-substep path with that move removed. Each control path-IMH chain below has 3000 saved states. The HMC comparison uses one continuous chain at each coupling; its length is increased from 20000 states at small β to as many as 250000 states where the slow tail first becomes visible. A constant Q^2 trace is reported as frozen rather than assigned $\tau_{\mathrm{int}} = 1/2$.

Table 22: Single-chain compact $U(1)$ topology diagnostics at $L = 32$. A dash denotes a frozen HMC trace with no sector changes. Values marked † have an unresolved long autocorrelation tail.

β	N_{HMC}	$\tau_{Q^2}^{\mathrm{HMC}}$	changes	$\tau_{Q^2}^{\mathrm{NHMC}}$	changes
1	20000	0.69(0.03)	15994	0.87(0.08)	2340
2	20000	1.68(0.09)	14606	0.92(0.09)	2178
3	20000	6.38(0.64)	11792	0.81(0.08)	1994
4	100000	34.1(3.5)	20650	0.81(0.08)	1820
5	100000	143(26)	3456	1.05(0.11)	1636
6	250000	731(173)	1233	1.84(0.25)	1542
7	150000	12419(5173)†	113	1.45(0.17)	1444
8	100000	3839(1698)†	6	1.60(0.20)	1388
9	50000	–	0	1.48(0.18)	1330
10	50000	–	0	1.82(0.24)	1165
11	50000	–	0	2.25(0.33)	1070
12	50000	–	0	1.61(0.20)	1137
13	50000	–	0	1.64(0.21)	999
14	100000	–	0	1.93(0.27)	1022

The two choices give comparable large-volume autocorrelation in these runs. At $L = 32$, the median ratio of the control value to the value with the flux move over $\beta = 1, \ldots, 14$ is 1.12.

F.5.4 Fixed-physical-area topology scaling

In two-dimensional compact $U(1)$, $\beta \propto a^{-2}$ at fixed continuum gauge coupling. Since the physical area is L^2a^2, keeping L^2/β fixed separates the approach to smaller lattice spacing from a change of physical area. We use $L^2/\beta = 80$ for the six points below, matching the winding-HMC pure-gauge comparison (Albandea et al., 2021). With $a/a_0 = \sqrt{\beta_0/\beta}$ and $\beta_0 = 5$, weighted least squares in $\log \tau$ gives

$$\tau_{\mathrm{int}}(Q^2) = A\beta^b, \qquad b = 0.90(18), \qquad z_{\mathrm{eff}} = 2b = 1.80(35), \tag{177}$$

with $\chi^2/\mathrm{dof} = 4.61/4$. The exponent describes the tested range rather than an asymptotic continuum limit. Albandea et al. report $b = 0.565(53)$ for wHMC in the same β parameterization.

Table 23: No-flux path-IMH scan along $L^2/\beta = 80$. Each chain contains 3000 saved states.

L	β	L^2/β	$\tau_{\mathrm{int}}(Q^2)$	sector changes
20	5.00	80.00	0.97(0.10)	1608
22	6.05	80.00	1.13(0.13)	1476
24	7.20	80.00	1.58(0.20)	1356
26	8.45	80.00	2.01(0.28)	1253
28	9.80	80.00	1.71(0.22)	1147
30	11.25	80.00	1.91(0.25)	1120

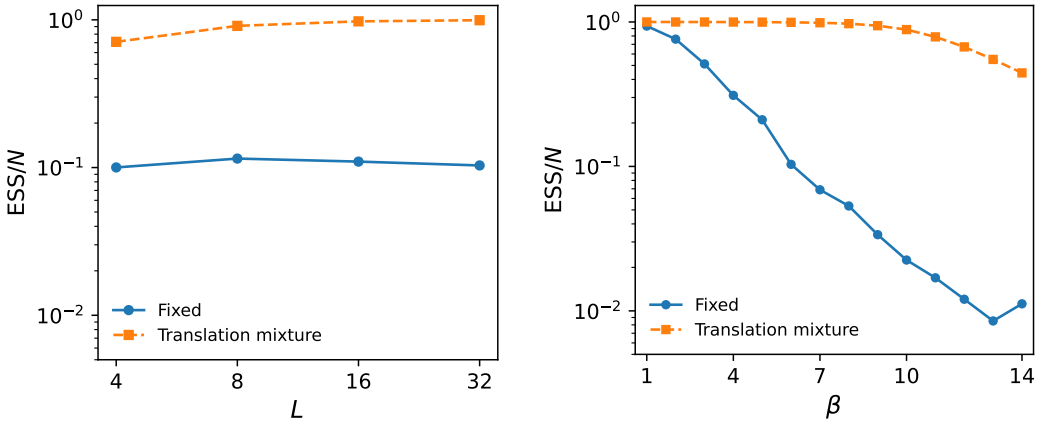

Figure 27: Direct importance-sampling ESS for the fixed structured reference and its translation mixture in SU(3). Left: volume dependence at $\beta = 6$. Right: coupling dependence at $L = 32$. Neither panel uses learned transport. The full comparison including Haar is shown below.

F.5.5 Comparison with the earlier $U(1)$ pilot

The earlier 8×8 round-trip calculation used a direct gauge-field base, a 20-stage path with five leapfrog steps per stage, and a chain-based round-trip estimator. The systematic calculation uses the structured prior, a 64-stage production path, and independent path-SNIS. Their overlapping $\beta = 4, 8$ observables provide a target-level comparison.

Table 24: Wilson-loop estimates from the earlier 8×8 round-trip pilot with a Haar reference and the systematic NHMC production calculation with a structured reference. Both target the same finite-volume $U(1)$ Wilson distribution using different algorithms, path lengths, and estimators.

β	observable	earlier RT pilot	systematic NHMC-SNIS	exact
4	W_{11}	0.8618 ± 0.0022	0.863157 ± 0.000262	0.863530
4	W_{22}	0.5545 ± 0.0107	0.555199 ± 0.000957	0.556099
4	W_{33}	0.2599 ± 0.0182	0.265567 ± 0.001582	0.267241
8	W_{11}	0.9359 ± 0.0010	0.935456 ± 0.000172	0.935477
8	W_{22}	0.7649 ± 0.0042	0.768482 ± 0.000839	0.768089
8	W_{33}	0.5568 ± 0.0106	0.559802 ± 0.001929	0.559419

The two protocols agree with the same finite-volume target within their reported uncertainties; the structured-prior scan supplies the production results used in the main text.

F.6 Compact $U(1)$ Round-Trip Pilot

We applied the round-trip construction to the trained compact $U(1)$ proposal on an 8×8 lattice. Each chain is initialized by an independent draw from the training base law followed by one forward

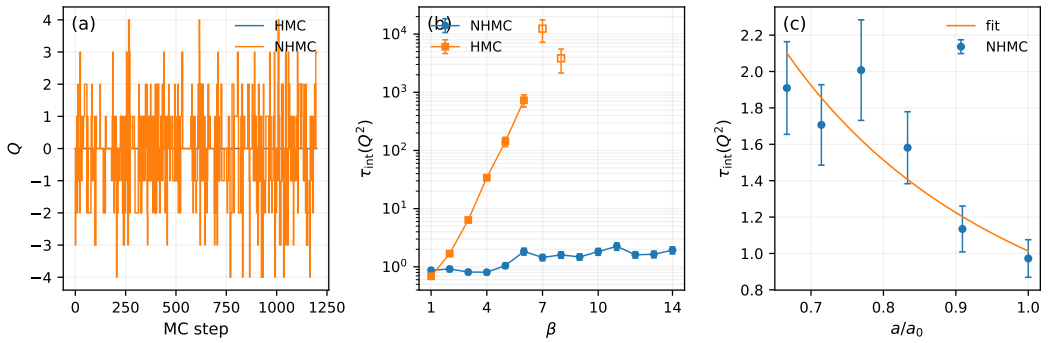

Figure 28: Compact-$U(1)$ charge histories and autocorrelations. Left: charge histories at $L = 32$, $\beta = 14$. Center: single-chain autocorrelation estimates versus coupling; open HMC markers have unresolved tails. Right: an exploratory fixed-area fit over six points. Axes count saved transitions, not equal wall-clock time. The NHMC control shown here disables the optional flux proposal.

NHMC path; no HMC or target-state warm start is used. The inverse-map and momentum-reversal residuals are below 4×10^{-15}, and the direct augmented-density ratio agrees with the work-difference expression to 2.3×10^{-13}.

Table 25: Compact $U(1)$ round-trip NHMC-MH pilot on an 8×8 lattice. Each chain starts from an independent draw from the training base law followed by one forward NHMC path. The $\beta = 4$ row uses 2000 transitions per chain and discards 400; the $\beta = 8$ row uses 5000 transitions per chain and discards 1000. Uncertainties are standard errors across chains. The final column counts rounded-sector changes over all post-discard traces.

β	Acceptance	Plaquette	W_{22}	W_{33}	Max. ratio error	Sector changes
4	1.05%	0.8618 ± 0.0022	0.5545 ± 0.0107	0.2599 ± 0.0182	2.3×10^{-13}	172
8	0.43%	0.9359 ± 0.0010	0.7649 ± 0.0042	0.5568 ± 0.0106	2.3×10^{-13}	83

At both couplings, the plaquette, W_{22}, and W_{33} estimates are within one standard error of their finite-volume character-expansion values: respectively $(0.86353, 0.55610, 0.26724)$ at $\beta = 4$ and $(0.93548, 0.76809, 0.55942)$ at $\beta = 8$. Acceptance is nevertheless only 1.05% and 0.43%. The post-discard traces contain 172 and 83 adjacent rounded-sector changes; at $\beta = 8$, 14 of 16 chains change sector at least once. Low acceptance and few accepted burn-in moves limit equilibration.

F.7 Learned proposals in two-dimensional SU(2)

The same recorded-work identities apply to a learned-force Hamiltonian proposal on two-dimensional $\mathrm{SU}(2)$ with product Haar reference measure and a normalized Haar source distribution. Link variables are unit quaternions, and the Wilson action is $S = \beta V(1 - P)$ with P the mean plaquette. Because the source density is one with respect to normalized product Haar measure, $\log q_0$ is constant and cancels in the path-SNIS weights and in the round-trip ratio, so the recorded work is

$$W = S + \sum_t \left(\log q_t^F - \log q_t^R\right). \tag{178}$$

The two-dimensional plaquette has the closed form $\langle P \rangle = I_2(\beta)/I_1(\beta)$. At $\beta = 2$ this equals 0.43313. Table 26 reports a 6×6 plaquette comparison against that value. The $T = 10$ raw proposal is 9.92% low. Path-SNIS with $N = 8192$ paths moves the estimate to 0.4341 ± 0.0029 (0.21% relative error) with Kish ESS 6.82%. The round-trip kernel, initialized by a short physics HMC run from Haar rather than by a single forward NHMC path as in the $U(1)$ pilot, accepts 22.6% of proposals and gives 0.4360 ± 0.0041. Inverse-map and work-ratio residuals remain below 10^{-12}. A longer $T = 20$ learned path gives a closer SNIS point estimate but a lower ESS (3.21%).

Table 26: Two-dimensional $\mathrm{SU}(2)$ plaquette estimates with learned NHMC on a 6×6 lattice at $\beta = 2$. The analytic plaquette reference is $I_2(\beta)/I_1(\beta) = 0.43313$. Independent estimates use $N = 8192$ forward paths initialized from Haar. Round-trip NHMC-MH uses eight chains of 300 transitions with 60 discarded; uncertainties are standard errors across chain means.

Kernel	Estimator	Plaquette	Rel. error	ESS$/N$ or acc.	N_{eff}
Reference	I_2/I_1	0.43313	–	–	–
Learned T=10	Raw	0.39018 ± 0.00085	9.92%	–	–
Learned T=10	Path-SNIS	0.4341 ± 0.0029	0.21%	6.82%	559
Learned T=10	Round-trip MH	0.4360 ± 0.0041	0.67%	22.6% acc.	–
Learned T=20	Raw	0.38718 ± 0.00090	10.61%	–	–
Learned T=20	Path-SNIS	0.4333 ± 0.0028	0.03%	3.21%	263
Learned T=20	Round-trip MH	0.4341 ± 0.0053	0.22%	14.5% acc.	–

F.8 Volume dependence with a Haar reference for $U(1)$, $\mathrm{SU}(2)$ and $\mathrm{SU}(3)$

The same Wilson convention is used on all three groups,

$$S[U] = \beta V\big(1 - P[U]\big), \tag{179}$$

with P the mean plaquette ($P = \cos\theta_p$ for $U(1)$, $P = \mathrm{Re}\,\mathrm{Tr}\,U_p/2$ for $\mathrm{SU}(2)$, and $P = \mathrm{Re}\,\mathrm{Tr}\,U_p/3$ for $\mathrm{SU}(3)$). The integration reference measure is product Haar measure and the source distribution is normalized Haar, so $\log q_0$ is constant and the recorded work reduces to

$$W = S + \sum_t \big(\log q_t^F - \log q_t^R\big). \tag{180}$$

Leapfrog uses a learned, momentum-independent force; the Wilson action enters the training objective through W, not as the integrator force. The analytic single-plaquette references used at the displayed precision are $\langle P\rangle = I_1(\beta)/I_0(\beta) = 0.69777$ for $U(1)$ at $\beta = 2$, $\langle P\rangle = I_2(\beta)/I_1(\beta) = 0.43313$ for $\mathrm{SU}(2)$ at $\beta = 2$, and, from the single-plaquette Weyl integral, $\langle P\rangle = 0.35395$ for $\mathrm{SU}(3)$ at $\beta = 5$ and 0.42253 at $\beta = 6$. All NHMC estimators below start from Haar followed by one forward path; unlike Table 26, there is no HMC warm start.

Table 27 collects the raw, path-SNIS, path-IMH, and round-trip results for all four scans. Table 28 gives the corresponding HMC and heat-bath plaquettes. Path-SNIS moves the raw proposal toward the reference value, while the Kish ESS and the path-IMH/round-trip acceptances fall with volume. At $L = 4$, increasing the $\mathrm{SU}(3)$ coupling from $\beta = 5$ to 6 reduces ESS from 5.23% to 0.45%, comparable to the $\beta = 5$ loss between $L = 4$ and $L = 6$.

Table 27: Learned-NHMC plaquette estimates from a Haar reference. Path-SNIS uses 8192 independent paths; the second line in each corrected entry gives ESS/N or acceptance. The analytic references are 0.69777 for $U(1)$ at $\beta = 2$, 0.43313 for SU(2) at $\beta = 2$, and 0.35395 and 0.42253 for SU(3) at $\beta = 5$ and 6. †The indicated chains have $\hat{R} \geq 1.1$, indicating persistent differences between chain segments.

Target	L	Raw	Path-SNIS	Path-IMH	Round-trip
			0.70270 ± 0.00396	0.7031 ± 0.0042	0.70013 ± 0.00161
$U(1), \beta = 2$	4	0.64144 ± 0.00124	ESS 4.10%	acc. 18.2%	acc. 24.1%
			0.69244 ± 0.00723	0.6992 ± 0.0038	0.69569 ± 0.00235
	6	0.65103 ± 0.00082	ESS 1.76%	acc. 10.7%	acc. 11.5%
			0.68943 ± 0.00612	$0.6933 \pm 0.0083^{\dagger}$	0.69678 ± 0.00221
	8	0.64416 ± 0.00061	ESS 0.62%	acc. 4.03%	acc. 3.52%
			0.42975 ± 0.00183	0.43170 ± 0.00128	0.43175 ± 0.00297
SU(2), $\beta = 2$	4	0.39848 ± 0.00117	ESS 27.6%	acc. 39.4%	acc. 42.1%
			0.43134 ± 0.00274	0.43159 ± 0.00267	0.43658 ± 0.00149
	6	0.39075 ± 0.00084	ESS 5.36%	acc. 21.2%	acc. 22.2%
			0.41519 ± 0.00391	0.42425 ± 0.00346	0.42838 ± 0.00330
	8	0.35668 ± 0.00063	ESS 0.57%	acc. 4.03%	acc. 3.41%
			0.35277 ± 0.00350	0.3623 ± 0.0106	0.35721 ± 0.00169
SU(3), $\beta = 5$	4	0.32557 ± 0.00080	ESS 5.23%	acc. 21.1%	acc. 25.0%
			0.34694 ± 0.00455	$0.3605 \pm 0.0099^{\dagger}$	0.35284 ± 0.00462
	6	0.30565 ± 0.00054	ESS 0.41%	acc. 3.49%	acc. 4.03%
			0.35959 ± 0.00760	0.3457 ± 0.0036	0.35047 ± 0.00333
	8	0.30842 ± 0.00045	ESS 0.09%	acc. 2.13%	acc. 2.17%
			0.41679 ± 0.00336	0.4226 ± 0.0041	0.42152 ± 0.00231
SU(3), $\beta = 6$	4	0.37546 ± 0.00075	ESS 0.45%	acc. 10.5%	acc. 11.8%
			0.42092 ± 0.00479	$0.4367 \pm 0.0086^{\dagger}$	0.42201 ± 0.00324
	6	0.37898 ± 0.00057	ESS 0.44%	acc. 3.64%	acc. 4.97%
			0.41563 ± 0.00334	$0.4151 \pm 0.0069^{\dagger}$	$0.4319 \pm 0.0095^{\dagger}$
	8	0.37055 ± 0.00049	ESS 0.16%	acc. 1.49%	acc. 0.57%

Table 28: Local-reference plaquette estimates for the Haar-base scans in Table 27. The $U(1)$ HMC values are from a separate longer run (5000 discarded and 20000 production transitions); the remaining HMC and heat-bath values are from the shared evaluation.

Target	L	Analytic reference	HMC	Heat bath
$U(1), \beta = 2$	4	0.69777	0.69878 ± 0.00049	0.69975 ± 0.00043
	6	0.69777	0.69746 ± 0.00019	0.69856 ± 0.00030
	8	0.69777	0.69738 ± 0.00027	0.69760 ± 0.00021
SU(2), $\beta = 2$	4	0.43313	0.43210 ± 0.00123	0.43324 ± 0.00049
	6	0.43313	0.43182 ± 0.00033	0.43323 ± 0.00031
	8	0.43313	0.43388 ± 0.00033	0.43282 ± 0.00013
SU(3), $\beta = 5$	4	0.35395	0.35374 ± 0.00074	0.35399 ± 0.00021
	6	0.35395	0.35452 ± 0.00020	0.35385 ± 0.00013
	8	0.35395	0.35391 ± 0.00036	0.35398 ± 0.00021
SU(3), $\beta = 6$	4	0.42253	0.42308 ± 0.00048	0.42262 ± 0.00021
	6	0.42253	0.42236 ± 0.00023	0.42262 ± 0.00015
	8	0.42253	0.42232 ± 0.00023	0.42256 ± 0.00013

Table 29 gives plaquette integrated autocorrelation times and the between-chain convergence statistic $\hat{R}$. On the eight points where both learned chains satisfy $\hat{R} < 1.1$, round-trip NHMC-MH decorrelates faster than independent path-IMH without exception, by factors of 1.7 to 25.9. Because one round-trip transition evaluates two NHMC paths against one for path-IMH, the ratio must exceed 2 for an improvement per evaluated path. This holds except for SU(2) at $L = 4$.

At three further points, path-IMH has $\hat{R} \geq 1.1$ while round-trip NHMC-MH has $\hat{R} < 1.1$. Both methods fail this criterion for SU(3) at $\beta = 6$ and $L = 8$, where round-trip acceptance is 0.57%, the lowest in this scan. Autocorrelation ratios are omitted for these four points. The local reference samplers have τ_{int} near unity at every volume.

Table 29: Integrated autocorrelation time of the plaquette, in units of one saved transition, for the two learned chains of Tables 27 and 28, with the HMC ($n_{\mathrm{lf}} = 10$, $\varepsilon = 0.08$) and heat-bath references. Estimates use the Sokal automatic window. Entries marked † have $\hat{R} \geq 1.1$; their autocorrelation estimates may miss fluctuations between regions that the chains have not explored. Ratios are omitted for these entries. The $U(1)$ HMC column is the longer re-run described in Table 28; the other autocorrelation estimates use the corresponding plaquette ensembles. At the eight points where both learned chains have $\hat{R} < 1.1$, the ratio of path-IMH to round-trip autocorrelation times ranges from 1.7 to 25.9. One round-trip transition evaluates two NHMC paths against one for path-IMH, so an improvement per evaluated path requires a ratio above 2. This holds except for $\mathrm{SU}(2)$ at $L = 4$. Both learned chains stay slower per transition than the local references, whose τ_{int} is near unity on every volume.

Group	β	L	Path-IMH	Round-trip	$\tau^{\mathrm{IMH}}/\tau^{\mathrm{RT}}$	HMC	Heat bath
$U(1)$	2	4	10.1	3.6	2.8	1.60	0.63
$U(1)$	2	6	29.0	8.2	3.5	1.45	0.60
$U(1)$	2	8	112.3†	26.2	—	1.46	0.59
$\mathrm{SU}(2)$	2	4	3.5	2.1	1.7	0.93	0.50
$\mathrm{SU}(2)$	2	6	9.1	4.4	2.1	0.96	0.52
$\mathrm{SU}(2)$	2	8	93.6	43.9	2.1	0.98	0.51
$\mathrm{SU}(3)$	5	4	80.4	3.1	25.9	0.90	0.50
$\mathrm{SU}(3)$	5	6	167.6†	36.1	—	0.93	0.50
$\mathrm{SU}(3)$	5	8	98.7	41.5	2.4	0.96	0.53
$\mathrm{SU}(3)$	6	4	43.4	8.3	5.2	0.88	0.52
$\mathrm{SU}(3)$	6	6	152.7†	23.9	—	0.93	0.53
$\mathrm{SU}(3)$	6	8	190.7†	176.3†	—	1.00	0.50

Gauge-equivariant architecture. The force and the forward/backward momentum means are gauge-equivariant drift networks. Each call first builds shared plaquettes and staples. From those it extracts gauge-invariant scalars—seven for $\mathrm{SU}(2)$ ($\mathrm{Re}\,\mathrm{Tr}(P)/2$, $|\mathrm{Im}(P)|^2$, $\mathrm{Re}\,\mathrm{Tr}(P^2)/2$, and per-direction $\mathrm{Re}\,\mathrm{Tr}(US)/2$, $|\mathrm{Im}(US)|^2$) and eleven for $\mathrm{SU}(3)$ (real and imaginary traces of P and P^2, $|\mathrm{Tr}(P)|^2$, and the analogous US traces). A three-block cyclic CNN with channels $48/48$, 3×3 kernels, circular padding, GroupNorm, SiLU, and a width-128 sinusoidal time embedding maps those scalars to four coefficients per link. The Lie-algebra output is the coefficient-weighted sum of four covariant bases. For $\mathrm{SU}(3)$ the Wilson-force matrix is first formed through its traceless anti-Hermitian projection; when expressed in the Hermitian momentum convention of Section 2, it is multiplied by $-i$. The other matrix primitives are projected to the corresponding traceless Hermitian tangent before a momentum kick. Under a gauge transformation $U_{x,\mu} \mapsto g(x) U_{x,\mu} g(x+\hat{\mu})^\dagger$ the output transforms in the adjoint at x. Forward and reverse momentum scales are separate invariant CNNs (two convolution blocks) with the same scalar features. The learned kernel therefore contains three equivariant drift nets (force, μ^F, μ^R) and two invariant scale nets, together 243,169 parameters for $\mathrm{SU}(2)$ and 251,809 for $\mathrm{SU}(3)$. Table 30 summarizes the layout. All lattice tensors are float64; $\mathrm{SU}(3)$ links are stored as complex 3×3 matrices.

Table 30: Gauge-equivariant learned-force kernel used for the Haar $\mathrm{SU}(2)$ and $\mathrm{SU}(3)$ runs. The same time embedding is shared by the force net and the forward/backward momentum nets. Parameter counts are from the trained checkpoints.

	$\mathrm{SU}(2)$	$\mathrm{SU}(3)$
Links	unit quaternions	$\mathrm{SU}(3)$ matrices
Lie-algebra force	$\mathfrak{su}(2) \simeq \mathbb{R}^3$	$\mathfrak{su}(3) \simeq \mathbb{R}^8$
Site invariants	3	5
Link invariants	2 per direction	3 per direction
Total scalar features	7	11
Covariant bases / link	4	4
CNN	$3\times$ cyclic 3×3, channels $48/48$	same
Time embedding	sinusoidal, width 128	same
Networks	force; fwd/bwd μ; fwd/bwd σ	same
Parameters	243,169	251,809
Path stages T / leapfrog per stage	10 / 5	10 / 5
Batch / epochs	16 / 8000	8 / 8000
Objective	$\mathbb{E}[W] + 0.01\,\mathrm{Var}(W)$	same

Training cost of equivariance. Each path has $T = 10$ stages and $n_{\mathrm{lf}} = 5$ leapfrog steps per stage. The force network is evaluated $n_{\mathrm{lf}} + 1 = 6$ times per stage, and each evaluation recomputes plaquettes, staples, and the four covariant bases before the CNN. The momentum nets are called once per stage at the endpoints. Training differentiates through that whole path in float64. Table 31 gives the training wall-clock cost: about 1.7–4.2 seconds per epoch for $\mathrm{SU}(2)$ (batch 16) and 2.6–6.6 seconds per epoch for $\mathrm{SU}(3)$ (batch 8). The networks have 0.24–0.25M parameters, but each leapfrog step rebuilds staples and four covariant bases. A non-equivariant CNN of the same width and leapfrog depth was not trained, so the relative cost of enforcing equivariance remains unmeasured.

Table 31: Observed training cost of the equivariant Haar kernels, 8000 epochs on one GPU. Seconds per epoch are wall time divided by 8000. Samples/s is training configurations per second. The $\mathrm{SU}(2)$ 6×6 job is an earlier run on a different device, so the L trend is not a FLOP measurement. A matched non-equivariant CNN was not trained.

Group	L	Batch	Wall (h)	s/epoch	Samples/s	Peak mem (MB)
$\mathrm{SU}(2)$	4	16	3.81	1.72	9.32	223
$\mathrm{SU}(2)$	6	16	9.41	4.24	3.78	327
$\mathrm{SU}(2)$	8	16	4.13	1.86	8.61	579
$\mathrm{SU}(3)$	4	8	5.71	2.57	3.11	148
$\mathrm{SU}(3)$	6	8	14.7	6.62	1.21	284

G Relative Path Densities and Gauge Covariance

This appendix collects the structural facts used by the gauge-field parameterization: relative path densities, gauge covariance, Gaussian momentum laws, and non-uniqueness of component parameterizations. Related analytic-force and gauge-equivariant constructions include Levy et al. (2018); Zhang & Chen (2022); Boyda et al. (2021).

G.1 Relative path densities

Let $(\Omega, \mathcal{F})$ denote the complete recorded path space, including the initial configuration, stochastic momentum records, and any recorded discrete variables.

Proposition 5 (Radon–Nikodym representation)**.** *Let Q_0 and Q_1 be probability measures on $(\Omega, \mathcal{F})$ with $Q_1 \ll Q_0$. Then*

$$\ell_{1|0}(\Gamma) = \log \frac{dQ_1}{dQ_0}(\Gamma) \tag{181}$$

is unique Q_0-almost everywhere and

$$dQ_1 = e^{\ell_{1|0}} dQ_0, \qquad \mathbb{E}_{Q_0}[e^{\ell_{1|0}}] = 1. \tag{182}$$

If $Q_2 \ll Q_1 \ll Q_0$, then

$$\log \frac{dQ_2}{dQ_0} = \log \frac{dQ_2}{dQ_1} + \log \frac{dQ_1}{dQ_0} \tag{183}$$

as an equality Q_2-almost everywhere on their positive-density sets. We use $\log 0 = -\infty$ in (181)*.*

Proof. The first statement is the Radon–Nikodym theorem and uniqueness of its derivative. Integrating $dQ_1 = (dQ_1/dQ_0)dQ_0$ gives $\mathbb{E}_{Q_0}[e^{\ell_{1|0}}] = Q_1(\Omega) = 1$. The last identity is the Radon–Nikodym chain rule followed by a logarithm. □

The measures must be compared on one specified measurable record space. If deterministic maps differ, their fully augmented (x, p) trajectory laws may have different supports and need not admit a density ratio. One must instead fix the maps or explicitly transport both measures to common record coordinates; changing coordinates also changes the endpoint map.

Thus a chosen reference path law can be factored out exactly. This identity does not imply that the remaining log-density ratio is small, nor that any one neural component must become simpler.

G.2 Gauge invariance of a relative density

Let $T_g : \Omega \to \Omega$ denote the action of a gauge transformation g on the complete path, including the adjoint transformation of momenta.

Proposition 6 (Gauge invariance of the density ratio)**.** *Assume $Q_1 \ll Q_0$ and, for a fixed g, $(T_g)_\# Q_0 = Q_0$ and $(T_g)_\# Q_1 = Q_1$. Then a Radon–Nikodym derivative $r = dQ_1/dQ_0$ satisfies*

$$r(T_g\Gamma) = r(\Gamma) \qquad Q_0\text{-a.e.} \tag{184}$$

and hence $\ell_{1|0} = \log r$ is invariant almost everywhere. For a compact group, under the usual measurability assumptions, an invariant version of r can be chosen by Haar averaging.

Proof. For every measurable A, invariance of Q_0 gives

$$\int_A r(T_g\Gamma)\, Q_0(d\Gamma) = \int_{T_g A} r(\Gamma)\, Q_0(d\Gamma) \tag{185}$$

$$= Q_1(T_g A) = Q_1(A) = \int_A r(\Gamma)\, Q_0(d\Gamma). \tag{186}$$

Uniqueness of the Radon–Nikodym derivative proves (184). For compact G, Haar averaging $\bar{r}(\Gamma) = \int_G r(T_g\Gamma) d\mu_G(g)$ gives an invariant representative of the same almost-everywhere class. □

This proposition concerns a scalar density ratio. A link-valued kick instead transforms covariantly. Invariance of the endpoint target alone is not enough: both path laws in the ratio must satisfy the symmetry. A maximal-tree choice removes local gauge redundancy, but non-Abelian variables retain simultaneous global conjugation. A generic matrix-input MLP is not automatically covariant under that remaining action. Equivariance must be checked for the implemented feature and field construction; path correction itself does not require it.

G.3 Gauge-covariant kick fields

For a link $\ell = (x, y)$, write $U_\ell \mapsto g_x U_\ell g_y^\dagger$ and let a link tangent field transform by the adjoint action at the source site.

Proposition 7 (Invariant coefficients and covariant fields)**.** *Suppose $I_j(U^g) = I_j(U)$ and $B_{k,\ell}(U^g) = g_x B_{k,\ell}(U) g_x^\dagger$. Then*

$$R_{\theta,\ell}(U,t) = \sum_k a_{\theta k}(I(U),t) B_{k,\ell}(U) \tag{187}$$

satisfies $R_{\theta,\ell}(U^g,t) = g_x R_{\theta,\ell}(U,t) g_x^\dagger$ for arbitrary scalar coefficient functions $a_{\theta k}$. If an analytic force $F_{0,\ell}$ is covariant, then $F_{0,\ell} + R_{\theta,\ell}$ is also covariant. At fixed U, the momentum kick

$$P'_\ell = P_\ell + \varepsilon[F_{0,\ell}(U) + R_{\theta,\ell}(U,t)] \tag{188}$$

is a translation in the Lie algebra and has unit Jacobian in the momentum coordinates.

Proof. The scalar coefficients are unchanged by a gauge transformation, while each $B_{k,\ell}$ transforms by the same adjoint action. The covariance of the sum follows immediately. The momentum update is a translation at fixed U, so its momentum–momentum Jacobian is the identity. □

The learned field need not be the gradient of a scalar potential for the path correction used here.

G.4 Gaussian momentum laws

The NHMC gauge calculations use conditional Gaussian momentum laws. Under matched deterministic dynamics, approximation of their conditional means has a direct path-space interpretation.

Proposition 8 (Conditional-mean error and path KL)**.** *Let $Q_\star$ and Q_θ have the same initial law, deterministic maps, and every stochastic conditional law other than the momentum means, including any discrete-label probabilities. At stage t, suppose their conditional momentum laws have the same positive definite covariance $\Sigma_t(H_t)$ but means $\mu_{\star,t}(H_t)$ and $\mu_{\theta,t}(H_t)$, where H_t is the history before the draw. If the expectations below are finite, then*

$$D_{\mathrm{KL}}(Q_\star \| Q_\theta) = \frac{1}{2} \sum_t \mathbb{E}_{Q_\star}\left[\|\mu_{\star,t}(H_t) - \mu_{\theta,t}(H_t)\|^2_{\Sigma_t(H_t)^{-1}} \right]. \tag{189}$$

Proof. The chain rule for relative entropy expresses the path KL as the sum of the conditional KL divergences at each stochastic draw, averaged over the preceding history under $Q_\star$. For equal-covariance Gaussian distributions,

$$D_{\mathrm{KL}}(\mathcal{N}(\mu_1,\Sigma)\|\mathcal{N}(\mu_2,\Sigma)) = \tfrac{1}{2}(\mu_1-\mu_2)^T \Sigma^{-1}(\mu_1-\mu_2).$$

Substitution gives (189). □

Equation (189) is conditional on matched deterministic maps, covariances, and all other random kernels. If any of these other kernels changes, its conditional KL contribution must also be retained. It does not make the decomposition of a general NHMC path law into momentum means, kicks, widths, and drift parameters unique.

G.5 NON-UNIQUENESS OF NHMC PARAMETERIZATIONS

A simple kick–drift–kick stage shows how a kick field and the conditional momentum laws can compensate one another.

Proposition 9 (Compensation in one KDK stage)**.** *Consider*

$$v = p + aF(x), \qquad y = \mathcal{D}_b(x, v), \qquad \bar{p} = v + aF(y), \tag{190}$$

with forward density $q_F(p \mid x)$ and reverse density $q_R(r \mid y)$ evaluated at $r = -\bar{p}$. For a position-dependent field Δ, define

$$\widehat{F} = F + \Delta, \tag{191}$$

$$\widehat{q}_F(\widehat{p} \mid x) = q_F(\widehat{p} + a\Delta(x) \mid x), \tag{192}$$

$$\widehat{q}_R(\widehat{r} \mid y) = q_R(\widehat{r} + a\Delta(y) \mid y). \tag{193}$$

Under the coupling $\widehat{p} = p - a\Delta(x)$, the two stages produce the same configuration y and the same realized forward–reverse log-density ratio.

Proof. The modified half-step momentum is $\widehat{v} = \widehat{p} + a\widehat{F}(x) = p + aF(x) = v$, so the drift gives the same y. The modified output momentum is $\widehat{\bar{p}} = \bar{p} + a\Delta(y)$ and hence $\widehat{r} = r - a\Delta(y)$. By construction, $\widehat{q}_F(\widehat{p} \mid x) = q_F(p \mid x)$ and $\widehat{q}_R(\widehat{r} \mid y) = q_R(r \mid y)$. The endpoints and the recorded momentum-density contribution to the work are therefore unchanged. □

The proposition preserves the configuration transition and recorded work under the displayed translation of the momentum record. It does not say that the laws on the original, unshifted momentum coordinates are identical. The example applies to the stated one-drift stage; interior kicks in a longer composition need not be absorbed by boundary momentum translations. It gives no general bound on neural representation cost.

H Additional Systems

We also consider the Lennard–Jones 13-particle cluster (LJ13) and four-dimensional $\mathrm{SU}(3)$ gauge fields.

H.1 Full-depth LJ13 comparison

LJ13 tests NHMC on the standard 13-particle Lennard–Jones target. Using the same 50000-path pool, path-SNIS reduces the energy-W_2 distance from 7.654 to 1.258 and the pooled pair-distance W_2 from 0.082 to 0.018. The geometric W_2 changes more modestly, from 1.809 ± 0.011 to 1.779 ± 0.015, and the path-SNIS ESS fraction is 0.63%. The result shows substantial correction of the energy and pair-distance distributions while also exposing the remaining high-dimensional overlap problem.

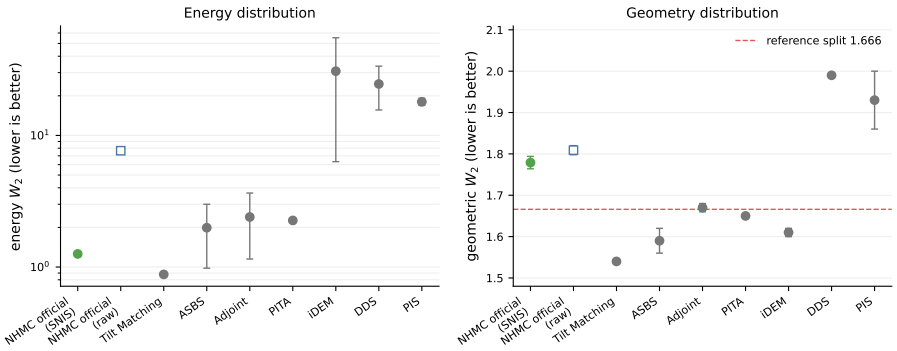

Figure 29: LJ13 distributional comparison on the standard full-depth target. Raw and path-SNIS NHMC use the same proposal pool; published baselines are shown using the aligned comparison of Potaptchik et al. (2025).

H.2 LJ13 Conditional-Momentum Ablation

The momentum-law ablation uses the historical half-depth LJ13 configuration; Section H.1 gives the full-depth benchmark. Within the ablation, the target, base distribution, path depth, optimizer, network backbone, and evaluation pool are fixed while the conditional momentum laws are changed.

Table 32: Conditional-momentum ablation on the historical half-depth LJ13 target. All rows use one training seed, 56 stages, 10 leapfrog updates per stage, and a common 100000-path evaluation. The counted force evaluations per path include dynamics and the position-dependent force features used by the active momentum networks. This diagnostic is separate from the full-depth official-potential comparison in Fig. 29.

Momentum laws	work std.	path-SNIS ESS	path-IMH acc.	raw energy-W_2	SNIS energy-W_2	force evals.
Fixed Gaussian	18.627	0.0104%	0.21%	29.667	4.916	616
Learned forward only	16.910	0.0387%	0.32%	23.321	4.092	672
Learned forward and reverse	2.050	2.5221%	21.69%	4.580	0.223	728

The learned forward-and-reverse model gives the narrowest work distribution, the largest path-SNIS ESS and path-IMH acceptance, and the smallest corrected energy-W_2 in this ablation.

H.3 Four-dimensional $\mathrm{SU}(3)$ extension

The four-dimensional periodic $\mathrm{SU}(3)$ calculations use the Wilson and tree-level Symanzik pure-gauge actions,

$$S_\beta(U) = \beta \left[c_0 \sum_P \left(1 - \frac{1}{3} \operatorname{ReTr} U_P \right) + c_1 \sum_R \left(1 - \frac{1}{3} \operatorname{ReTr} U_R \right) \right], \tag{194}$$

with $(c_0, c_1) = (1, 0)$ and $(5/3, -1/12)$, respectively, and with the usual four-dimensional pure-gauge convention $\beta = 6/g_0^2$. Link momenta are traceless Hermitian matrices. The learned force and forward/reverse momentum means transform in the adjoint representation at the source site of each link. The models include scalar force coefficients, a wider scalar network, a multi-operator covariant basis, and a rank-two positive low-rank correction to the momentum covariance. All stochastic momentum densities retain their normalization and log determinant in the recorded work. Nonequilibrium lattice-gauge calculations based on Jarzynski work, stochastic normalizing flows, and topology-oriented out-of-equilibrium transport provide the closest physics context for this extension (Caselle et al., 2016; 2018; 2022; Bulgarelli et al., 2025; Bonanno et al., 2024; 2026). These calculations measure proposal overlap on small lattices. Large-volume topology and continuum scaling have not been evaluated.

Table 33: Four-dimensional $\mathrm{SU}(3)$ calculations included in this work.

Volume	Action/couplings	Calculation	Evaluation
4^4	Wilson and tree-level Symanzik, $\beta = 1, \ldots, 14$	models shared across β; SNIS overlap and five Wilson loops; full-lattice and block round-trip chains	short paths and Markov chains
8^4	Wilson and Symanzik, $\beta = 6$	transferred multi-operator model; 16-candidate SNIS and short Markov chains	low proposal overlap
$16^4, 32^4$	$\beta = 6$	one gradient update within a conditional block, retaining neighboring boundary links	equilibrium observables not measured

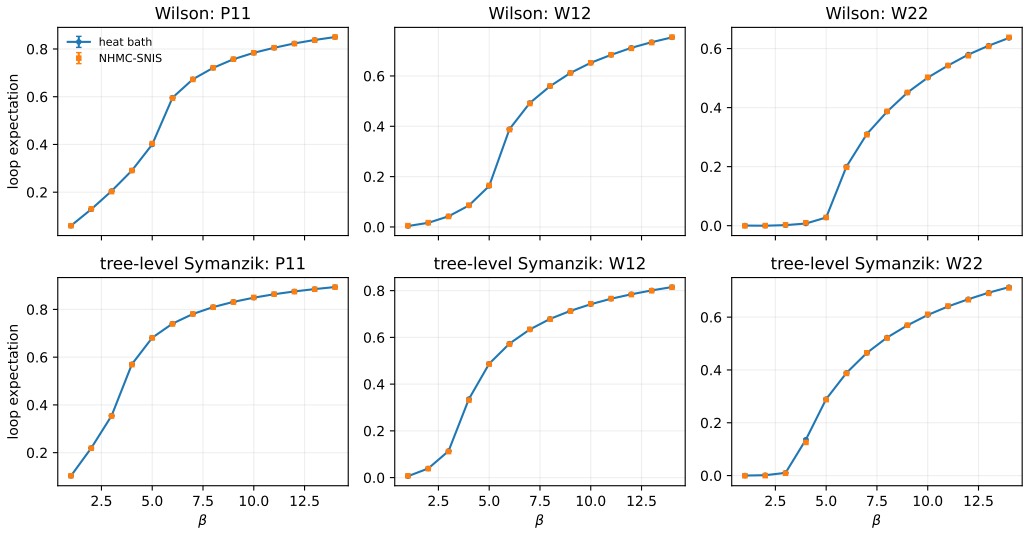

Figure 30: Four-dimensional $\mathrm{SU}(3)$ comparison on 4^4. The learned multi-operator NHMC proposal is evaluated by path-SNIS and compared with the independent heat-bath banks used by this study. Columns show the plaquette, W_{12}, and W_{22} for the Wilson (top) and tree-level Symanzik (bottom) actions over integer $\beta = 1, \ldots, 14$. Several of the largest-loop heat-bath references, estimated from short chains, have non-negligible between-chain variation.

Table 34: Four-dimensional $\mathrm{SU}(3)$ path-SNIS ESS at $L = 4$, summarized over integer $\beta = 1, \ldots, 14$. Each point uses 32 held-out candidates from the finite heat-bath banks.

NHMC proposal variant	Wilson median ESS ($N = 32$)	Symanzik median ESS ($N = 32$)
direct reweighting	26.70	28.41
untrained physical force, coarse	19.90	7.29
untrained physical force, fine	26.83	27.94
learned scalar coefficients	24.86	20.14
wider scalar network	25.78	22.96
learned multi-operator	24.69	18.00
multi-operator + low-rank covariance	24.69	18.00
sparse refresh	24.26	21.87

At $L = 4$, the learned proposals improve on the coarse untrained physical-force path in several settings and are comparable to direct reweighting or a finer physical integrator. At $L = 8$, $\beta = 6$, the low ESS and visible loop shifts show that overlap has become the dominant limitation. At $L = 16, 32$, the calculation consists of one gradient update within a conditional block, with neighboring links retained for boundary interactions. Equilibrium observables were not measured at these volumes.

Accelerated parallel tempering is used as an external comparison. The short block calculations also show that a high local acceptance rate can coexist with poor full-volume split-chain statistics, so acceptance and global mixing are reported separately.

H.4 Nearby-coupling Wilson calculation

A separate four-dimensional Wilson calculation uses $\beta_0 = 5.95$ and $\beta = 6.0$. The model is a pointwise MLP with 16 hidden units, a multi-operator covariant input, two stages, fixed step size, and 982 parameters. Each volume's heat-bath reference has four chains with 2048 warm-up sweeps and 4096 saved measurements per chain, separated by four sweeps. The source pool contains 1024 held-out fields, 256 from each chain. Their correlations enter the uncertainty estimate; weight ESS measures concentration of the weights rather than independence of the stored fields.

Table 35: Wilson-loop estimates on 4^4 at $\beta = 6.0$. The last column is the largest absolute standardized difference over five loops, combining each method's reported uncertainty with that of the heat-bath reference.

Method	Records	P_{11}	W_{22}	W_{33}	$\max \lvert z \rvert$
Heat bath	16384	.596800	.200362	.066553	—
HMC	16384	.596829	.200355	.066575	.15
Direct SNIS	1024	.596704	.200347	.066733	.33
Learned SNIS	1024	.596878	.200706	.066801	.89
Path-IMH, stored pool	768	.596080	.199610	.066950	1.34
Round-trip MH	2048	.600157	.208367	.076299	6.12

The HMC calculation has four chains, 2048 warm-up and 4096 production trajectories per chain, step size 0.02, and 25 leapfrog steps. The maximum $\widehat{R}$ is 1.005. The stored-pool path-IMH calculation uses each held-out source field once. The round-trip run has one chain and 256 warm-up updates. Its five loop estimates are displaced upward by 4.1–6.1 reported standard errors, despite an acceptance rate of 0.812. This round-trip chain remains displaced from the reference within the sampling budget.

For a dedicated 6^4 model, direct and learned weight ESS fractions are 0.207 and 0.300; the latter has $\max |z| = 2.29$ over the five loops. The $4^4 \to 8^4$ transferred model has ESS $11/1024$ and maximum normalized weight 0.200. HMC and path-Metropolis chains were not evaluated at these larger volumes in this calculation.