

RESEARCH ARTICLE | OCTOBER 30 2026

Boltzmann generators for amorphous particle systems

Louis Grenioux  ; Leonardo Galliano  ; Ludovic Berthier   ; Giulio Biroli; Marylou Gabrié 



J. Chem. Phys. 165, 164135 (2026)

<https://doi.org/10.1063/5.0353653>



Articles You May Be Interested In

AVS Quantum Science publishes original research and more in 2021

AVS Quantum Sci. (March 2021)

Seven nonstandard models coupling quantum matter and gravity

AVS Quantum Sci. (May 2022)



The Journal of Chemical Physics

Special Topics Open for Submissions

[Learn More](#)

Boltzmann generators for amorphous particle systems

Cite as: J. Chem. Phys. 165, 164135 (2026); doi: 10.1063/5.0353653

Submitted: 23 July 2026 • Accepted: 17 September 2026 •

Published Online: 30 October 2026



Louis Grenioux,^{1,2} Leonardo Galliano,^{3,4} Ludovic Berthier,^{5,a)} Giulio Biroli,³ and Marylou Gabrié³

AFFILIATIONS

¹CMAP, CNRS, École polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France

²Center for Computational Mathematics, Flatiron Institute, New York, New York 10010, USA

³Laboratoire de Physique de l'École normale supérieure ENS, Université PSL, CNRS, Sorbonne Université, Université de Paris, 75005 Paris, France

⁴Dipartimento di Fisica, Università di Trieste, Strada Costiera 11, 34151 Trieste, Italy

⁵Gulliver, CNRS UMR 7083, ESPCI Paris, PSL Research University, 75005 Paris, France

^{a)}Author to whom correspondence should be addressed: ludovic.berthier@espci.fr

ABSTRACT

Sampling configurations in thermodynamic equilibrium is a long-standing challenge in statistical physics. Boltzmann generators address this problem by employing generative models to propose independent configurations, which are then reweighted via importance sampling using exact likelihood evaluations. Recent Boltzmann generators based on continuous normalizing flows and flow matching have achieved significant success for particle systems and biomolecules. However, these approaches have not been extended to amorphous materials (glasses), for which equilibrium sampling is notoriously slow. Because of their disordered structure, the invariances and geometrical constraints of amorphous materials differ from those of crystals and biomolecules, preventing the direct use of existing generative models. Here, we develop Boltzmann generators tailored to amorphous materials by building the required equivariances directly into Riemannian stochastic interpolants. Our framework incorporates periodic boundary conditions and particle symmetries using equivariant graph neural networks. Numerical experiments demonstrate that enforcing physical symmetries significantly improves the accuracy of Boltzmann generators but also reveal an intrinsic limitation of the continuous-flow formulation: accumulated numerical errors during likelihood integration break time-reversibility, compromising exact thermodynamic reweighting. These results reveal a fundamental challenge for continuous-flow generative models in statistical mechanics and call for alternative approaches that preserve exact thermodynamic consistency.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0353653>

I. INTRODUCTION

Sampling configurations at thermodynamic equilibrium is a long-standing challenge in statistical physics. The ability to generate equilibrium samples is essential for progress across a wide range of scientific applications, from biomolecular function and drug discovery to materials design.^{1–4} In particular, accurate equilibrium sampling underlies the modeling of protein folding and binding, as well as crystal nucleation and growth,^{5–9} with implications spanning global health and energy-storage technologies.¹⁰ Consequently, developing methods for accurate and unbiased Boltzmann sampling remains a central objective for applications in chemistry, biology, and materials science.

The difficulty of Boltzmann sampling stems from the structure of the energy landscapes that define the equilibrium measure. For many systems of interest, the effective energy is high-dimensional, rugged, and populated by numerous metastable states sometimes separated by large free energy barriers. As a consequence, classical simulation-based methods, such as Molecular Dynamics (MD) and Markov chain Monte Carlo (MCMC), can become trapped in local minima and require a computationally prohibitive number of steps to mix between relevant modes, thus producing strongly correlated samples and large statistical inefficiencies.

A recent line of work proposes to address this challenge by sampling through a trainable generator with tractable likelihoods, enabling importance-sampling reweighting to the target energy and

unbiased estimation of observables. Early advances leveraged normalizing flows^{11,12} and autoregressive models,¹³ and subsequent work has explored more expressive ordinary differential equation (ODE)-based continuous normalizing flows (CNFs).^{14–17} Overall, this paradigm, sometimes referred to as Boltzmann generators (BGs) following Ref. 11, has yielded promising results for molecular systems and many-particle benchmarks.

Among the broad class of particle systems, amorphous materials remain a particularly challenging frontier. Amorphous particle systems are disordered arrangements of interacting particles (possibly with multiple species) that lack the atomic periodicity of crystals. Prominent examples include structural glasses and glass-forming liquids.¹⁸ A major obstacle to their theoretical understanding is the extremely long equilibration timescales, even at moderate system sizes. This severely limits the reach of standard numerical sampling approaches.¹⁹ Swap Monte Carlo extends the accessible regime but remains restricted to specific models.²⁰ These limitations make amorphous materials a natural and important target for equilibrium samplers based on learnable generators.

A critical requirement for extending Boltzmann generators to amorphous materials is to explicitly encode the geometry and invariances of the underlying physical system into the models. Prior work has developed efficient equivariant BGs,^{14,15} but these models were not designed to capture the specific invariances and geometric constraints arising from periodic boundary conditions (PBCs) used to simulate amorphous systems. More recently, equivariant CNFs trained by maximum likelihood have been applied to amorphous materials²¹ but at a prohibitive computational cost due to the necessity to estimate their likelihoods in training, as will be further discussed below. Diffusion-based models provide a cheaper alternative and can generate visually realistic amorphous samples.²² However, their lack of a tractable likelihood prevents principled reweighting and therefore precludes exact Boltzmann sampling.

Motivated by these limitations, we focus on stochastic interpolants (SIs), a generalization of flow matchings (FMs).^{23–25} These methods admit a cheap training objective while retaining an expressive CNF generation mechanism, allowing (i) straightforward incorporation of symmetries via equivariant velocity fields^{14,26} and (ii) tractable likelihoods that enable importance weight corrections and unbiased estimation of expectations of observables.²⁷ Moreover, recent extensions to Riemannian manifolds provide a principled route to handle periodic boundary conditions by working directly with the appropriate geometry,^{28,29} as already exploited for crystalline settings.³⁰ Building on these ideas, our goal is to tailor Boltzmann generators to amorphous materials by enforcing the relevant geometric structure and invariances by design. Guided by these goals, we make the following contributions:

- We introduce the equivariant Riemannian stochastic interpolant (ERSI) framework, which extends Riemannian stochastic interpolants by incorporating invariance constraints tailored to amorphous particle systems.
- We prove that the optimal marginal paths and velocity fields induced by our objective respect the full symmetry group relevant to multi-species amorphous systems on the torus, providing novel theoretical guarantees within the stochastic interpolant framework.

- We adapt the graph neural network architecture of Ref. 31 to respect the full set of symmetries relevant for amorphous materials.
- We apply our framework to two canonical models of multi-component metallic glass formers and benchmark against symmetry- and geometry-agnostic baselines. Our approach yields higher-quality generations and improved agreement of physical observables after importance-sampling reweighting, while exhibiting improved scalability with system size.
- However, our study also uncovers a fundamental limitation to the use of CNF generators for reweighting due to the discretization errors arising in computing the model's likelihood, raising doubts about its general applicability for applications in statistical physics.

Incorporating invariances into push-forward generative models has been studied in Refs. 14 and 32–35, with Ref. 15 adapting FM to build equivariant transport maps, deriving results akin to our theoretical results, but limited to linear actions on $\mathbb{R}^n$. In parallel, Denoising Diffusion Probabilistic Models (DDPMs) were extended with equivariances,^{36,37} enabling generation but not likelihood evaluation. All these advances relied on equivariant GNNs,^{26,31} which also underpin our approach. Interestingly, later works showed that strong generative performance could be obtained even without explicit invariance constraints,^{38–40} suggesting that symmetries enhance efficiency and generalization, though they may not be strictly required for generation. Our experiments below demonstrate that incorporating invariances is advantageous.

Generative modeling for materials has already attracted significant attention, in particular, for crystalline structures. Like the systems considered here, crystals exhibit a number of symmetries and non-Euclidean representations. Most approaches build on Riemannian extensions of DDPMs,^{41–45} while works such as 30 and 46 resonate more closely with our setting by embedding invariances into Riemannian extensions of FMs. In this body of work, success is typically assessed by visual or structural fidelity to the training configurations, which makes the task substantially easier than checking the fidelity to the equilibrium Boltzmann distribution. Moreover, unlike amorphous materials, crystals can be described by unit cells and fractional coordinates, due to their atomic periodicity, which entails qualitatively distinct invariance structures.

This paper is organized as follows. In Sec. II, we recall the characteristics of particle mixtures and provide some background on generative models for sampling particle systems. In Sec. III, we present our equivariant Riemannian stochastic interpolant framework. In Sec. IV, we present the detailed setup of our numerical evaluation of ERSI, before presenting the results of our investigation in Sec. V.

II. GENERATIVE MODELS FOR MULTI-COMPONENT PARTICLE SYSTEMS

A. Modeling mixtures of particles

1. Definitions and equilibrium distribution

We consider classical models of multi-component particle systems, which describe a system of interacting particles at a constant temperature and volume. The microscopic state of the system is fully

specified by a configuration $C = (s, X)$, where s encodes the discrete chemical species of each particle from a finite set $\mathcal{S}$ and X denotes their spatial coordinates.

To minimize finite-size effects and accurately mimic bulk behavior, simulations must employ periodic boundary conditions. Mathematically, this corresponds to embedding the particle coordinates on a d -dimensional flat torus $\mathcal{M} = [0, L)^d$, where L is the linear size of the simulation box. This defines a Riemannian manifold, where particles exiting one face of the box smoothly re-enter from the opposite side. The natural metric on $\mathcal{M}$ is the nearest image distance,

$$d_{\mathcal{M}}(X_{(i)}, X_{(j)}) = \min_{k \in \mathbb{Z}^d} \|X_{(i)} - X_{(j)} + kL\|, \quad (1)$$

where $X_{(i)} \in \mathcal{M}$ denotes the coordinates of the i th particle. Note that $d_{\mathcal{M}}$ is a distance on $\mathcal{M}$ but not on $\mathbb{R}^d$ since it does not satisfy the triangular inequality. The full state space of the system is the product manifold $\mathcal{C} = \mathcal{S}^N \times \mathcal{M}^N$. On the torus, particle coordinates are not unique: for any $(s, X) \in \mathcal{C}$, all configurations of the form (s', X') with $s' = s$, $X' = X + kL$, $k \in \mathbb{Z}^{Nd}$ represent the same physical state. This ambiguity is lifted by applying the modulo operator component-wise,

$$A \% L = A - \left\lfloor \frac{A}{L} \right\rfloor L, \quad A \in \mathbb{R},$$

mapping all equivalent coordinates back into the fundamental domain.

The thermodynamic equilibrium of the system is determined by a potential energy function $U: \mathcal{C} \rightarrow \mathbb{R}$. For typical particle models, this is expressed as a sum of pairwise interactions based on inter-particle distances,

$$U(s, X) = \sum_{i=1}^N \sum_{j < i}^N W(s_i, s_j, d_{\mathcal{M}}(X_{(i)}, X_{(j)})), \quad (2)$$

where $W: \mathcal{S} \times \mathcal{S} \times \mathbb{R}^+ \rightarrow \mathbb{R}$ is the interaction potential. At temperature T , the equilibrium distribution is the Boltzmann measure,

$$p_{\star}(\mathrm{d}s, \mathrm{dvol}_X) = \frac{1}{\mathcal{Z}} \exp\left(-\frac{U(s, X)}{k_B T}\right) \mathrm{d}s \mathrm{dvol}_X, \quad (3)$$

where k_B is the Boltzmann constant, dvol_X is the volume element over $\mathcal{M}^N$, and the partition function $\mathcal{Z} = \int \exp(-U(s, X)/k_B T) \mathrm{d}s \mathrm{dvol}_X$ ensures the normalization. This joint distribution allows species to vary freely and corresponds to sampling from a “semi-grand” canonical ensemble, where the total number of particles is fixed, but the composition is not. Since standard models of glass-forming mixtures are defined at a fixed composition, we will also be interested in sampling from $p_{\star}$ conditioned on a composition, where the number of particles per species is strictly conserved. Sampling from this conditioned distribution corresponds to sampling the canonical ensemble. Our objective is to build a generative model that approximates $p_{\star}$ (or its conditional on the composition) to allow efficient sampling of equilibrium configurations.

2. Symmetries

For a particle mixture at equilibrium, the potential energy and, consequently, the Boltzmann distribution are invariant under transformations that preserve relative distances and species composition. In a periodic box, these symmetries correspond to a specific set of transformations on the configuration space $\mathcal{C}$, denoted collectively by the group $G_{\mathcal{C}}$ (illustrated in Fig. 1). This group comprises the following:

- **Particle permutations**, exchanging the indices of any two identical particles. For any permutation $\sigma \in S_N$, this transformation swaps both the species assignment and the spatial coordinates of the particles,

$$g_{\sigma}(s, X) = \begin{pmatrix} s_{\sigma(1)}, \dots, s_{\sigma(N)} \\ X_{(\sigma(1))}, \dots, X_{(\sigma(N))} \end{pmatrix}.$$

- **Translations**, shifting all particle coordinates by the same vector and wrapping them back onto $\mathcal{M}$. Denoting by $\mathbf{1}_N \in \mathbb{R}^N$ the vector with all coordinates equal to 1, for any $u \in \mathbb{R}^d$,

$$g_u(s, X) = \begin{pmatrix} s \\ (X + \mathbf{1}_N \otimes u) \% L \end{pmatrix}.$$

- **Box symmetries**, combining permutations of axes and sign flips of coordinates along these axes. Mathematically, this corresponds to multiplying the coordinates by a signed permutation matrix M in the d -dimensional hyperoctahedral group B_d ,

$$g_M(s, X) = \begin{pmatrix} s \\ ((\mathbf{1}_N \otimes M)X) \% L \end{pmatrix},$$

where $\mathbf{1}_N$ is the identity matrix of size $N \times N$.

While isolated particle systems, such as single molecules, may exhibit full continuous rotational invariance, restricting coordinates to a flat torus rigorously reduces these symmetries to signed permutations of the coordinate axes. We prove in Appendix A that $p_{\star}$ defined in Eq. (3) and its conditional on the composition are $G_{\mathcal{C}}$ -invariant, which means that applying any action $g \in G_{\mathcal{C}}$ to a configuration leaves its probability density unchanged.

B. Boltzmann generators

We refer to Boltzmann generators for the set of enhanced sampling schemes designed to produce unbiased equilibrium configurations using normalizing flows.^{11,12,32,47–49} Normalizing flows are generative models that consist of learning a deterministic invertible transport map that pushes a simple, tractable base distribution p_{base} (such as an ideal gas) toward a proposal distribution $\hat{q}$ that closely approximates a complex target π (such as the equilibrium Boltzmann distribution). Thanks to the invertibility of the transport map, the likelihood $\hat{q}$ of the generated samples can be evaluated exactly through a change of variable formula (see the continuous version below), and the proposal distribution can be formally corrected to the true target measure via importance sampling. In particular, denoting by $X \in \mathcal{X}$ the configurations in this generic setup, the

expectation value of an observable $\phi(X)$ under the target distribution π is estimated using the self-normalized importance sampling estimator,

$$\bar{\phi}_\pi^R = \sum_{i=1}^R \phi(X_i) \frac{\bar{w}(X_i)}{\sum_{j=1}^R \bar{w}(X_j)}, \quad X_i \sim \hat{q} \text{ i.i.d.}, \quad (4)$$

where $\bar{w}(X_i) = \bar{\pi}(X_i)/\hat{q}(X_i)$ is the unnormalized importance weight of the i th configuration, requiring the knowledge of the unnormalized density $\bar{\pi}$ of π . This estimator implicitly assumes that the support of the target distribution is completely contained within the support of the model proposal. While the self-normalized estimator $\bar{\phi}_\pi^R$ introduces a finite-sample bias, it remains asymptotically consistent and converges to the exact target expectation as the number of proposal samples $R \rightarrow \infty$. Beyond computing expectation values, the unnormalized importance weights $\bar{w}$ directly yield estimates of free energy differences between the base and target systems,^{2,17} making the framework naturally suited for targeted free energy perturbation applications.

A highly expressive framework for parameterizing these transport maps is Continuous Normalizing Flows (CNFs).^{27,50} In a CNF, the transformation is defined by an ordinary differential equation (ODE),

$$d\hat{X}_t = \hat{v}(t, \hat{X}_t)dt, \quad \hat{X}_0 \sim p_{\text{base}}, \quad (5)$$

where $\hat{v}$ is a learned time-dependent velocity field for $t \in [0, 1]$ and the output of the map is $\hat{X}_1$, the configuration reached at the final time.

This formulation yields a tractable likelihood by integration of the instantaneous change-of-variables formula (Ref. 27, Theorem 1). Specifically, the logarithm of the density $\hat{q}_t$ of $\hat{X}_t$ evolves along the ODE,

$$\begin{aligned} \frac{d}{dt} \log \hat{q}_t(\hat{X}_t) &= -\text{div } \hat{v}(t, \hat{X}_t), \\ \log \hat{q}_0(\hat{X}_0) &= \log p_{\text{base}}(\hat{X}_0), \end{aligned} \quad (6)$$

from which $\hat{q} = \hat{q}_1$ can be computed.

While access to this exact likelihood theoretically allows the model to be trained directly via maximum likelihood estimation, doing so is computationally prohibitive for high-dimensional many-body systems. The bottleneck is twofold. First, evaluating the divergence term in Eq. (6) requires computing the trace of the velocity field's Jacobian, demanding expensive auto-differentiation

at every step (although this can be somewhat mitigated by restricting the model to specific, less expressive architectures³²). Second, maximum-likelihood training is fundamentally not “simulation-free” as it requires numerically integrating the ODE and tracking its gradients at every single optimization step. Whether this is achieved through memory-efficient techniques, such as the adjoint method,²⁷ or by discretizing the ODE prior to optimization,⁵¹ the fundamental requirement of repeatedly simulating the flow trajectories renders traditional maximum likelihood training of CNFs exceptionally slow. The recently introduced stochastic interpolant framework offers an alternative, as discussed next.

C. Stochastic interpolants

Stochastic interpolants (SIs)^{23,24} are a generative modeling framework generalizing flow matching (FM)^{25,52} designed to overcome the limitations of training CNFs via maximum likelihood. SI relies on an interpolation process $(X_t)_{t \in [0,1]}$ between the base distribution $(X_0 \sim p_{\text{base}})$ and the target distribution $(X_1 \sim p_*)$. This process is defined through an interpolation function $X_t = I(t, X_0, X_1)$ satisfying the boundary conditions $I(0, X_0, X_1) = X_0$ and $I(1, X_0, X_1) = X_1$. SIs then seek to estimate a time-dependent velocity field $\hat{v}$ such that the process $(\hat{X}_t)_{t \in [0,1]}$ defined as the integration of the ODE equation (5) shares the same time-marginal distributions as $(X_t)_{t \in [0,1]}$. An exact solution v_t^* is given by the conditional expectation of the path velocities,

$$v^*(t, x) = \mathbb{E}[\partial_t I(t, X_0, X_1) \mid X_t = x]. \quad (7)$$

While this conditional expectation is analytically intractable, v^* is also a minimizer of a mean-squared regression loss $v^* \in \arg \min_{\hat{v}} \mathcal{L}(\hat{v})$, with

$$\mathcal{L}(\hat{v}) = \int_0^1 \mathbb{E}[\|\hat{v}(t, X_t) - \partial_t I(t, X_0, X_1)\|^2] dt. \quad (8)$$

This optimization problem allows us to build an empirical loss from samples of p_{base} and p_* to learn a velocity field $\hat{v}$ to approximate v^* , completely avoiding the expensive divergence computation during training by maximum likelihood. We emphasize that while SI training removes the cost of likelihood evaluation during optimization—thereby enabling the use of highly expressive network architectures—Eq. (6) must still be integrated at sampling time to compute the exact density needed to reweight physical observables.

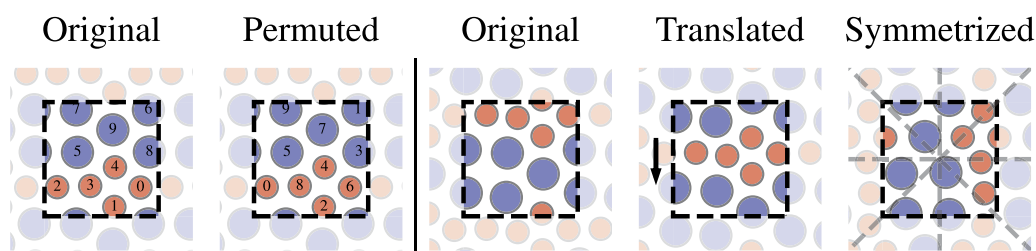


FIG. 1. Illustration of invariance group actions on a particle configuration. The system contains two particle species with different effective diameters.

Recently, the authors of Ref. 28 extended FM (SI with linear interpolants) to manifold-supported distributions, introducing Riemannian flow matching. This was shortly followed by the authors of Ref. 29, who formulated Riemannian stochastic interpolants (RSIs) as the corresponding manifold generalization of SI. In standard Euclidean space, the simplest interpolation is a linear path. The natural generalization to non-Euclidean geometries is to interpolate instead along the shortest path on the manifold, namely, a geodesic.

Following Ref. 29, the geodesic path on the flat torus can be compactly written using the manifold's exponential and logarithmic maps as $I(t, X_0, X_1) = \exp_{X_0}(t \log_{X_0}(X_1))$. These geometric operations have direct physical interpretations. The logarithmic map computes the shortest displacement vector between two configurations using the nearest-image convention,

$$\log_{X_0}(X_1) = \left(X_1 - X_0 + \frac{L}{2} \right) \% L - \frac{L}{2}, \quad (9)$$

while the exponential map applies this displacement and wraps the coordinates back into the primary simulation box,

$$\exp_X(V) = (X + V) \% L. \quad (10)$$

Because the velocity along this geodesic is constant and equal to the initial displacement $\log_{X_0}(X_1)$, the training objective from Eq. (8) simplifies directly to

$$\mathcal{L}(\hat{v}) = \int_0^1 \mathbb{E} \left[\left\| \hat{v}(t, X_t) - \log_{X_0}(X_1) \right\|^2 \right] dt. \quad (11)$$

D. Equivariant flows

Learning an accurate approximation of a complex, high-dimensional target distribution with a CNF formulation typically demands highly expressive vector fields, which, in turn, require vast parameterizations and large amounts of training data. However, the efficiency of this regression task can be drastically improved by guaranteeing by design that the generated distribution $\hat{q}$ strictly inherits the physical invariances of the target π . Equivariant normalizing flows^{14,26} build on this principle. The central result is that coupling a symmetry-invariant base distribution p_{base} with an equivariant vector field ensures that the resulting time-dependent density $\hat{q}_t$ remains strictly invariant under the symmetry group at all times $t \in [0, 1]$. Formally, a velocity field v is equivariant with respect to a symmetry group G if, for any transformation $g \in G$, it satisfies

$$v(t, g(X)) = J_g(X) v(t, X), \quad (12)$$

where J_g denotes the Jacobian of the transformation g .

When training via maximum likelihood over the family of equivariant velocity fields with a fixed invariant base distribution, as in Refs. 21 and 32, imposing the equivariance of $\hat{v}$ within the neural network architecture is sufficient: by construction, the generated density inherits the invariance of the base at all times, guaranteeing that the final distribution belongs to the correct invariance class.

In this work, however, we will consider regression-based objectives of the form (8), which are computationally cheaper but

impose additional structure: the intermediate distributions are no longer free-form but are instead constrained to follow a prescribed stochastic interpolant path. This raises two questions: (i) are the intermediate densities $(q_t)_{t \in [0,1]}$ of the interpolation process $(X_t)_{t \in [0,1]}$ invariant under the symmetry group? and (ii) if so, is the unique minimizer v^* of the regression objective equivariant? We address both questions in Sec. III, focusing on the case of particle mixtures.

III. EQUIVARIANT RIEMANNIAN STOCHASTIC INTERPOLANTS (ERSI)

To correctly model a glass-forming mixture with stochastic interpolants, the generated density must be defined on $\mathcal{C} = \mathcal{S} \times \mathcal{M}$ and be invariant under the physical symmetries introduced in Sec. II A 2. As argued below, this requires not only invariant initial and final distributions but also an interpolation path that is explicitly equivariant. We then show how to adapt an equivariant graph neural network (EGNN) architecture³¹ to the torus to guarantee that the learned generative process respects these symmetries.

A. Equivariance of the optimal velocity field

Considering the group G_C of symmetries of the Boltzmann distribution p_* , we say an interpolation function is G_C -equivariant if

$$I(t, g(C_0), g(C_1)) = g(I(t, C_0, C_1)), \quad \forall t \in [0, 1]. \quad (13)$$

Physically, this means that applying any symmetry transformation $g \in G_C$ to both the initial and final configurations must result in an identical transformation of the entire intermediate trajectory.

We establish that if both the base distribution p_{base} and the target distribution p_* are G_C -invariant, employing a G_C -equivariant interpolant guarantees that the exact target velocity field v^* defined in Eq. (7) is G_C -equivariant (see Appendix B for the proof). Moreover, we also show that the time-marginal densities of the interpolation process q_t remain G_C -invariant at all times.

These results show that the equivariance of the interpolant is crucial to ensure the consistency of the regression task in Eq. (8), where the learned velocity field $\hat{v}$ is also equivariant. To the best of our knowledge, these are original theoretical contributions to the FM and SI literature. Crucially, they show that, unlike in the optimal transport (OT) setting of Ref. 15, aligning the boundary distributions alone is insufficient.

B. Designing the interpolant

Based on the requirements derived above, we now specify our choice for the interpolation path. For the full configuration space $\mathcal{C}$, comprising both the continuous spatial coordinates and the discrete chemical species, the most natural interpolation is the shortest path, or geodesic, on $\mathcal{C}$.²⁹ Assuming the discrete species are lifted into continuous variables during the flow, the geodesic interpolant decomposes independently across the two spaces (Ref. 30, see Appendix A 3),

$$I(t, (s_0, X_0), (s_1, X_1)) = \left((1-t)s_0 + ts_1, \exp_{X_0}(t \log_{X_0}(X_1)) \right), \quad (14)$$

where the exponential and logarithmic maps are defined in Sec. II C. The first component represents the geodesic in the Euclidean space of the species variables, which corresponds to a linear interpolation. The second component is the geodesic on the flat torus, which displaces particles along the shortest nearest-image vector between X_0 and X_1 . We prove in Proposition B.8 that this interpolant is G_C -equivariant.

In practice, typical models of amorphous materials fix the composition of the mixture. We therefore target the canonical Boltzmann distribution conditioned on a specific particle composition. To enforce this in our framework, we keep the chemical species fixed along the interpolation path, setting $s_0 = s_1$. Consequently, the base distribution p_{base} is constructed as the product of a uniform spatial distribution over the torus $\mathcal{M}^N$ and a uniform distribution over all particle permutations that match the required target composition. During training, a permutation can always be found such that the initial and final species exactly align. As a result, the chemical composition remains constant along the entire interpolation path.

C. Graph neural network architecture

In order to approximate the G_C -equivariant v^* , we parameterize the velocity field $\hat{v}$ using a family of GNNs, following recent work on equivariant architectures.^{14,30,31,42} Compared to prior continuous flow models for amorphous materials relying on Gaussian radial basis functions,²¹ GNN architectures provide a far more expressive parameterization. This increased capacity enables transporting samples from a completely uniform base distribution, whereas simpler architectures only manage to start from a physically structured high-temperature liquid.²¹

The construction adapts the architecture in Ref. 31 to the torus geometry. Position variables are initialized in the input configuration $X_{(i)}^0 = X_{(i)}$, and particle features embed time and particle

species $H_i^0 = (t, s_i)$. The GNN architecture then iterates from layer k to layer $k+1$,

$$\begin{aligned} M_{ij}^k &= \hat{\phi}_e(H_i^k, H_j^k, d_{\mathcal{M}}(X_{(i)}, X_{(j)})^2), \\ P_i^k &= \sum_{j \neq i} \hat{\phi}_m(M_{ij}^k) M_{ij}^k, \quad H_i^{k+1} = \hat{\phi}_h(H_i^k, P_i^k), \\ X_{(i)}^{k+1} &= \exp_{X_{(i)}^k} \left(\sum_{j \neq i} \frac{\log_{X_{(i)}^k} X_{(j)}^k}{d_{\mathcal{M}}(X_{(i)}, X_{(j)}) + 1} \hat{\phi}_d(M_{ij}^k) \right), \end{aligned}$$

where the $\hat{\phi}_\bullet$ functions are neural networks such that $\hat{\phi}_e$ outputs edge features in $\mathbb{R}^n$ representing pairwise interactions between particles, $\hat{\phi}_m$ transforms edge messages while preserving dimension before aggregation, $\hat{\phi}_h$ updates the particle features, and $\hat{\phi}_d: \mathbb{R}^n \rightarrow \mathbb{R}^d$ decodes interaction features into displacements on the torus. For a GNN of depth K , the velocity field is finally obtained as

$$\hat{v}(t, C) = \left(\begin{matrix} \mathbf{0}_{N \times d_s} \\ \log_{X_{(1)}} X_{(1)}^K, \dots, \log_{X_{(N)}} X_{(N)}^K \end{matrix} \right), \quad (15)$$

where the null component corresponds to the preserved species-composition along the interpolation.

Unlike the linear $E(n)$ symmetries considered by the original architecture of Ref. 31, the symmetry group G_C is nonlinear. Extending equivariance from linear Euclidean spaces to G_C requires some care. We establish this rigorously in Appendix B, proving that this modified architecture defines a Lipschitz-bounded, G_C -equivariant velocity field on $\mathcal{C}$. To illustrate this, Fig. 2 compares two ODE trajectories whose initial configurations are related by a symmetry transformation. Because the learned velocity field is equivariant, this relationship is preserved throughout the entire integration. At any intermediate time t , the configurations in the two paths remain related by the exact same initial symmetry operation.

Finally, the combination of a symmetry-preserving interpolation path with a geometry-aware equivariant architecture defines the generative modeling framework used in this work, which we refer to as the Equivariant Riemannian Stochastic Interpolant (ERSI).

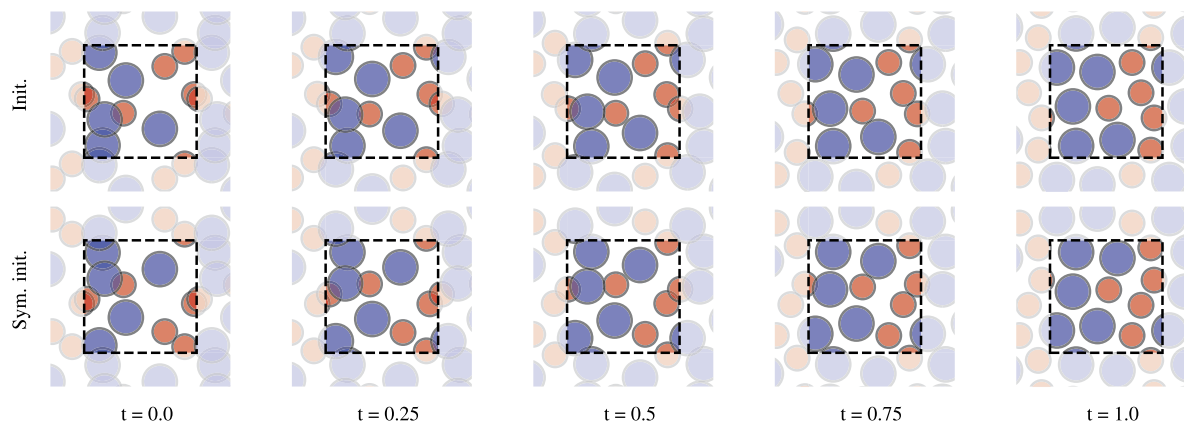


FIG. 2. Two trajectories of particle configurations generated by integrating Eq. (5) with an equivariant velocity field $\hat{v}$ started from symmetric initial configurations (first column). The intermediate configurations are related by the same initial transformation.

IV. MODELS AND METHODS

In this section, we describe the physical models and additional implementation choices made in our evaluation of the ERSI used in Sec. V.

A. Two model systems

We consider two-dimensional systems composed of $N = 10$ and $N = 44$ particles in a square box with periodic boundary conditions at constant temperature T in the canonical ensemble.

The second system^{55,56} is a ternary (5/11, 3/11, 3/11) variation of the Kob–Andersen mixture Lennard-Jones⁵⁷ introduced for efficient simulations. The interaction is defined as

$$W_{KA}(s_1, s_2, r) = \begin{cases} 4\epsilon_{s_1 s_2} W_{LJ}(s_1, s_2, r) + W_0 + W_2 \left(\frac{r}{\sigma_{s_1 s_2}} \right)^2 + W_4 \left(\frac{r}{\sigma_{s_1 s_2}} \right)^4, & r < r_{s_1 s_2}^{\text{cut}}, \\ 0, & \text{otherwise,} \end{cases}$$

with

$$W_{LJ}(s_1, s_2, r) = \left[\left(\frac{\sigma_{s_1 s_2}}{r} \right)^{12} - \left(\frac{\sigma_{s_1 s_2}}{r} \right)^6 \right],$$

$$\epsilon = \begin{pmatrix} 1.0 & 1.5 & 0.75 \\ 1.5 & 0.5 & 1.5 \\ 0.75 & 1.5 & 0.75 \end{pmatrix}, \quad \sigma = \begin{pmatrix} 1.0 & 0.8 & 0.9 \\ 0.8 & 0.88 & 0.8 \\ 0.9 & 0.8 & 0.94 \end{pmatrix},$$

$r_{s_1 s_2}^{\text{cut}} = 2.5\sigma_{s_1 s_2}$, and correction terms (W_0 , W_2 , W_4) chosen as in Ref. 55. The number density is fixed at $N/L^2 = 1.192075$, and we consider two temperatures, $T = 1.0$, which marks the onset of glassy dynamics, and $T = 0.32$, representing a deeply supercooled state.

B. Training dataset

We assume access not only to the target energy function U but also to a dataset of configurations distributed according to the equilibrium measure p_* . This auxiliary data set enables an initial learning phase in which a generative model is trained to define a proposal distribution that can later be corrected via importance sampling or even improved in an adaptive loop.⁴⁹

We generated the training datasets with the Metropolis–Hastings Monte Carlo algorithm.⁵⁸ Starting from particles uniformly distributed in the box, the update kernel consists of selecting one particle randomly with equal probability and attempting a displacement drawn from a Gaussian distribution centered at the current position with a standard deviation of 0.065.⁵⁹ The move is then accepted or rejected according to the standard Metropolis criterion. We define one unit of time as N (number of particles) attempted moves. For the binary system at temperature $T = 0.1$ and for the ternary system at temperature $T = 1.0$, we run 100 independent chains, each initialized from a different random configuration, for 10^4 time units to reach equilibrium. In this time scale, the potential energy rapidly relaxes to a steady value, and the self-intermediate scattering function¹⁸ (a standard time-correlation function used to quantify structural relaxation in liquids) decays to zero within 10^4

The first system^{53,54} is a binary (50:50) mixture interacting via the inverse power law interaction potential,

$$W_{IPL}(s_1, s_2, r) = \begin{cases} \epsilon \left(\frac{\sigma_{s_1 s_2}}{r} \right)^{12} + W_0, & r < r_{s_1 s_2}^{\text{cut}}, \\ 0, & \text{otherwise,} \end{cases}$$

with $\sigma = \begin{pmatrix} 1.0 & 1.2 \\ 1.2 & 1.4 \end{pmatrix}$ and $\epsilon = 1$. The value W_0 is chosen to shift the potential continuously to zero at $r_{s_1 s_2}^{\text{cut}} = 2.5\sigma_{s_1 s_2}$. The number density is fixed to $N/L^2 = 0.5$ and the temperature to $T = 0.1$, corresponding to a dense fluid state.

time units, confirming that the system is fully equilibrated. From these equilibrated configurations, each chain was then propagated for an additional 10^7 time units, storing one configuration every 10^4 steps.

For the ternary model at temperature $T = 0.32$, we augmented the displacement moves with particle swap moves,²⁰ applied with probability p_{swap} . In a swap proposal, two particles of different species are randomly selected, and the two species are exchanged. The proposal is then accepted or rejected according to the usual Metropolis criterion. Swap moves are known to dramatically accelerate equilibration in this specific system.⁵⁵ In this setting, we run 100 independent chains for 5×10^4 time units to reach equilibrium, followed by 5×10^7 time units for data collection, storing one configuration every 5×10^4 steps.

Both procedures yield a total of 10^5 uncorrelated equilibrium configurations, which we use as training data. The datasets are available on Zenodo.⁶⁰

C. Equivariant optimal transport

To improve training efficiency and shorten the transport trajectories, recent works replace the independent endpoint sampling $(X_0, X_1) \sim p_{\text{base}} \otimes p_*$ with a coupling closer to the optimal transport (OT) coupling $(X_0, X_1) \sim \Pi(p_{\text{base}}, p_*)$.^{61,62} To ease computations, the OT problem is solved between mini-batches of samples from p_{base} and p_* used when computing the training objective (11), as described in Ref. 63, which is an idea widely adopted for particle systems.^{15,64,65} Given two configurations $C_0 = (\vec{s}, X_0)$ and $C_1 = (\vec{s}, X_1)$ with identical composition, each is partitioned by species, and within each group, the Hungarian algorithm is applied using d_M [see Eq. (1)] as cost. This OT-based matching, consistent with the particle-permutation invariance, shortens the interpolation paths significantly. While previous work¹⁵ has accounted for additional invariances in the OT coupling, we found that further applying discrete box rotations to optimally align the spatial orientations

of C_0 and C_1 before interpolation substantially increased the computational overhead without producing measurable improvements in the thermodynamic observables. Therefore, in the results presented below, the OT coupling is restricted to particle permutations.

D. Baseline methods

To isolate the respective roles of spatial geometry and coordinate invariances, we evaluate the performance of ERSI against two reference baselines designed to separately relax these two constraints.

The first reference model is a geometry-aware but symmetry-agnostic baseline, which we refer to as Riemannian Stochastic Interpolants (RSIs). This model accounts for periodic boundary conditions by employing the geodesic interpolation from Eq. (14) directly on the flat torus $\mathcal{M}^N$ as ERSI, but it does not structurally enforce the invariances of the group G_C . The velocity field is parameterized by a standard multilayer perceptron (MLP) that maps raw coordinate arrays without permutation or translation constraints, allowing one to test whether a symmetry-agnostic network can implicitly infer these physical conservation laws from the training configurations.

The second reference model is a geometry-agnostic Equivariant Flow Matching (EFM) baseline, following the general formulation of Ref. 15. This method uses the equivariant graph neural network introduced in Ref. 31 to enforce the full $E(d)$ symmetry group of permutations, translations, and rotations within flat Euclidean space $\mathbb{R}^{Nd}$. Crucially, this baseline implicitly targets a different Boltzmann distribution than ERSI: its interaction potential is evaluated using the Euclidean distance $\|X_{(i)} - X_{(j)}\|$ rather than the torus distance $d_{\mathcal{M}}(X_{(i)}, X_{(j)})$ [Eq. (1)], and it is trained with a linear interpolant that minimizes this Euclidean distance between X_0 and X_1 , without constraining the interpolation trajectory to remain on the manifold $\mathcal{M}$. By neglecting the toroidal topology of the periodic box in both its target distribution and its interpolant, this baseline allows us to directly monitor the physical consequences of ignoring periodic boundary conditions during the generative process.

For a fair comparison, the total number of trainable parameters across the three generative models is matched, and all models are trained for the same number of epochs (see Appendix C for details).

E. Relevant physical observables

To assess the physical validity of the generated configurations, we monitor three standard structural and thermodynamic observables. First, we compute the average potential energy,

$$\langle U \rangle = \mathbb{E}_{p_*} [U_*], \quad (16)$$

of the generated configurations. Second, we evaluate the specific heat capacity at constant volume,

$$c_V = \frac{\mathbb{E}_{p_*} [U_*^2] - \mathbb{E}_{p_*} [U_*]^2}{Nk_B T^2}, \quad (17)$$

which probes energy fluctuations and is notoriously sensitive to sampling inaccuracies.^{21,66} Third, we determine the liquid structure via the radial distribution function,⁶⁷

$$g(r) = \mathbb{E}_{p_*} \left[\frac{L^2}{N^2} \sum_{i=1}^N \sum_{j \neq i}^N \delta(r - d_{\mathcal{M}}(X_{(i)}, X_{(j)})) \right], \quad (18)$$

which measures the density profile around a reference particle.

The statistical efficiency of the self-normalized reweighting protocol is determined by the configuration-space overlap between the proposal density $\hat{q}$ and the target measure p_* . If the proposal distribution deviates significantly from the target, the importance weights become highly unevenly distributed, causing the variance of the estimator to diverge. We track this sampling quality by measuring the normalized effective sample size (ESS), defined via the inverse participation ratio of the normalized weights $\bar{w}(C_i) = e^{-U_*(C_i)/k_B T} / \hat{q}(C_i)$,

$$\text{ESS}(C_{1:R}) = \frac{(\sum_{i=1}^R \bar{w}(C_i))^2}{R \sum_{i=1}^R \bar{w}_i(C_i)^2} \text{ for i.i.d. } C_i \sim \hat{q}. \quad (19)$$

The value of the ESS is bounded between $1/R$ and 1. This metric is inversely proportional to the relative variance of the importance weights, directly reflecting the asymptotic variance of the self-normalized estimator in Eq. (4).⁶⁸ A low ESS indicates an imbalanced weight distribution where the physical averages are dominated by a very small fraction of rare configurations, signaling that the proposal distribution does not generate a sufficient number of relevant states of the target system.

V. NUMERICAL RESULTS

A. Role of geometric and symmetry constraints in the binary mixture

We evaluate the performance of the Equivariant Riemannian Stochastic Interpolant (ERSI) by comparing it against the two reference baselines introduced in Sec. IV D: the geometry-agnostic but symmetry-aware EFM model and the geometry-aware but symmetry-agnostic RSI model. The efficiency of each architecture is verified by tracking the convergence of the quantities introduced in Sec. IV E as a function of the number of generated samples R . For all models, expectations are computed both directly from generated configurations and by reweighting the generated proposal samples through the importance sampling estimator in Eq. (4). Reference equilibrium values, denoted as the target, are obtained directly from the training set.

We first examine a binary mixture containing a small number of particles, $N = 10$. As shown in Fig. 3, the symmetry-agnostic RSI baseline fails to capture meaningful structure, yielding a completely flat radial distribution function $g(r)$. This indicates that a simple architecture cannot implicitly infer the particle permutation and translation invariances from the training data alone. As a consequence, reweighted estimates of physical quantities are incorrect even for a large number of samples. In contrast, the geometry-agnostic EFM performs better: the average potential energy $\langle U \rangle$ and the specific heat c_V can be recovered through reweighting, along with a noisy estimate of $g(r)$. The proposed ERSI framework displays the highest sample efficiency, converging rapidly to the exact target values for all quantities with significantly fewer samples than the EFM model.

To investigate how these architectural choices scale with the system size, we increase the number of particles to $N = 44$. Moving

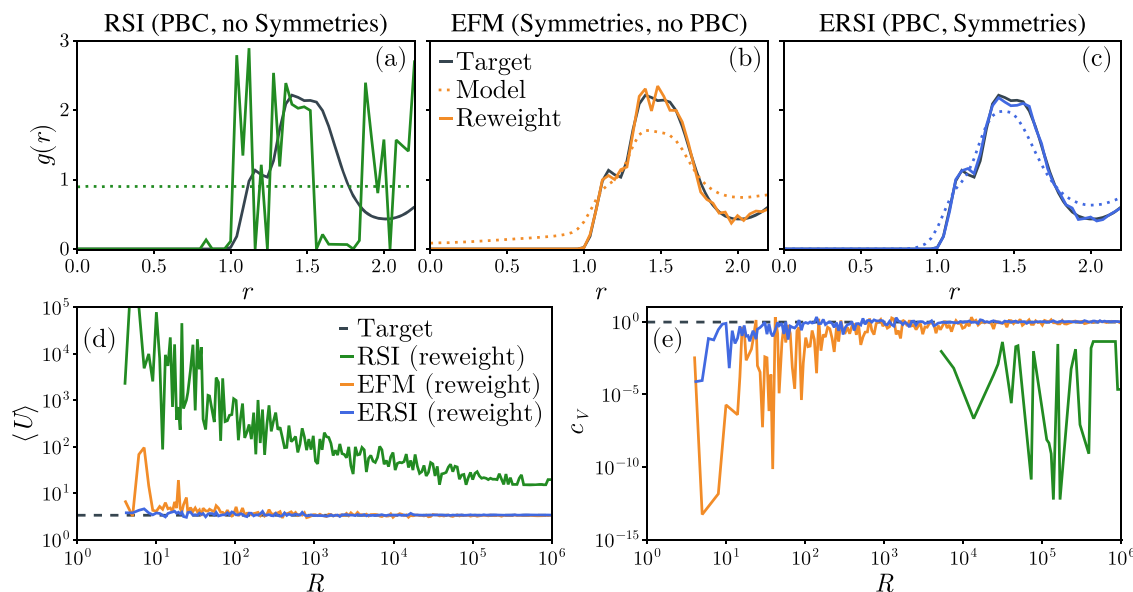


FIG. 3. Physical observables for the $N = 10$ binary mixture. Radial distribution function $g(r)$ computed for (a) RSI, (b) EFM, and (c) ERSI. Dotted, solid colored, and solid black lines represent uncorrected model proposals, reweighted estimates, and reference target values, respectively. Convergence of (d) the mean potential energy $\langle U \rangle$ and (e) the specific heat c_v as a function of the number of generated configurations R .

to a larger system increases the dimensionality and complexity of the configuration space, exposing the scaling limitations of the baselines. Typical snapshots of the generated configurations are presented in Fig. 4, alongside typical training configurations from p_* . RSI fails to move particles significantly, producing configurations that remain close to the uniform base distribution, hence its poor performance. Augmenting the training data with random actions from G_C to help the RSI model learn the invariances was attempted but did not improve performance. EFM generates more realistic configurations, but many particle overlaps occur near the box edges because periodic boundary conditions are ignored. These overlaps cause three issues: (i) averages obtained without reweighting display unphysical characteristics, such as a much higher value of $g(r < 1)$, (ii) many samples are discarded during the correction, as configurations with overlapping particle pairs receive zero weight in Eq. (4), and (iii) the likelihood of overlaps at the boundaries increases with system size, limiting the scalability of the EFM model. In contrast, configurations produced by ERSI correctly incorporate periodic boundary conditions and do not suffer from this problem, resulting in realistic configurations. This is quantitatively confirmed in Fig. 5, where ERSI is the only method that allows recovery of correct estimates of physical observables.

The difference in reweighting efficiency can be quantified by monitoring the normalized effective sample size (ESS/ R) as a function of the number of generated proposals R . In our simulations, we observe that the normalized ESS initially decays as $1/R$ over a wide range of sample sizes before eventual stabilization on a finite plateau at large R . This transient regime implies that for small R , the importance sampling estimator is structurally dominated by a singular configuration carrying nearly all the statistical weight. For the small system ($N = 10$), Fig. 6(a) shows

that the crossover scale $\bar{R}$ at which the estimator escapes this $1/R$ scaling is strongly model-dependent. For the symmetry-agnostic RSI baseline, no crossover is observed, and the estimator remains trapped in the $1/R$ regime over the entire sampling range. For the geometry-agnostic EFM model, the plateau is reached at $\bar{R} \approx 10^4$, whereas the proposed ERSI framework achieves stabilization much earlier, at $\bar{R} \approx 10^3$, demonstrating a significantly higher sample efficiency.

To rationalize this observation, we compare the energy distributions of samples from the three models with that of the target, shown in Fig. 6(c). For clarity, we discard configurations with energies larger than twice the maximum observed in the target dataset. The discarded fraction is 100% for RSI (hence no data are shown for this model), about 84% for EFM, and about 3% for ERSI. For small R , the two distributions $P(U)$ (target) and $\hat{P}(U)$ (model) overlap poorly: $\hat{P}(U)$ is concentrated at higher energies and only begins to penetrate the region of significant $P(U)$ weight from the right tail. As R increases, rare samples from this region appear with very small $\hat{P}(U)$ but relatively large $P(U)$, producing very large reweighting factors. A few such configurations dominate the averages, driving the $1/R$ behavior of the ESS. Once R exceeds $\bar{R}$, $\hat{P}(U)$ has infiltrated sufficiently deep into the bulk of $P(U)$ so that ratios $P(U)/\hat{P}(U)$ are less extreme, and the ESS stabilizes to a plateau. We repeat the analysis for $N = 44$ in Fig. 6, using up to $R = 8 \times 10^3$ samples for RSI and EFM and up to $R = 10^6$ samples for ERSI. This choice is practical: since the baselines produce heavily overlapping configurations that carry vanishing statistical weight, extending their sample sizes provides no additional statistical benefit. The results are qualitatively the same, although EFM samples are significantly worse compared to $N = 10$ due to overlapping particles at the boundaries of the physical domain (see Fig. 4).

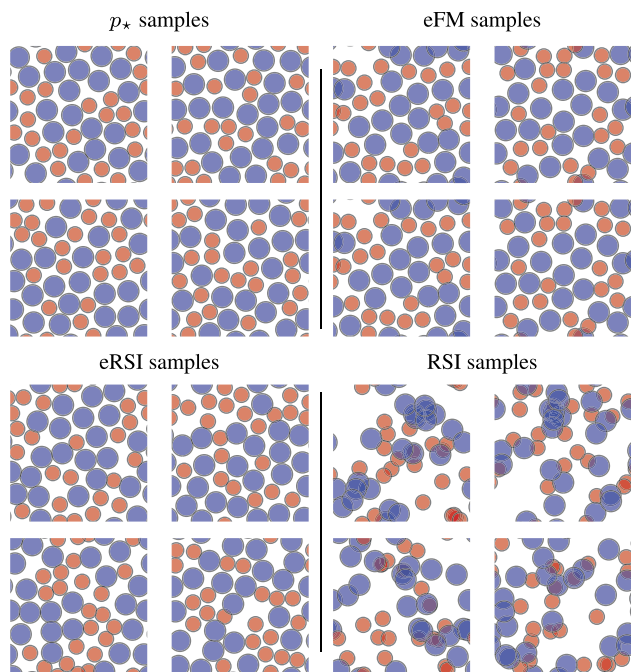


FIG. 4. Typical particle configurations for the $N = 44$ binary mixture. Snapshots obtained from the target distribution (p_*) are compared with proposals generated by the EFM, RSI, and ERSI models. The symmetry-agnostic RSI baseline fails to move particles, resulting in random configurations, while the geometry-agnostic EFM model generates realistic local packing but exhibits particle overlaps across the periodic boundaries. The ERSI framework simultaneously preserves both the local liquid structure and the boundary constraints.

B. Equilibrium sampling of the ternary mixture

Having demonstrated that incorporating explicit geometric and symmetry constraints into the generative model significantly increases sampling efficiency for a binary mixture, we extend our analysis to a more complex glass-forming model: the ternary mixture described in Sec. IV A. This system is advantageous because equilibrium reference configurations can be produced at temperatures deep into the supercooled regime using swap Monte Carlo moves (see Sec. IV B). This enables the generation of large training datasets at

low temperatures where conventional molecular dynamics struggles to equilibrate, which allows for assessing the model's performance in a regime where relaxation dynamics is very slow.

We first focus on a state point at temperature $T = 1.0$, which corresponds approximately to the onset temperature of slow dynamics for this mixture. In this state point, we systematically investigate the effect of network capacity on generative performance, evaluating whether scaling up the number of model parameters improves the convergence of the reweighted estimators. In Fig. 7(a), we plot the potential energy distribution $P(U)$ of model proposals generated by networks of varying sizes against the target reference ensemble. Figure 7(b) shows the corresponding radial distribution function $g(r)$, comparing the unweighted proposals (dotted lines) and the reweighted estimates (solid lines) against the target reference (gray). These data confirm that increasing the network capacity (ranging from $\sim 22 \times 10^3$ parameters for the smallest model to 58×10^4 parameters for the largest) systematically leads to a closer agreement with the target structure. However, this performance is saturating: while increasing the parameter count from 22×10^3 to 117×10^3 yields a substantial improvement in the unweighted mean potential energy, a further fivefold increase yields a much smaller gain. A rough extrapolation of this trend suggests that bridging the remaining gap to the average target energy purely by scaling up the network capacity would require an architecture with more than 10^7 parameters, making brute-force scaling practically unfeasible. Nevertheless, the improved quality of these proposal samples directly accelerates the convergence of the reweighted quantities. As displayed in Figs. 7(c) and 7(d), the estimators for both the mean potential energy $\langle U \rangle$ and the specific heat c_V converge toward the target values at lower values of R with larger architectures. This improved efficiency is also confirmed by the evolution of the normalized effective sample size in Fig. 7(e), where the crossover scale $\bar{R}$ at which the ESS departs from the $1/R$ regime and stabilizes on a finite plateau is lowered by a factor of 10 for the largest network compared to the smallest network. This improved performance, however, comes at the cost of ~ 1.5 times slower training and ~ 6 times slower sampling on an NVIDIA A100 GPU.

We now turn to a significantly more challenging regime by lowering the temperature to $T = 0.32$, which corresponds to a supercooled state point where the structural relaxation time of the system increases by a factor of ~ 300 compared to the onset temperature at $T = 1.0$. Numerical tests indicated that the network architecture with 58×10^4 parameters fails to achieve sufficient accuracy to allow for

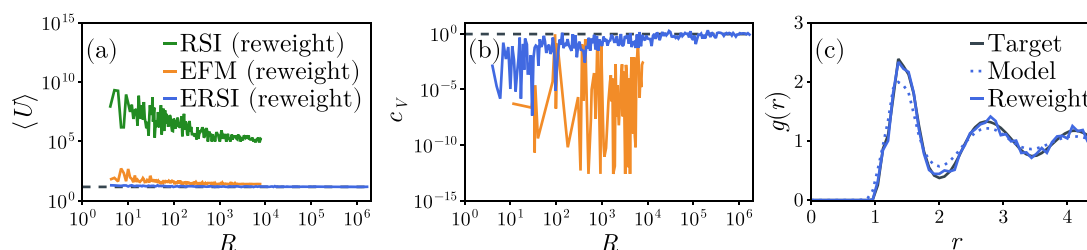


FIG. 5. Physical observables for the $N = 44$ binary mixture. Convergence of (a) the mean potential energy $\langle U \rangle$ and (b) the specific heat c_V as a function of the number of generated configurations R . (c) Radial distribution function $g(r)$ computed from 1.8×10^6 configurations generated by the ERSI framework. Reference baselines are truncated or omitted where unphysical boundary configurations or numerical instabilities prevent reliable estimates.

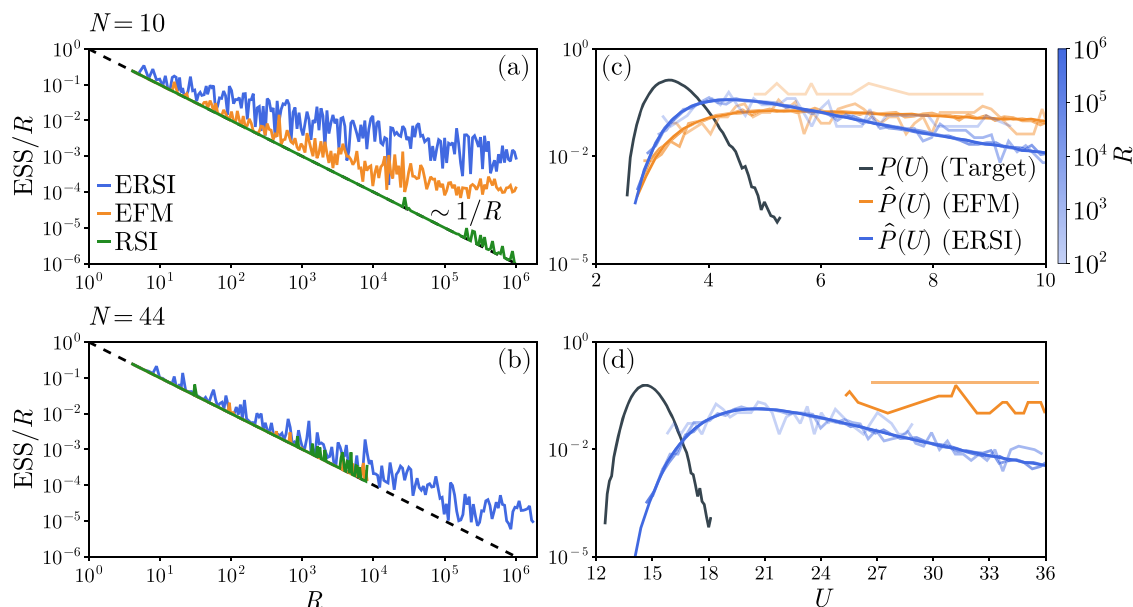


FIG. 6. Sampling efficiency and energy distribution overlap. (a) and (b) Normalized effective sample size ESS/R as a function of the number of generated configurations R for (a) $N = 10$ and (b) $N = 44$ particles. Results are shown for the symmetry-agnostic RSI baseline (green), the geometry-agnostic EFM baseline (orange), and the proposed ERSI framework (blue). Dashed black lines indicate the asymptotic $1/R$ scaling regime where a single configuration dominates the reweighted estimator. (c) and (d) Corresponding potential energy histograms $P(U)$ showing the overlap with the equilibrium target reference (dark gray solid curve) for different sample volumes R (indicated by the color bar). For clarity, unphysical high-energy configurations exceeding twice the maximum target energy are omitted from the histograms; this truncation discards 100%, $\approx 84\%$, and $\approx 3\%$ of the proposals generated at $N = 10$ by RSI, EFM, and ERSI, respectively. For $N = 44$, the maximum sample size for RSI and EFM is bounded at $R = 8 \times 10^3$ as generating more samples for these poor baselines does not provide further statistical benefit.

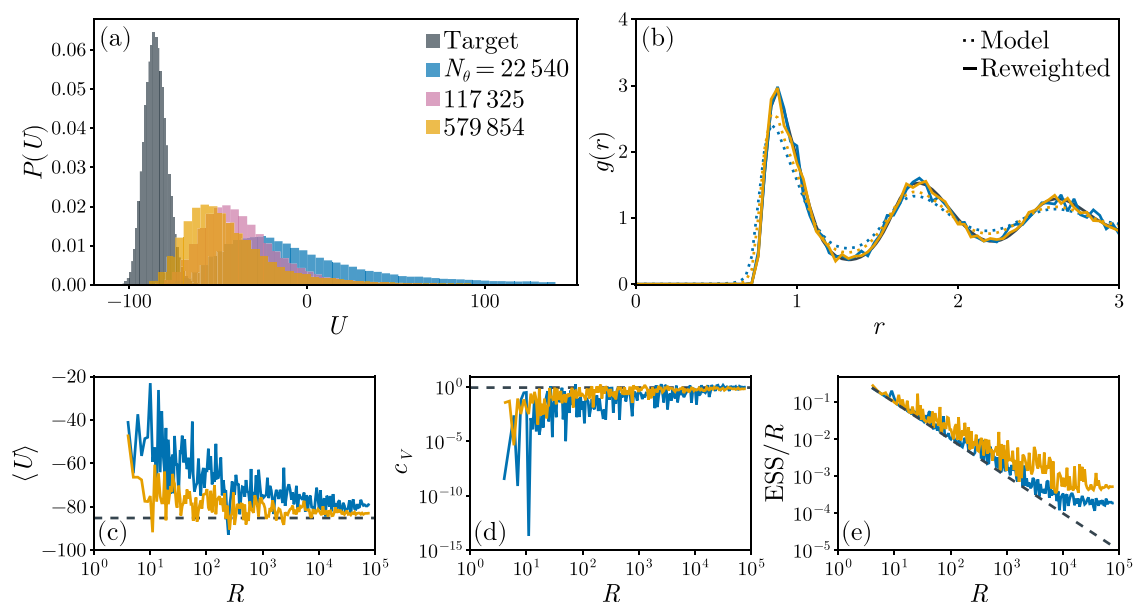


FIG. 7. Comparison of model sizes for the ternary mixture at $T = 1.0$. (a) Potential energy distribution $P(U)$ comparing the target reference (gray) with raw model proposals generated by architectures of varying numbers of trainable parameters N_θ . (b) Radial distribution function $g(r)$ comparing the target reference (gray), the unweighted model proposals (dotted lines), and the reweighted estimates (full lines) across different model sizes. (c) and (d) Convergence of the reweighted mean potential energy $\langle U \rangle$ (c) and specific heat c_V (d) as a function of the number of configurations R . (e) Evolution of the normalized effective sample size ESS/R with R .

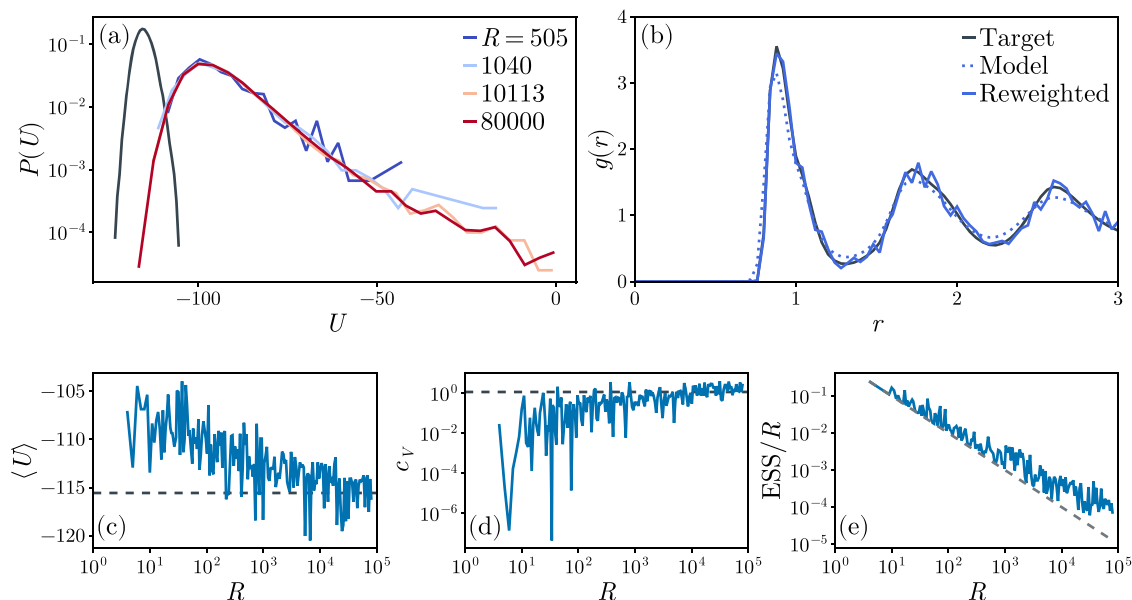


FIG. 8. Sampling efficiency of the ternary mixture in the supercooled regime at $T = 0.32$. (a) Potential energy distribution $P(U)$, comparing the target reference (gray) with the raw model proposals for different sample sizes R . (b) Radial distribution function $g(r)$ comparing the equilibrium reference (gray), the unweighted model proposals (dotted blue), and the reweighted estimates (solid blue). (c) and (d) Convergence of the reweighted mean potential energy $\langle U \rangle$ (c) and specific heat c_V (d) as a function of the number of generated configurations R . (e) Evolution of the normalized effective sample size with R settling at $\sim 10^{-4}$.

reweighting. Consequently, we scale up the network capacity to an architecture with $\sim 10^6$ parameters.

The results are summarized in Fig. 8. As shown in the potential energy distribution $P(U)$ in Fig. 8(a), the overlap between the model proposals and the equilibrium target is not large. The two distributions intersect only through their tails, where the low-energy proposals match the high-energy target samples. This poor overlap leads to low statistical efficiency when reweighting physical quantities. The radial distribution function $g(r)$ in Fig. 8(b) shows that the reweighted estimator successfully corrects the unweighted model proposals to match the target curve, although the resulting estimate remains somewhat noisy. Figures 8(c) and 8(d) show the convergence of the reweighted estimators for the mean potential energy $\langle U \rangle$ and the specific heat c_V as a function of R . Both quantities exhibit relatively large fluctuations up to large sample sizes $R \gtrsim 10^4$. This slow convergence across both structural and thermodynamic quantities is reflected in the normalized effective sample size ESS/R shown in Fig. 8(e), which follows the $1/R$ decay over most of the sample sizes, before plateauing to $\sim 10^{-4}$ above $R \approx 10^4$. This trend indicates that the reweighted averages are dominated by a small fraction of rare, low-energy configurations, and this explains the statistical noise in the pair correlation function.

C. Uncovering a fundamental limitation of continuous flows

1. Pre-filtering configurations before evaluating the density

Boltzmann generators based on continuous normalizing flows offer an elegant route to exact reweighting, but their practical

application to challenging scenarios reveals a severe computational bottleneck. Generating configurations by integrating Eq. (5) is relatively cheap, but evaluating the exact generated density $\hat{q}(C)$ requires solving a coupled ODE system with Eq. (6), which requires computing the divergence of the parameterized velocity field at each integration step. This divergence computation scales quadratically with the number of particles, and using the approximation trick of Hutchinson was shown to be too crude to perform statistical reweighting.⁶⁹ As a result, in challenging settings, where many samples are required to populate the rare low-energy states, the integration of the coupled ODE giving the likelihood becomes computationally prohibitive.

To circumvent this limitation, we can leverage the fact that, in our application, the majority of generated proposals have much higher energies compared to target samples and thus carry vanishingly small importance weights that do not contribute to equilibrium averages. We can thus implement a simple filtering strategy, where we could generate a large number of configurations using Eq. (5) without computing their densities. We then evaluate their potential energies U (this is numerically cheap) in order to restrict the expensive density evaluation to a small fraction α corresponding to the lowest-energy samples, i.e., to samples that have a chance to contribute to the equilibrium average. The density $\hat{q}$ for these selected configurations is obtained by integrating Eq. (6) backward from $t = 1$ to $t = 0$. A similar approach was formally introduced in Ref. 70, where analytical bounds for the resulting truncation error were derived. Here, we evaluate the robustness of this protocol empirically using our results for the ternary mixture at $T = 0.32$. Specifically, we fix a target threshold for the retained partition function mass, $\mathcal{Z}_\alpha / \mathcal{Z} \approx \sum_{i=1}^{R_\alpha} \bar{w}(C_i) / \sum_{i=1}^R \bar{w}(C_i)$, and track the minimum

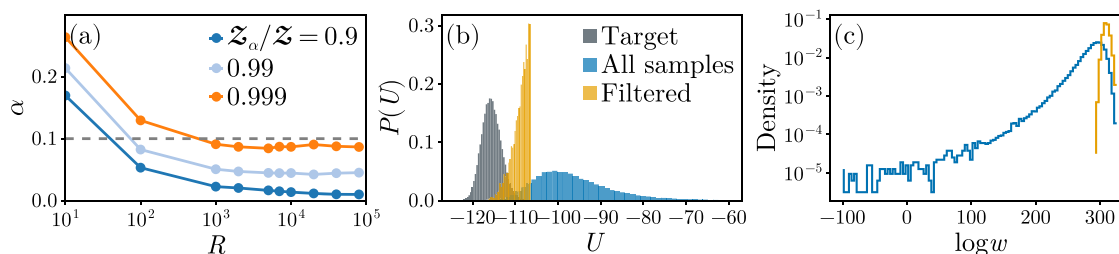


FIG. 9. Energy-filtering protocol. (a) Fraction α of lowest-energy samples needed to preserve a given fraction Z_α/Z of the estimated target partition function as a function of the total number of samples R . (b) Truncated potential energy histogram showing the discarded high-energy configurations. (c) Distribution of the unnormalized importance weights before and after applying the 10% filtering threshold.

fraction of configurations α required to satisfy this threshold across different total sample sizes R . As shown in Fig. 9(a), retaining only the lowest-energy 10% of the generated configurations consistently preserves more than 99.9% of the total partition function across sample sizes R . The corresponding truncated potential energy histogram and the log-weight distributions before and after filtering are displayed in Figs. 9(b) and 9(c) for a total of 8×10^4 samples. These distributions confirm that the filtering process safely removes configurations with negligible statistical weight without altering the largest weights, which would otherwise introduce an uncontrolled bias.

Motivated by this validation, we generate 64×10^4 independent samples from Eq. (5) and evaluate the density $\hat{q}(C)$ via backward integration of Eq. (6), which is ~ 150 times more expensive, only for the 10% lowest-energy configurations. The statistical effect of this selection is illustrated in Fig. 10(a) where the normalized ESS stabilizes near 10^{-3} rather than at 10^{-4} [recall Fig. 8(e)].

These results demonstrate that pre-filtering configurations using the potential energy before performing the more costly computation of their density is a valid strategy that allows for a much faster and more efficient convergence of equilibrium measurements via importance sampling.

2. Fundamental limitation of density estimation

However, this increased statistical resolution reveals that the estimate for $\langle U \rangle$ and c_V is systematically biased. Figures 10(b) and 10(c) show that while the plateau values remain within the fluctuations of the target data, the estimators clearly converge to a value

that is slightly larger than the true target mean. Since our prior analysis demonstrated that the filtering protocol itself introduces no detectable bias, this systematic shift must represent an intrinsic limitation of the flow architecture that becomes visible only because the statistical variance has been considerably reduced.

We found that the source of this bias lies in the numerical integration of the flow trajectories needed to estimate densities. In a high-dimensional configuration space, integrating trajectories under a finite numerical tolerance introduces small truncation errors at each step. This numerical drift breaks time-reversibility: comparing the importance weights obtained via forward and backward integration for the same physical configurations reveals a mean relative discrepancy of $\sim 10\%$. Crucially, both integration directions are equally affected, and one of them cannot be assumed to be better than the other, as both represent distinct numerical approximations of the same underlying continuous flow.⁷¹

The numerical drift appears as the interplay between system complexity and model capacity. Sampling at lower temperatures T requires deeper, heavier networks to capture finer details of the underlying energy landscape. These highly parameterized velocity fields exhibit strong local curvature and numerical stiffness, making ODE trajectories more prone to discretization errors that accumulate across all dimensions. To quantitatively characterize this integration difficulty, we evaluated the maximum singular value (spectral norm) of the Jacobian of the velocity field, $\sigma_{\max}(t) = \|\partial \dot{v} / \partial X\|_2$, along the flow trajectories. For the onset temperature $T = 1.0$, the maximum spectral norm averages 6.3 at $t = 0.5$. Deep in the supercooled regime at $T = 0.32$, this value more than

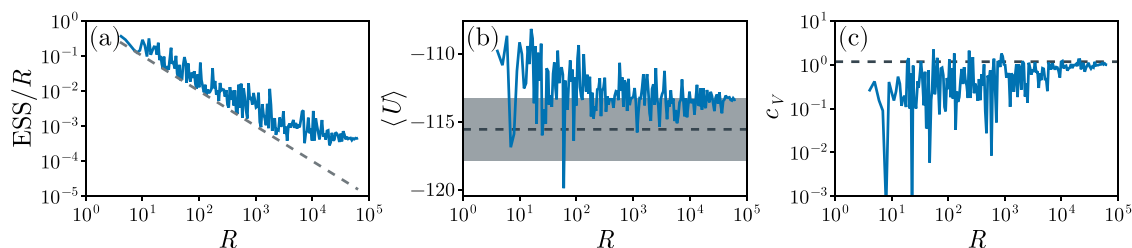


FIG. 10. Reweighting with filtered samples. (a) Evolution of the normalized effective sample size ESS/R as a function of R with energy filtering. (b) Convergence of the reweighted mean potential energy $\langle U \rangle$ using the filtered samples. The shaded gray band denotes the equilibrium energy fluctuations, exposing a systematic numerical bias. (c) Convergence of the specific heat c_V as a function of R .

doubles to 13.8. This larger stiffness directly impacts the integration cost, as the adaptive ODE solver requires 40% more steps at the lower temperature to maintain the same numerical tolerance. The numerical errors are therefore larger at lower temperatures because the integration space is larger. For highly demanding tasks such as sampling glass-forming mixtures, these accumulated errors alter the weight distribution, injecting a systematic bias that compromises the reliability of exact reweighting.

In addition, attempting to suppress these errors by tightening the ODE solver tolerances rapidly becomes unfeasible: evaluating the exact density requires computing the divergence $\nabla \cdot \dot{v}$ at each step via dN backward passes. For stiff flows where the required number of solver steps is already large, tightening tolerances drastically increases the computational cost. Because this drift cannot be overcome by brute-force integration, it constitutes a fundamental bottleneck for continuous-time Boltzmann generators. We conclude that the strategy will be limited by the large number of parameters needed for larger and more complex systems.

VI. CONCLUSION AND PERSPECTIVES

We have developed the Equivariant Riemannian Stochastic Interpolant (ERSI) framework, a generative modeling approach tailored to sample equilibrium configurations of bulk amorphous particle systems with multiple species. In doing so, we establish the exact mathematical requirements under which continuous normalizing flows trained via stochastic interpolants preserve physical symmetries. Specifically, we proved that when the boundary distributions are invariant, restricting the interpolation to a geodesic path guarantees that the optimal time-marginal densities and the optimal velocity field respect the full symmetry group G_C of the multi-component mixture on C . This provides a justification for parameterizing the vector field using a geometry-aware, equivariant graph neural network.

Benchmark simulations on dense binary and supercooled ternary mixtures demonstrate that explicitly encoding geometric structure and physical symmetries yields more accurate equilibrium sampling and more reliable estimates of physical observables, using fewer generated samples and scaling to larger system sizes than architectures that neglect these properties.

Because Boltzmann generators rely on importance sampling for thermodynamic reweighting, accessing the exact likelihood is essential. In our continuous flow formulation, tracking this density requires integrating a coupled ODE system whose computational cost is much higher than simply generating samples. This presents a significant bottleneck when gathering large sample volumes.

Even more crucially, our work reveals an intrinsic limitation of continuous normalizing flows. Since the adaptive ODE solver is imperfect, local truncation errors accumulate along the continuous trajectories. This breaks time-reversibility and corrupts the importance sampling, which eventually produces biased thermodynamic averages. The limitations become more acute for larger models that are needed to generate configurations of larger systems with many-body interactions. Our results therefore reveal a

fundamental challenge for continuous-flow generative models in statistical mechanics and call for alternative approaches that preserve exact thermodynamic consistency.

Looking forward, our results point to interesting directions for future research. First, the systematic bias introduced by continuous ODE solvers could be bypassed by adopting discrete-time formulations. Recent developments in diffusion-based sampling^{72,73} directly discretize the equations of motion to compute the exact likelihood natively through discrete transition kernels, offering a direct route to eliminate integration errors.

Second, our approach assumes access to an initial dataset of biased samples for training the Boltzmann generator. Relaxing this assumption and developing training strategies that do not rely on such *a priori* data is an important next step. Promising directions include adaptive MCMCs,⁴⁹ possibly combined with sequential tempering,^{13,74,75} as proposed for normalizing flows and autoregressive models

ACKNOWLEDGMENTS

This work was partially performed using HPC resources from GENCI-IDRIS (Grant Nos. AD011015234R1, AD011015234R2, and AD011015234). G.B. and M.G. acknowledge support from PRAIRIE-PSAI Grant No. ANR-23-IACL-0008, managed under the France 2030 program. L.B. acknowledges the support of the French Agence Nationale de la Recherche (ANR) under Grant No. ANR-24-CE30-0442 (project GLASSGO). Leonardo Galliano acknowledges co-funding from the European Union - Next Generation EU. Louis Grenioux acknowledges funding from Hi! Paris. We thank Gerhard Jung for useful discussions and Jérémy Diharce for his precious involvement in previous versions of this project.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Louis Grenioux and Leonardo Galliano contributed equally to this work.

Louis Grenioux: Conceptualization (equal); Writing – original draft (equal). **Leonardo Galliano:** Conceptualization (equal); Writing – original draft (equal). **Ludovic Berthier:** Conceptualization (equal); Writing – original draft (equal). **Giulio Biroli:** Conceptualization (equal); Writing – original draft (equal). **Marylou Gabrié:** Conceptualization (equal); Writing – original draft (equal).

DATA AVAILABILITY

The data that support the findings of this study are openly available in Zenodo at <https://doi.org/10.5281/zenodo.23388791>.⁶⁰

APPENDIX A: TARGET DISTRIBUTION'S SYMMETRY GROUP

In this appendix, we establish that the target Boltzmann distribution p_* is strictly invariant under the group of physical symmetry transformations G_C introduced in Sec. II A 2. We first connect the nearest image distance d_M with the logarithmic map on the torus in Proposition A.9. We then establish the G_C -equivariance of both the logarithmic and exponential maps (Proposition A.12, Corollary A.13, and Proposition A.11). Finally, Corollaries A.14 and A.15 synthesize these geometric results to prove our main outcome: that the metric d_M , the potential energy function U , the target density p_* , and the corresponding canonical distribution p_*^u at fixed composition are all G_C -invariant.

Definition A.1. Let G denote a set of group actions acting on the configuration space $\mathcal{C}$. A probability density q on $\mathcal{C}$ is said to be G -invariant if, for all $g \in G$ and all $C \in \mathcal{C}$, $q(g(C)) = q(C)$.

Lemma A.2. Let $g \in G_C$. The action can be decomposed as

$$g : \begin{pmatrix} s \\ X \end{pmatrix} \mapsto \begin{pmatrix} g^s(C) \\ g^X(C) \end{pmatrix} = \begin{pmatrix} A_\sigma s \\ (A_\sigma \otimes I_d) h_{M,c,L}(X) \end{pmatrix},$$

$$h_{M,c,L}(X) = ((I_N \otimes M)X + (\mathbf{1}_N \otimes c)) \% L,$$

where $\sigma \in S_N$ is a permutation, A_σ is the associated permutation matrix of size $N \times N$, M is a signed permutation matrix of size $d \times d$, and c is a vector in $\mathbb{R}^d$. The function $h_{M,c,L}$ could be also un-vectorized as

$$h_{M,c,L}(X) = \{\tilde{h}_{M,c,L}(X_{(i)})\}_{i=1}^N, \quad \tilde{h}_{M,c,L}(y) = (My + c) \% L,$$

which also means that for all $i \in \llbracket 1, N \rrbracket$,

$$g^X(C)_{(i)} = \tilde{h}_{M,c,L}(X_{(\sigma(i))}).$$

Note that for any $g \in G_C$, g can be written as $g = f_L \circ h_{A,b}$, where

$$f_L : (s, X) \mapsto \begin{pmatrix} s \\ X \% L \end{pmatrix} \text{ and } h_{A,b} : C \mapsto AC + b, \quad (\text{A1})$$

with A being an orthogonal matrix and b being a vector in $\mathbb{R}^{N|s|+Nd}$, which are defined in Lemma A.2. The set of actions induced by $h_{M,c,L}$ on $\mathcal{M}^N$ is denoted $\mathcal{G}_{\mathcal{M}^N}$, and the one induced by $\tilde{h}_{M,c,L}$ on $\mathcal{M}$ is denoted $\mathcal{G}_{\mathcal{M}}$. From Lemma A.2, we also deduce that A and b in Eq. (A1) can be written as

$$A = \begin{pmatrix} A_\sigma & 0 \\ 0 & (A_\sigma \otimes I_d)(I_N \otimes M) \end{pmatrix}, \quad b = \mathbf{1}_N \otimes c. \quad (\text{A2})$$

Lemma A.3. Let $X, Y \in \mathbb{R}^{Nd}$. We have

$$((X \% L) + Y) \% L = (X + Y) \% L. \quad (\text{A3})$$

Proof. We have that $X = (X \% L) + k_X L$, where $k_X \in \mathbb{Z}^{Nd}$, which leads to

$$((X \% L) + Y) \% L = (X + Y - k_X) \% L = (X + Y) \% L.$$

Lemma A.4. Let $y \in \mathbb{R}$ and $\epsilon \in \{-1, +1\}$, then $(\epsilon y) \% L = (\epsilon(y \% L)) \% L$.

Proof. If $\epsilon = 1$, this is obviously true. If $\epsilon = -1$, we are trying to prove that

$$(-y) \% L = -(y \% L) \% L.$$

By definition, we have that

$$(-y) \% L = -y - \left\lfloor \frac{-y}{L} \right\rfloor L,$$

$$(-(y \% L)) \% L = -y - \left(\left\lfloor \frac{-y}{L} + k \right\rfloor - k \right) L,$$

where $k = \lfloor y/L \rfloor$. Using the fact that for any $x \in \mathbb{R}$ and $n \in \mathbb{N}$,

$$\lfloor x + n \rfloor - n = \lfloor x \rfloor,$$

we get the intended result. $\square$

Lemma A.5. Let $M \in \mathcal{B}_d$ and $X \in \mathbb{R}^d$, then $(MX) \% L = (M[X \% L]) \% L$.

Proof. For any $M \in \mathcal{B}_d$, there exist $\epsilon \in \{-1, 1\}^d$ and $\sigma \in S_d$ such that for all $i, j \in \llbracket 1, d \rrbracket$, $M_{ij} = \epsilon(i)\delta_{\sigma(i),j}$. For all $i \in \llbracket 1, d \rrbracket$, we have

$$((MX) \% L)_i = (\epsilon(i)X_{\sigma(i)}) \% L,$$

$$(M[X \% L])_i = \epsilon(i)[X \% L]_{\sigma(i)}.$$

Use Lemma A.4 on each coordinate to conclude. $\square$

Proposition A.6. The logarithmic map can be written for any $X \in \mathcal{M}$ and $Y \in \mathcal{T}_X \mathcal{M}$,

$$\log_X(Y) = \frac{L}{2\pi} \text{atan2} \left(\sin \left(\frac{2\pi}{L} (X - Y) \right), \cos \left(\frac{2\pi}{L} (X - Y) \right) \right).$$

Proof. For any $X \in \mathcal{M}$ and $Y \in \mathcal{T}_X \mathcal{M}$, we have that

$$\begin{aligned} \log_X(Y) &= \left(Y - X + \frac{L}{2} \right) \% L - \frac{L}{2} = Y - X - L \left\lfloor \frac{Y - X + \frac{L}{2}}{L} \right\rfloor = \frac{L}{\pi} \arctan \left(\tan \left(\frac{\pi}{L} (Y - X) \right) \right) \quad \left(\text{Using } \arctan \tan \alpha = \alpha - \pi \left\lfloor \frac{\alpha}{\pi} + \frac{1}{2} \right\rfloor \right) \\ &= \frac{L}{\pi} \arctan \left(\frac{\sin \left(\frac{2\pi}{L} (Y - X) \right)}{1 + \cos \left(\frac{2\pi}{L} (Y - X) \right)} \right) \quad \left(\text{Using } \tan \alpha = \frac{\sin 2\alpha}{1 + \cos 2\alpha} \right) \\ &= \frac{L}{2\pi} \text{atan2} \left(\sin \left(\frac{2\pi}{L} (X - Y) \right), \cos \left(\frac{2\pi}{L} (X - Y) \right) \right). \end{aligned}$$

$\square$

Lemma A.7. For all $a \in \mathbb{R}$, $-\lfloor \frac{a}{L} + \frac{1}{2} \rfloor = \arg \min_{k \in \mathbb{Z}} |a + kL|$ holds.

Proof. By definition, we have that

$$\left\lfloor \frac{a}{L} \right\rfloor L \leq a \leq \left(\left\lfloor \frac{a}{L} \right\rfloor + 1 \right) L,$$

which means that

$$k^* = \arg \min_{k \in \mathbb{Z}} |a + kL| \in \left\{ -\left\lfloor \frac{a}{L} \right\rfloor, -\left(\left\lfloor \frac{a}{L} \right\rfloor + 1 \right) \right\}.$$

Case 1. $k^* = -\lfloor a/L \rfloor$ This means that a/L is closer to $\lfloor a/L \rfloor$ than to $\lfloor a/L \rfloor + 1$, which implies that $a/L + 1/2$ is also closer to $\lfloor a/L \rfloor$, which means that

$$\left\lfloor \frac{a}{L} \right\rfloor = \left\lfloor \frac{a}{L} + \frac{1}{2} \right\rfloor$$

and that $k^* = -\lfloor a/L + 1/2 \rfloor$.

Case 2. $k^* = -(\lfloor a/L \rfloor + 1)$ This means that a/L is closer to $\lfloor a/L \rfloor + 1$ than to $\lfloor a/L \rfloor$, which implies that $a/L + 1/2$ is also closer to $\lfloor a/L \rfloor + 1$, which means that

$$\left\lfloor \frac{a}{L} \right\rfloor + 1 = \left\lfloor \frac{a}{L} + \frac{1}{2} \right\rfloor$$

and that $k^* = -\lfloor a/L + 1/2 \rfloor$. $\square$

Lemma A.8. Let $f : \mathbb{R}^k \rightarrow \mathbb{R}^+$ be defined for all $x \in \mathbb{R}^k$ as

$$f(x) = \sum_{i=1}^k f_i(x_i), \quad f_i : \mathbb{R} \rightarrow \mathbb{R}^+.$$

Then, the minimum of f decomposes as

$$\arg \min_{x \in \mathbb{R}^k} f(x) = \arg \min_{x_1 \in \mathbb{R}} f_1(x) \times \cdots \times \arg \min_{x_k \in \mathbb{R}} f_k(x).$$

Proof. Suppose that there exist $\tilde{x} \in \arg \min_{x \in \mathbb{R}^k} f(x)$ and $\tilde{x} \notin \arg \min_{x_1 \in \mathbb{R}} f_1(x) \times \cdots \times \arg \min_{x_k \in \mathbb{R}} f_k(x)$. Then, there exists $i \in \{1, \dots, k\}$ such that $\tilde{x}_i \notin \arg \min_{x_i \in \mathbb{R}} f_i(x)$. Let $\hat{x}_i \in \arg \min_{x_i \in \mathbb{R}} f_i(x)$, then $\hat{x} \in \mathbb{R}^k$ is defined as

$$\hat{x}_j = \begin{cases} \tilde{x}_j & \text{if } j \neq i, \\ \hat{x}_i & \text{otherwise.} \end{cases}$$

Then, $f(\hat{x}) \leq f(\tilde{x})$, which contradicts the assumptions. This gives the left-right inclusion, and the right-left one is trivial. $\square$

Proposition A.9. Let $X, Y \in \mathcal{M}$, then the pre-metric $d_{\mathcal{M}}$ can be written as

$$d_{\mathcal{M}}(X, Y) = \|\log_X(Y)\| \quad \text{for any } X, Y \in \mathcal{M}.$$

Proof. We have that $\log_X(Y) = Y - X + k_{X,Y}L$, where

$$k_{X,Y} = -\left\lfloor \frac{Y - X + \frac{L}{2}}{L} \right\rfloor = -\left\lfloor \frac{Y - X}{L} + \frac{1}{2} \right\rfloor.$$

Moreover, we have that

$$d_{\mathcal{M}}(X, Y) = \min_{k \in \mathbb{Z}^d} \|X - Y + kL\| = \min_{k \in \mathbb{Z}^d} \|Y - X + kL\|$$

$$= \min_{k \in \mathbb{Z}^d} \|Y - X + kL\|^2,$$

which implies that

$$d_{\mathcal{M}}(X, Y) = \min_{k \in \mathbb{Z}^d} \sum_{i=1}^d (Y_i - X_i + k_i L)^2.$$

Using Lemma A.7 and the monotonicity of the square root on $\mathbb{R}^+$, we have that for all $i \in \llbracket 1, d \rrbracket$,

$$\arg \min_{k \in \mathbb{Z}} (Y_i - X_i + kL)^2 = [k_{X,Y}]_i.$$

By Lemma A.8, this implies that

$$\arg \min_{k \in \mathbb{Z}^d} \sum_{i=1}^d (Y_i - X_i + k_i L)^2 = k_{X,Y},$$

leading to

$$d_{\mathcal{M}}(X, Y) = \|Y - X + k_{X,Y}L\| = \|\log_X(Y)\|. \quad \square$$

Proposition A.10. Let $X, Y \in \mathcal{M}$, then a pre-metric (see Ref. 28, Sec. 3.2 for the definition) between X and Y in $\mathcal{M}$ can be defined as

$$d_{\mathcal{M}}(X, Y) = \min_{k \in \mathbb{Z}^d} d^E(X, Y + kL), \quad (\text{A4})$$

where $d^E(X, Y) = \|X - Y\|$ is the Euclidean distance.

Proof. Let $X, Y \in \mathcal{M}$; we obviously have that $d_{\mathcal{M}}(X, Y) \geq 0$, checking the non-negativity. Moreover, using the fact that d^E is a distance,

$$d_{\mathcal{M}}(X, Y) = 0 \iff \exists k \in \mathbb{Z}^d,$$

$$d^E(X, Y + kL) = 0 \iff X = Y \% L \iff X = Y,$$

which checks positivity. Using Proposition A.9, we have that $\nabla d_{\mathcal{M}}(X, Y) = \log_X Y$, which leads to

$$\nabla d_{\mathcal{M}}(X, Y) \neq 0 \iff \log_X Y \neq 0 \iff \left(Y - X + \frac{L}{2} \right) \% L - \frac{L}{2} \neq 0$$

$$\iff \exists ! k \in \mathbb{Z}^{Nd}, Y - X + \frac{L}{2} + kL - \frac{L}{2} \neq 0$$

$$\iff (Y - X) \% L \neq 0 \iff X \neq Y. \quad \square$$

Proposition A.11. Let $A, B \in \mathbb{R}^d$, $M \in B_d$, and $u \in \mathbb{R}^d$, then

$$\exp_{(MA+u) \% L}(MB) = (M \exp_A B + u) \% L$$

holds.

Proof. Using Lemma A.3, we have that

$$\exp_{(MA+u) \% L}(MB) = \left(\exp_{(MA) \% L} MB + u \right) \% L$$

Moreover, given that the signed permutation matrix M is characterized by permutation σ and a vector $\epsilon \in \{-1, 1\}^d$, we have for all $i \in \llbracket 1, d \rrbracket$,

$$[M(A+B)]_i = \epsilon_i(A_{\sigma(i)} + B_{\sigma(i)}).$$

Using Lemmas A.3 and A.4, we get that

$$\begin{aligned} [\exp_{(MA)} \% L MB]_i &= (\epsilon_i(A_{\sigma(i)} + B_{\sigma(i)})) \% L \\ &= (\epsilon_i[(A_{\sigma(i)} + B_{\sigma(i)}) \% L]) \% L. \end{aligned}$$

Additionally, note that

$$[(M \exp_A B) \% L]_i = (\epsilon_i[(A_{\sigma(i)} + B_{\sigma(i)}) \% L]) \% L,$$

which proves that $\exp_{(MA)} \% L MB = (M \exp_A B) \% L$. If we put this expression in the first equation and use Lemma A.3, we get

$$\begin{aligned} \exp_{(MA+u)} \% L (MB) &= ((M \exp_A B) \% L + u) \% L \\ &= (M \exp_A B + u) \% L. \end{aligned}$$

Proposition A.12. Let $X, Y \in \mathcal{M}$. For any $M \in B_d$,

$$\log_{(MX)} \% L ((MY) \% L) = M \log_X(Y)$$

holds.

Proof. Let $x, y \in [0, L]$, we have that

$$\begin{aligned} \log_{(-x)} \% L ((-y) \% L) &= \frac{L}{\pi} \arctan \left(\frac{\sin \left(\frac{2\pi}{L} ((-y) \% L - (-x) \% L) \right)}{1 + \cos \left(\frac{2\pi}{L} ((-y) \% L - (-x) \% L) \right)} \right). \end{aligned}$$

By definition, there exists $k, k' \in \mathbb{Z}$ such that

$$(-x) \% L = -x + kL \quad \text{and} \quad (-y) \% L = -y + k'L.$$

This leads to

$$\begin{aligned} \frac{2\pi}{L} ((-y) \% L - (-x) \% L) &= \frac{2\pi}{L} ([-y + k'L] - [-x + kL]), \\ &= -\frac{2\pi}{L} (y - x) + 2\pi \underbrace{(k - k')}_{\in \mathbb{Z}}. \end{aligned}$$

Using the periodicity of the sine and cosine in the above formula, we obtain

$$\begin{aligned} \log_{(-x)} \% L ((-y) \% L) &= \frac{L}{\pi} \arctan \left(\frac{-\sin \left(\frac{2\pi}{L} (y - x) \right)}{1 + \cos \left(\frac{2\pi}{L} (y - x) \right)} \right) = -\log_X(y). \end{aligned}$$

Let $M \in B_d$ and $X, Y \in \mathcal{M}$. By definition, for all $i \in \llbracket 1, d \rrbracket$, $(MX)_i = \epsilon_i X_{\sigma(i)}$. Moreover, using the previous result,

$$\begin{aligned} [\log_{(MX)} \% L ((MY) \% L)]_i &= \log_{[(MX) \% L]_i}([(MY) \% L]_i) \\ &= \log_{[(MX)_i] \% L}([(MY)_i] \% L), \end{aligned}$$

$$\begin{aligned} &= \log_{(\epsilon_i X_{\sigma(i)}) \% L}((\epsilon_i Y_{\sigma(i)}) \% L) = \epsilon_i \log_{X_{\sigma(i)}}(Y_{\sigma(i)}), \\ &= [M \log_X(Y)]_i, \end{aligned}$$

which concludes the proof. $\square$

Corollary A.13. Let $X, Y \in \mathcal{M}$. For any $M \in B_d$ and $u \in \mathbb{R}^d$,

$$\log_{(MX+u)} \% L ((MY + u) \% L) = M \log_X(Y). \quad (\text{A5})$$

Proof. Using Lemma A.3 applied twice, we have

$$\begin{aligned} \log_{(X+u)} \% L ((Y+u) \% L) &= \left((Y+u) \% L - (Y+u) \% L + \frac{L}{2} \right) \% L - \frac{L}{2} \\ &= \left(Y+u - Y-u + \frac{L}{2} \right) \% L - \frac{L}{2} \\ &= \log_X \% L (Y \% L). \end{aligned}$$

The result comes by applying this remark in Proposition A.12 with $MX \in \mathcal{M}$ and $MY \in \mathcal{M}$. $\square$

Corollary A.14. The pre-metric $d_{\mathcal{M}}$ is $\mathcal{G}_{\mathcal{M}}$ -invariant, i.e., for all $X, Y \in \mathcal{M}$,

$$d_{\mathcal{M}}(g(X), g(Y)) = d_{\mathcal{M}}(X, Y) \quad \text{for all } g \in \mathcal{G}_{\mathcal{M}}.$$

Consequently, the potential U (2) and the induced density p_* (3) are both G_c -invariant.

Proof. This comes from the fact that, according to Proposition A.9, the distance can be written as the norm of the logarithmic map, which itself is equivariant as per Corollary A.13, using the decomposition of the action in Lemma A.2. Using the fact that the scaling matrix is always orthogonal, we get the invariances with respect to symmetries and translations. For the permutation invariance, we just use the permutation invariance of the sum. $\square$

Corollary A.15. Given a fixed composition $\{n_s\}_{s \in \mathcal{S}}$, the canonical distribution associated to p_* is

$$\begin{aligned} p_*^n(ds, d\text{vol}_X) &= \frac{1}{\mathcal{Z}^n} \exp \left(-\frac{U(s, X)}{k_B T} \right) \\ &\times \prod_{s' \in \mathcal{S}} \delta_{\sum_{i=1}^N \delta_{s'}(s_i)}(n_{s'}) ds d\text{vol}_X. \end{aligned}$$

This density is G_c -invariant.

Proof. Use the G_c -invariance of p_* and the permutation invariance of the sum. $\square$

APPENDIX B: EQUIVARIANT RIEMANNIAN STOCHASTIC INTERPOLANTS

This appendix provides the formal theoretical guarantees for the Equivariant Riemannian Stochastic Interpolant (ERSI) framework introduced in Sec. III, establishing end-to-end equivariance from the underlying stochastic process down to the neural velocity field architecture. First, Proposition B.5 demonstrates that an equivariant interpolant paired with an invariant base distribution produces invariant marginal distributions across all $t \in [0, 1]$, while

Proposition B.7 proves that this construction yields an equivariant optimal velocity field. Proposition B.8 then verifies that our specific geodesic interpolation defined in Eq. (14) satisfies this condition. Next, we generalize the results of Ref. 14 to our non-linear symmetry group G_C in Propositions B.12 and B.14. Finally, Proposition B.15 proves by induction that our equivariant graph neural network architecture (Sec. III C) is G_C -equivariant at every layer, closing the gap between our theoretical framework and its practical implementation.

Lemma B.1. G_C is a group for the composition operation.

Proof. The identity is in G_C as it is a special case of the invariances. Let $g^1, g^2 \in G_C$; we can decompose them as $g^1 = f_L \circ h_{A^1, b^1}$ and $g^2 = f_L \circ h_{A^2, b^2}$, where A_1 (respectively, A^2) can be decomposed into A_σ^1, M^1 (respectively, A_σ^2, M^2) and b^1 (respectively, b^2) into $b^1 = \mathbf{1}_N \otimes c^1$ (respectively, $b^2 = \mathbf{1}_N \otimes c^2$). Using both Lemmas A.5 and A.3, $g^1 \circ g^2$ can be written for any $C \in \mathcal{C}$,

$$(g^1 \circ g^2)(C) = f_L \left[\begin{pmatrix} A_{\sigma^1} A_{\sigma^2} & 0 \\ 0 & (A_{\sigma^1} A_{\sigma^2} \otimes I_d)(I_N \otimes M^1 M^2) \end{pmatrix} C + (\mathbf{1}_N \otimes (c^1 + M^1 c^2)) \right],$$

which shows, by Lemma A.2, that $g^1 \circ g^2 \in G_C$. Finally, let $g \in G_C$ be written under the same decomposition; using Lemmas A.5 and A.3 again, we have for any $C \in \mathcal{C}$,

$$g^{-1}(C) = f_L \left[\begin{pmatrix} A_{\sigma^{-1}} & 0 \\ 0 & (I_N \otimes M^{-1})(A_{\sigma^{-1}} \otimes I_d) \end{pmatrix} C - (\mathbf{1}_N \otimes c) \right] = f_L \left[\begin{pmatrix} A_{\sigma^{-1}} & 0 \\ 0 & (A_{\sigma^{-1}} \otimes I_d)(I_N \otimes M^{-1}) \end{pmatrix} C + (\mathbf{1}_N \otimes (-c)) \right].$$

which shows, by Lemma A.2, that $g^{-1} \in G_C$. $\square$

Lemma B.2. For all $g = f_L \circ h_{A,b} \in G_C$, as in Eq. (A1), let J_g be the Jacobian of g . Then, $J_g(C) = A$, for almost every $C \in \mathcal{C}$, and $|J_g(C)| = 1$ almost everywhere.

Proof. Let $C \in \mathcal{C}$, then by the chain rule,

$$J_g(C) = J_{f_L}(h_{A,b}(C))J_{h_{A,b}}(C).$$

The modulo has unit derivative (except at jump points), so $J_{f_L} = I_{(N+1)d}$ almost everywhere. Since $h_{A,b}$ is affine, $J_{h_{A,b}} = A$, so $J_g(C) = A$ for almost every $C \in \mathcal{C}$. Moreover, $|J_g(C)| = |A|$. Using Eq. (A2), we can show that $A \in \mathcal{B}_{Nd}$, which implies that $|A| = 1$. $\square$

Lemma B.3. Let $g \in G_C$ and $C, D \in \mathcal{C}$, then

$$\int_{\mathcal{C}} \varphi(C) \delta_{g(D)}(g(C)) dC = \int_{\mathcal{C}} \varphi(C) \delta_D(C) dC$$

for all $\varphi \in C^\infty(\mathcal{C})$.

Proof. Make the change of variable $C = g^{-1}(U)$, since g is a group element,

$$U = g(C), \quad dC = |J_{g^{-1}}(U)| dU.$$

Using Lemma B.2, $|J_{g^{-1}}(U)| = 1$; thus,

$$\begin{aligned} \int_{\mathcal{C}} \varphi(C) \delta_{g(D)}(g(C)) dC &= \int_{\mathcal{C}} \varphi(g^{-1}(U)) \delta_{g(D)}(U) dU = \varphi(D) \\ &= \int_{\mathcal{C}} \varphi(C) \delta_D(C) dC. \end{aligned}$$

$\square$

Definition B.4. An interpolation function $I : [0, 1] \times \mathcal{C} \times \mathcal{C} \rightarrow \mathcal{C}$ is said to be G_C -equivariant if for all $g \in G_C$, $I(t, g(C_0), g(C_1)) = g(I(t, C_0, C_1))$ holds for any $t \in [0, 1]$ and $C_0, C_1 \in \mathcal{C}$.

Proposition B.5. Given a base p_{base} and the target $p_\star$ distribution both G_C -invariant and a G_C -equivariant interpolant I , then the induced marginal densities $(p_t)_{t=0}^1$ are all G_C -invariant.

Proof. The marginal densities are defined as

$$p_t(C) = \int_{\mathcal{C}} \delta_{I(t, C_0, C_1)}(C) p_{\text{base}}(C_0) p_\star(C_1) dC_0 dC_1.$$

Let $g \in G_C$, then

$$p_t(g(C)) = \int_{\mathcal{C}} \delta_{I(t, C_0, C_1)}(g(C)) p_{\text{base}}(C_0) p_\star(C_1) dC_0 dC_1.$$

Make the change of variable $C_0 = g(U_0)$ and $C_1 = g(U_1)$; using Lemma B.2, we get

$$\begin{aligned} p_t(g(C)) &= \int_{\mathcal{C}} \delta_{I(t, g(U_0), g(U_1))}(g(C)) p_{\text{base}}(g(U_0)) p_\star(g(U_1)) dU_0 dU_1 \\ &= \int_{\mathcal{C}} \delta_{I(t, g(U_0), g(U_1))}(g(C)) p_{\text{base}}(U_0) p_\star(U_1) dU_0 dU_1 \end{aligned} \quad (\text{B1})$$

$$= \int_{\mathcal{C}} \delta_{g(I(t, U_0, U_1))}(g(C)) p_{\text{base}}(U_0) p_\star(U_1) dU_0 dU_1 \quad (\text{B2})$$

$$= \int_{\mathcal{C}} \delta_{I(t, U_0, U_1)}(C) p_{\text{base}}(U_0) p_\star(U_1) dU_0 dU_1 = p_t(C). \quad (\text{B3})$$

We used the invariance of p_{base} and $p_\star$ to get (B1), the equivariance of I to go from (B1) to (B2), and Lemma B.3 to go from (B2) to (B3). $\square$

Lemma B.6. Let $g \in G_C$ and $I : [0, 1] \times \mathcal{C} \times \mathcal{C} \rightarrow \mathcal{C}$ be a differentiable, G_C -equivariant interpolation function. Then, the time derivative $\partial_t I$ is $\mathcal{C}$ -equivariant, i.e., it verifies for any $t \in [0, 1]$, $C_0, C_1 \in \mathcal{C}$,

$$\begin{aligned} \partial_t I(t, g(C_0), g(C_1)) &= A \partial_t I(t, C_0, C_1) \\ \text{for all } g \in G_C \text{ such that } g &= f_L \circ h_{A,b}. \end{aligned}$$

Proof. Let $C_0, C_1 \in \mathcal{C}$, then

$$\begin{aligned} \partial_t I(t, g(C_0), g(C_1)) &= \partial_t g(I(t, C_0, C_1)) && (\text{Equivariance of } I) \\ &= J_g(I(t, C_0, C_1)) \partial_t I(t, C_0, C_1) && (\text{Chain rule}) \\ &= A \partial_t I(t, C_0, C_1) && (\text{Lemma B.2}), \end{aligned}$$

where A is defined in Eq. (A1). $\square$

Proposition B.7. Given a G_C -equivariant interpolation function I , if p_{base} and p_* are G_C -invariant, then the corresponding optimal velocity field is G_C -equivariant.

Proof. The optimal velocity field is defined as

$$u^*(t, C) = \frac{1}{p_t(C)} \int_C \partial_t I(t, C_0, C_1) \delta_{I(t, C_0, C_1)}(C) p_{\text{base}}(C_0) p_*(C_1) dC_0 dC_1.$$

Since p_t is G_C -invariant (see Proposition B.5), we need to show that the integral

$$\mathcal{I}(C) = \int_C \partial_t I(t, C_0, C_1) \delta_{I(t, C_0, C_1)}(C) p_{\text{base}}(C_0) p_*(C_1) dC_0 dC_1$$

is a G_C -equivariant vector field. Let $g \in G_C$,

$$\mathcal{I}(g(C)) = \int_C \partial_t I(t, C_0, C_1) \delta_{I(t, C_0, C_1)}(g(C)) p_{\text{base}}(C_0) p_*(C_1) dC_0 dC_1.$$

Making the change of variable $C_0 = g(U_0)$ and $C_1 = g(U_1)$, using Lemma B.2,

$$\mathcal{I}(g(C)) = \int_C \partial_t I(t, g(U_0), g(U_1)) \delta_{I(t, g(U_0), g(U_1))}(g(C)) p_{\text{base}}(U_0) p_*(U_1) dU_0 dU_1$$

$$= \int_C \partial_t I(t, g(U_0), g(U_1)) \delta_{I(t, g(U_0), g(U_1))}(g(C)) p_{\text{base}}(U_0) p_*(U_1) dU_0 dU_1 \quad (\text{B4})$$

$$= \int_C A \partial_t I(t, U_0, U_1) \delta_{g(I(t, U_0, U_1))}(g(C)) p_{\text{base}}(U_0) p_*(U_1) dU_0 dU_1 \quad (\text{B5})$$

$$= \int_C A \partial_t I(t, U_0, U_1) \delta_{I(t, U_0, U_1)}(C) p_{\text{base}}(U_0) p_*(U_1) dU_0 dU_1 = A \mathcal{I}(C), \quad (\text{B6})$$

where we use the invariance of p_{base} and p_* in (B4), then the equivariance of $\partial_t I$ due to Lemma B.6 to go from (B4) to (B5) and Lemma B.3 to go from (B5) to (B6). $\square$

Proposition B.8. The interpolant in Eq. (14) is G_C -equivariant.

Proof. Let $g \in G_C$. Consider the decomposition of g introduced in Lemma A.2. If we denote

$$I_L : (t, C_0, C_1) \mapsto \begin{pmatrix} I_L^S(t, C_0, C_1) \\ I_L^X(t, C_0, C_1) \end{pmatrix},$$

then

$$I_L^S(t, g(C_0), g(C_1)) = (1-t)A_\sigma s_0 + tA_\sigma s_1 = A_\sigma I_L^S(t, C_0, C_1).$$

Moreover, for every $i \in \llbracket 1, N \rrbracket$, by definition,

$$\begin{aligned} I_L^X(t, g(C_0), g(C_1))_{(i)} &= \exp_{\tilde{h}_{M, \epsilon, L}((X_0)_{(\sigma(i))})} \left(t \log_{\tilde{h}_{M, \epsilon, L}((X_0)_{(\sigma(i))})} \tilde{h}_{M, \epsilon, L}((X_1)_{(\sigma(i))}) \right) = \exp_{\tilde{h}_{M, \epsilon, L}((X_0)_{(\sigma(i))})} \left(t M \log_{(X_0)_{(\sigma(i))}} (X_1)_{(\sigma(i))} \right) \\ &= \left(M \exp_{(X_0)_{(\sigma(i))}} \left(t \log_{(X_0)_{(\sigma(i))}} (X_1)_{(\sigma(i))} \right) + c \right) \% L = \tilde{h}_{M, \epsilon, L}(I^X(t, C_0, C_1))_{(\sigma(i))}, \end{aligned}$$

where we successively used Corollary A.13 and Proposition A.11, which concludes the proof. $\square$

We consider a velocity field $\hat{v}$ on $\mathcal{C}$ and denote by $\hat{T}_t$ the map transporting an initial configuration of particles along the velocity field between time 0 and time t . In other words, $\hat{T}_t(c_0)$ gives the solution of the ODE $dC_u = \hat{v}(u, C_u) du$ at time t for the initial condition $C_0 = c_0$. Given a base distribution p_{base} , we seek to learn $\hat{v}$ such that the push-forward of p_{base} through the transport map $\hat{T}_t$, which is the marginal distribution of the RSI model denoted by $\hat{q}_t$ below, is G_C -invariant. For this, we rely on Proposition B.12, extending the result of Ref. 14, Theorem 1 to the invariance group G_C .

Definition B.9. A diffeomorphism $T : \mathcal{C} \rightarrow \mathcal{C}$ is G_C -equivariant if for all $g \in G_C$, $T \circ g = g \circ T$ holds.

Definition B.10. A diffeomorphism $T : \mathcal{C} \rightarrow \mathcal{C}$ is G_C -equivariant if for all $g \in G_C$, $T \circ g = g \circ T$ holds.

Lemma B.11. Let $g \in G_C$ and $T : \mathcal{C} \rightarrow \mathcal{C}$ be a diffeomorphism. T is G_C -equivariant if and only if T^{-1} is G_C -equivariant.

Proof. Let $Y \in \mathcal{C}$ and set $X = T(Y)$, then

$$\begin{aligned} T^{-1}(g(X)) &= g(T^{-1}(X)) \iff T^{-1}(g(T(Y))) = g(Y) \\ &\iff g(T(Y)) = T(g(Y)). \end{aligned}$$

The following proposition generalizes (Ref. 14, Theorem 1) to G_C , which is nonlinear. $\square$

Proposition B.12. Let $q : \mathcal{C} \rightarrow \mathbb{R}$ be a density of $\mathcal{C}$, and let $T : \mathcal{C} \rightarrow \mathcal{C}$ be a diffeomorphism. If q is G_C -invariant and T is a G_C -equivariant, then the push-forward of q through T is also G_C -invariant.

Proof. Let $C \in \mathcal{C}$, the push-forward of q through T writes as

$$p(C) = q(T^{-1}(C)) |J_{T^{-1}}(C)|.$$

Thus,

$$p(g(C)) = q(T^{-1}(g(C))) |J_{T^{-1}}(g(C))|.$$

By Lemma B.11 and the invariance of q , we have

$$\begin{aligned} q(T^{-1}(g(C))) &= q(g(T^{-1}(C))) \\ &= q(T^{-1}(C)). \end{aligned}$$

Moreover,

$$\begin{aligned} |J_{T^{-1}}(g(C))| &= \frac{|J_{T^{-1} \circ g}(C)|}{|J_g(C)|} && \text{(Chain rule)} \\ &= |J_{T^{-1} \circ g}(C)| && \text{(Lemma B.2)} \\ &= |J_{g \circ T^{-1}}(C)| && \text{(Lemma B.11)} \\ &= |J_g(T^{-1}(C))| |J_{T^{-1}}(C)| && \text{(Chain rule)} \\ &= |J_{T^{-1}}(C)| && \text{(Lemma B.2)}. \end{aligned}$$

Combining these results, we get $p(g(C)) = p(C)$. $\square$

Definition B.13. A velocity field $\hat{v} : [0, 1] \times \mathcal{C} \rightarrow \mathcal{TC}$ is G_C -equivariant if, for all $g \in G_C$, using the decomposition $g = f_L \circ h_{A,b}$ (with A and b defined in Lemma A.2),

$$\hat{v}(t, g(C)) = A \hat{v}(t, C), \quad \forall C \in \mathcal{C}.$$

$$\begin{aligned} T(g(Z)) &= f_L \left(g(Z) + \int_0^t v(u, C(t, g(Z))) du \right) \\ &= f_L \left(g(Z) + \int_0^t v(u, g(C(t, Z))) du \right) && \text{(Using } C(\cdot, g(Z)) = g(C(\cdot, Z)) \text{)} \\ &= f_L \left(f_L(h_{A,b}(Z)) + A \int_0^t v(u, C(t, Z)) du \right) && \text{(Using the Equivariance of } v \text{)} \\ &= f_L \left(h_{A,b}(Z) + A \int_0^t v(u, C(t, Z)) du \right) && \text{(Using Lemma A.3)} \\ &= f_L \left(h_{A,b} \left(Z + \int_0^t v(u, C(t, Z)) du \right) \right) = g(T(Z)). \end{aligned}$$

$\square$

Proposition B.15. The velocity field in Eq. (15) is G_C -equivariant.

Proof. Let $t \in [0, 1]$ and $C \in \mathcal{C}$. We define

$$\begin{aligned} \psi_k(t, C) &= (\psi_k^s(t, C), \psi_k^x(t, C)) = (s, X^k), \\ \Gamma_k(t, C) &= H^k, \quad \text{and} \quad \Lambda_k(t, C) = M^k. \end{aligned}$$

We start by showing by induction on $k \in \mathbb{N}$ that for all $g \in G_C$ and $i, j \in \llbracket 1, N \rrbracket$, we have

$$\begin{aligned} \psi_k(t, g(C)) &= g(\psi_k(t, C)), \quad [\Gamma_k(t, g(C))]_i = [\Gamma_k(t, C)]_{\sigma(i)}, \quad \text{and} \\ [\Lambda_k(t, g(C))]_{i,j} &= [\Lambda_k(t, C)]_{\sigma(i), \sigma(j)}, \end{aligned}$$

where, building on Eq. (A1), we decompose g as

$$\begin{aligned} g : \begin{pmatrix} s \\ X \end{pmatrix} &\mapsto \begin{pmatrix} g^s(C) \\ g^x(C) \end{pmatrix} = \begin{pmatrix} A_\sigma s \\ (A_\sigma \otimes I_d) h_{M,b,L}(X) \end{pmatrix}, \\ h_{M,b,L}(X) &= ((I_N \otimes M)X + (\mathbf{1}_N \otimes b)) \% L, \end{aligned}$$

Proposition B.14. If $v : [0, 1] \times \mathcal{C} \rightarrow \mathcal{TC}$ is a Lipschitz-bounded G_C -equivariant velocity field, then the transport map induced by its ODE flow is G_C -equivariant.

Proof. Let $g \in G_C$ be decomposed (as per Lemma A.2) as $g = f_L \circ h_{A,b}$. Let $Z \in \mathcal{C}$. Consider that $C(\cdot, Z)$ denotes the solution of the ODE with initial condition $Z \in \mathcal{C}$. Let $\tilde{C}(\cdot, Z) = C(\cdot, g(Z))$. Using the chain rule and the equivariance of v , we have almost everywhere that

$$\begin{aligned} \frac{d\tilde{C}(t, Z)}{dt} &= A \frac{dC(t, Z)}{dt} = Av(t, C(t, Z)) \\ &= v(t, g(C(t, Z))) = v(t, \tilde{C}(t, Z)). \end{aligned}$$

Moreover, $\tilde{C}(0, Z) = g(Z)$. Using the unicity of the ODE's solutions (due to the velocity field being Lipschitz-bounded), we get

$$C(t, g(Z)) = g(C(t, Z)), \quad \text{almost everywhere, for all } t \in [0, 1].$$

Using this in the integral definition of the transport map, for all $t \in [0, 1]$, we get

where $h_{M,b,L}$ describes the group of invariances restricted to $\mathcal{M}^N$ (denoted $\mathcal{G}_{\mathcal{M}^N}$), $\sigma \in S_N$ is a permutation, A_σ is the associated permutation matrix of size $N \times N$, M is a signed permutation matrix of size $d \times d$, and b is a vector in $\mathbb{R}^d$.

At $k = 0$, we have $\psi_0(t, g(C)) = g(C) = g(\psi_0(t, C))$, and for all $i \in \llbracket 1, N \rrbracket$,

$$[\Gamma_0(t, g(C))]_i = (t, [g^s(C)]_i) = (t, s_{\sigma(i)}) = [\Gamma_0(t, C)]_{\sigma(i)}.$$

Moreover, for all $i, j \in \llbracket 1, N \rrbracket$,

$$\begin{aligned} [\Lambda_0(t, g(C))]_{i,j} &= \hat{\phi}_e \left([\Gamma_0(t, g(C))]_i, [\Gamma_0(t, g(C))]_j \right) \\ &\quad \times d_{\mathcal{M}}(h_{M,b,L}(X)_{(\sigma(i))}, h_{M,b,L}(X)_{(\sigma(j))}). \end{aligned}$$

Using the previous statement on Γ_0 as well as the invariance of $d_{\mathcal{M}}$ to $\mathcal{G}_{\mathcal{M}}$ (see Corollary A.14),

$$[\Lambda_0(t, g(C))]_{ij} = \hat{\phi}_e([\Gamma_0(t, C)]_{\sigma(i)}, [\Gamma_0(t, C)]_{\sigma(j)}, d_M(X_{\sigma(i)}, X_{\sigma(j)})), \quad \psi_k(t, g(C)) = g(\psi_k(t, C)), \quad [\Gamma_k(t, g(C))]_i = [\Gamma_k(t, C)]_{\sigma(i)}, \text{ and} \\ = [\Lambda_k(t, g(C))]_{\sigma(i), \sigma(j)}. \quad [\Lambda_k(t, g(C))]_{ij} = [\Lambda_k(t, C)]_{\sigma(i), \sigma(j)},$$

Let $k \in \mathbb{N}$ and assume that

Let $i, j \in \llbracket 1, N \rrbracket$. Using the recursion assumption and the permutation invariance of the sum, we have that

$$\begin{aligned} [\Gamma_{k+1}(t, g(C))]_i &= \hat{\phi}_h\left([\Gamma_k(t, g(C))]_i, \sum_{i \neq j} \hat{\phi}_m([\psi_k(t, g(C))]_{i,j})[\psi_k(t, g(C))]_{i,j}\right) \\ &= \hat{\phi}_h\left([\Gamma_k(t, C)]_{\sigma(i)}, \sum_{i \neq j} \hat{\phi}_m([\psi_k(t, C)]_{\sigma(i), \sigma(j)})[\psi_k(t, C)]_{\sigma(i), \sigma(j)}\right) \\ &= \hat{\phi}_h\left([\Gamma_k(t, C)]_{\sigma(i)}, \sum_{i \neq j} \hat{\phi}_m([\psi_k(t, C)]_{\sigma(i), j})[\psi_k(t, C)]_{\sigma(i), j}\right) = [\Gamma_{k+1}(t, C)]_{\sigma(i)}. \end{aligned}$$

Similarly, using the same argument as in $k = 0$, it is easy to show that

$$[\Lambda_{k+1}(t, g(C))]_{ij} = [\Lambda_k(t, C)]_{\sigma(i), \sigma(j)}.$$

Additionally, we have

$$\psi_{k+1}^s(t, g(C)) = g^s(C) = g^s(\psi_k^s(t, C)).$$

Moreover, using the definition and the previous statement on Λ_k ,

$$\begin{aligned} \psi_{k+1}^X(t, g(C))_{(i)} &= \exp_{\psi_k^X(t, g(C))_{(i)}}\left(\sum_{i \neq j} \frac{\log_{\psi_k^X(t, g(C))_{(j)}} \psi_k^X(t, g(C))_{(i)}}{d_M(g^X(C)_{(i)}, g^X(C)_{(j)}) + 1} \phi_d([\Lambda_k(t, g(C))]_{i,j})\right) \\ &= \exp_{\psi_k^X(t, g(C))_{(i)}}\left(\sum_{i \neq j} \frac{\log_{\psi_k^X(t, g(C))_{(j)}} \psi_k^X(t, g(C))_{(i)}}{d_M(g^X(C)_{(i)}, g^X(C)_{(j)}) + 1} \phi_d([\Lambda_k(t, C)]_{\sigma(i), \sigma(j)})\right). \end{aligned}$$

Using the recursion assumption together with Corollary A.13, we have

$$\begin{aligned} \log_{\psi_k^X(t, g(C))_{(j)}} \psi_k^X(t, g(C))_{(i)} &= \log_{g^X(\psi_k(t, C))_{(j)}} g^X(\psi_k(t, C))_{(i)} \\ &= \log_{h_{M,b,L}(\psi_k^X(t, C))_{(\sigma(j))}} h_{M,b,L}(\psi_k^X(t, C))_{(\sigma(i))} \\ &= M \log_{\psi_k^X(t, C)_{(\sigma(j))}} \psi_k^X(t, C)_{(\sigma(i))}. \end{aligned}$$

Using Proposition A.11 and the recursion assumption again, we get that

$$\psi_{k+1}^X(t, g(C))_{(i)} = \left(M \exp_{\psi_k^X(t, C)_{(\sigma(i))}} \sum_{i \neq j} \frac{\log_{\psi_k^X(t, C)_{(\sigma(j))}} \psi_k^X(t, C)_{(\sigma(i))}}{d_M(g^X(C)_{(i)}, g^X(C)_{(j)}) + 1} \phi_d([\Lambda_k(t, C)]_{\sigma(i), \sigma(j)}) + b \right) \% L.$$

Together with the invariance of d_M and the permutation invariance of the sum, it leads to

$$\begin{aligned} \psi_{k+1}^X(t, g(C))_{(i)} &= \left(M \exp_{\psi_k^X(t, C)_{(\sigma(i))}} \sum_{i \neq j} \frac{\log_{\psi_k^X(t, C)_{(j)}} \psi_k^X(t, C)_{(\sigma(i))}}{d_M(X_{\sigma(i)}, X_{\sigma(j)}) + 1} \phi_d([\Lambda_k(t, C)]_{\sigma(i), j}) + b \right) \% L \\ &= \left[(M \psi_{k+1}^X(t, g(C))_{(i)} + b) \% L \right]_{(\sigma(i))} = g^X(\psi_{k+1}^X(t, C))_{(i)}, \end{aligned}$$

which concludes the recursive proof. We get the overall proof using Corollary A.13 as

TABLE I. Hyperparameter settings, network architectures (K | HF), batch sizes, training steps, learning rates, and gradient clipping thresholds across all studied glass models, state points, and baseline methods.

System	N	T	Model	Architecture (K HF)	Batch size	Training steps	Learning rate	Gradient clipping
Binary mixture	10	0.10	ERSI	EGNN 3 32	2048	5×10^4	1×10^{-4}	2.0
	10	0.10	EFM	EGNN 3 32	2048	5×10^4	5×10^{-3}	2.0
	10	0.10	RSI	MLP	2048	5×10^4	5×10^{-4}	2.0
	44	0.10	ERSI	EGNN 3 32	1024	10^5	5×10^{-3}	2.0
	44	0.10	EFM	EGNN 3 32	1024	10^5	1×10^{-3}	2.0
	44	0.10	RSI	MLP	1024	10^5	5×10^{-5}	2.0
Ternary mixture	44	1.00	ERSI	EGNN 3 32	1024	5×10^4	5×10^{-3}	None
	44	1.00	ERSI	EGNN 4 64	1024	5×10^4	5×10^{-3}	None
	44	1.00	ERSI	EGNN 5 128	1024	5×10^4	1×10^{-3}	None
	44	0.32	ERSI	EGNN 10 128	512	2.5×10^4	1×10^{-4}	None

$$\begin{aligned}
\hat{v}(t, g(C)) &= \left(\begin{matrix} \mathbf{0}_{N \times d_s} \\ \left\{ \log_{g^X(X)_{(i)}} \psi_K^X(t, g(C))_{(i)} \right\}_{i=1}^N \end{matrix} \right) \\
&= \left(\begin{matrix} \mathbf{0}_{N \times d_s} \\ \left\{ \log_{g^X(X)_{(i)}} g^X(\psi_K^X(t, C))_{(i)} \right\}_{i=1}^N \end{matrix} \right) && \text{(Recursion)} \\
&= \left(\begin{matrix} \mathbf{0}_{N \times d_s} \\ \left\{ M \log_{X(\sigma(i))} \psi_K^X(t, C)_{(\sigma(i))} \right\}_{i=1}^N \end{matrix} \right) && \text{(Corollary A.13)} \\
&= A \hat{v}(t, C).
\end{aligned}$$

□

APPENDIX C: NUMERICAL DETAILS

In this appendix, we report the implementation, architectural, and training details for all numerical simulations on the physical systems defined in Sec. IV. For the Equivariant Riemannian Stochastic Interpolant (ERSI), the velocity field is parameterized via our Equivariant Graph Neural Network (EGNN) architecture [Eq. (15)]. In the base $K = 3$, HF = 32 network (containing $\approx 22k$ parameters), particle species are embedded via one-hot encoding, $\hat{\phi}_e$ and $\hat{\phi}_h$ are three-layer MLPs of width 32, and $\hat{\phi}_m$ is a two-layer MLP of width 32. For scaled ERSI architectures (4 | 64, 5 | 128, and 10 | 128), the MLP hidden widths are set equal to HF, and the number of message-passing layers is set to K . For the EFM baseline, we adopt the architecture of Ref. 31 with standard Euclidean operators. For the non-equivariant RSI baseline, particle configurations are processed directly by a three-layer MLP with 64 hidden units per layer, projecting coordinates onto the torus following Ref. 28 [Eq. (26)]; testing explicit data augmentation via random group actions of G_c during training yielded no noticeable performance recovery. All models were trained on 10^5 equilibrium configurations using the Adam optimizer. Hyperparameters—including learning rates, batch sizes, and gradient clipping settings—were selected via validation loss and are summarized in Table I. Generative sampling and exact

likelihood evaluations [Eq. (6)] are performed using the adaptive `dopri5` Runge–Kutta solver from `torchdiffeq`⁷⁶ with absolute and relative integration tolerances set to 10^{-5} .

REFERENCES

- J. S. Liu and J. S. Liu, *Monte Carlo Strategies in Scientific Computing* (Springer, 2001), Vol. 10.
- T. Lelièvre, M. Rousset, and G. Stoltz, *Free Energy Computation: A Mathematical Perspective* (Imperial College Press, 2010).
- D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Elsevier, 2023).
- K. Ohno, K. Esfarjani, and Y. Kawazoe, *Computational Materials Science: From Ab Initio to Monte Carlo Methods* (Springer, 2018).
- M. Parrinello and A. Rahman, “Crystal structure and pair potentials: A molecular-dynamics study,” *Phys. Rev. Lett.* **45**, 1196 (1980).
- M. Matsumoto, S. Saito, and I. Ohmine, “Molecular dynamics simulation of the ice nucleation and growth process leading to water freezing,” *Nature* **416**, 409 (2002).
- F. Noé, C. Schütte, E. Vanden-Eijnden, L. Reich, and T. R. Weikel, “Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations,” *Proc. Natl. Acad. Sci. U. S. A.* **106**, 19011 (2009).
- K. Lindorff-Larsen, S. Piana, R. O. Dror, and D. E. Shaw, “How fast-folding proteins fold,” *Science* **334**, 517 (2011).

- ⁹I. Buch, T. Giorgino, and G. De Fabritiis, "Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations," *Proc. Natl. Acad. Sci. U. S. A.* **108**, 10184 (2011).
- ¹⁰V. L. Deringer, "Modelling and understanding battery materials with machine-learning-driven atomistic simulations," *J. Phys.: Energy* **2**, 041003 (2020).
- ¹¹F. Noé, S. Olsson, J. Köhler, and H. Wu, "Boltzmann generators: Sampling equilibrium states of many-body systems with deep learning," *Science* **365**, eaaw1147 (2019).
- ¹²M. Albergo, G. Kanwar, and P. Shanahan, "Flow-based generative models for Markov chain Monte Carlo in lattice field theory," *Phys. Rev. D* **100**, 034515 (2019).
- ¹³D. Wu, L. Wang, and P. Zhang, "Solving statistical mechanics using variational autoregressive networks," *Phys. Rev. Lett.* **122**, 080602 (2019).
- ¹⁴J. Köhler, L. Klein, and F. Noé, "Equivariant flows: Exact likelihood generative learning for symmetric densities," in *Proceedings of the 37th International Conference on Machine Learning, Proceedings of the Machine Learning Research Vol. 119*, edited by H. Daumé III and A. Singh (PMLR, 2020), pp. 5361–5370.
- ¹⁵L. Klein, A. Krämer, and F. Noé, "Equivariant flow matching," in *Advances in Neural Information Processing Systems* edited by A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine (Curran Associates, Inc., 2023), Vol. 36, pp. 59886–59910.
- ¹⁶D. Rehman, T. Akhound-Sadegh, A. Gazizov, Y. Bengio, and A. Tong, "FALCON: Few-step accurate likelihoods for continuous flows," in *International Conference on Learning Representations (ICLR, 2026)*, Vol. 2026, pp. 70449–70473.
- ¹⁷M. Schebek, J. He, E. Hoffmann, Y. Du, F. Noé, and J. Rogal, "Assessing generative modeling approaches for free energy estimates in condensed matter," *J. Chem. Phys.* **164**, 184111 (2026).
- ¹⁸L. Berthier and G. Biroli, "Theoretical perspective on the glass transition and amorphous materials," *Rev. Mod. Phys.* **83**, 587–645 (2011).
- ¹⁹L. Berthier and D. R. Reichman, "Modern computational studies of the glass transition," *Nat. Rev. Phys.* **5**, 102–116 (2023).
- ²⁰A. Ninarello, L. Berthier, and D. Coslovich, "Models and algorithms for the next generation of glass transition studies," *Phys. Rev. X* **7**, 021039 (2017).
- ²¹G. Jung, G. Biroli, and L. Berthier, "Normalizing flows as an enhanced sampling method for atomistic supercooled liquids," *Mach. Learn.: Sci. Technol.* **5**, 035053 (2024).
- ²²K. Yang and D. Schwalbe-Koda, "A generative diffusion model for amorphous materials," *npj Comput. Mater.* **12**, 29 (2025).
- ²³M. S. Albergo and E. Vanden-Eijnden, "Building normalizing flows with stochastic interpolants," in *The Eleventh International Conference on Learning Representations*, 2023.
- ²⁴M. Albergo, N. M. Boffi, and E. Vanden-Eijnden, "Stochastic interpolants: A unifying framework for flows and diffusions," *J. Mach. Learn. Res.* **26**, 1 (2025).
- ²⁵Y. Lipman, R. T. Q. Chen, H. Ben-Hamu, M. Nickel, and M. Le, "Flow matching for generative modeling," in *The Eleventh International Conference on Learning Representations*, 2023.
- ²⁶V. Garcia Satorras, E. Hoogeboom, F. Fuchs, I. Posner, and M. Welling, "E(n) equivariant normalizing flows," in *Advances in Neural Information Processing Systems*, edited by M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan (Curran Associates, Inc., 2021), Vol. 34, pp. 4181–4192.
- ²⁷R. T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. Duvenaud, "Neural ordinary differential equations," *Adv. Neural Inf. Proc. Syst.* **31**, 6571–6583 (2018).
- ²⁸R. T. Q. Chen and Y. Lipman, "Flow matching on general geometries," in *The Twelfth International Conference on Learning Representations*, 2024.
- ²⁹J. Wu, B. Chen, Y. Zhou, Q. Meng, R. Zhu, and Z.-M. Ma, "Riemannian neural geodesic interpolant," *arXiv:2504.15736* (2025).
- ³⁰B. K. Miller, R. T. Q. Chen, A. Sriram, and B. M. Wood, "FlowMM: Generating materials with Riemannian flow matching," in *Proceedings of the 41st International Conference on Machine Learning, Proceedings of Machine Learning Research, Vol. 235*, edited by R. Salakhutdinov, Z. Kolter, K. Heller, A. Weller, N. Oliver, J. Scarlett, and F. Berkenkamp PMLR, 2024, pp. 35664–35686.
- ³¹V. G. Satorras, E. Hoogeboom, and M. Welling, "E(n) equivariant graph neural networks," in *Proceedings of the 38th International Conference on Machine Learning, Proceedings of Machine Learning Research Vol. 139*, edited by M. Meila and T. Zhang (PMLR, 2021), pp. 9323–9332.
- ³²J. Köhler, L. Klein, and F. Noé, "Equivariant flows: Sampling configurations for multi-body systems with symmetric energies," *arXiv:1910.00753* (2019).
- ³³P. Wirnsberger, G. Papamakarios, B. Ibarz, S. Racanière, A. J. Ballard, A. Pritzel, and C. Blundell, "Normalizing flows for atomic solids," *Mach. Learn.: Sci. Technol.* **3**, 025009 (2022).
- ³⁴L. Midgley, V. Stimper, J. Antorán, E. Mathieu, B. Schölkopf, and J. M. Hernández-Lobato, "SE(3) Equivariant Augmented Coupling Flows," in *Advances in Neural Information Processing Systems*, edited by A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine (Curran Associates, Inc., 2023), Vol. 36, pp. 79200–79225.
- ³⁵M. Biloš and S. Günnemann, "Scalable normalizing flows for permutation invariant densities," in *Proceedings of the 38th International Conference on Machine Learning, Proceedings of Machine Learning Research Vol. 139* edited by M. Meila and T. Zhang (PMLR, 2021), pp. 957–967.
- ³⁶M. Xu, L. Yu, Y. Song, C. Shi, S. Ermon, and J. Tang, "GeoDiff: A geometric diffusion model for molecular conformation generation," in *International Conference on Learning Representations*, 2022.
- ³⁷E. Hoogeboom, V. G. Satorras, C. Vignac, and M. Welling, "Equivariant diffusion for molecule generation in 3D," in *Proceedings of the 39th International Conference on Machine Learning, Proceedings of Machine Learning Research Vol. 162*, edited by K. Chaudhuri, S. Jegelka, L. Song, C. Szepesvari, G. Niu, and S. Sabato (PMLR, 2022), pp. 8867–8887.
- ³⁸K. Martinkus, J. Ludwiczak, W.-C. Liang, J. Lafrance-Vanashe, I. Hotzel, A. Rajpal, Y. Wu, K. Cho, R. Bonneau, V. Gligorijevic, and A. Loukas, "AbDiffuser: Full-atom generation of in-vitro functioning antibodies," in *Thirty-Seventh Conference on Neural Information Processing Systems*, 2023.
- ³⁹A. E. Chu, J. Kim, L. Cheng, G. El Nesr, M. Xu, R. W. Shuai, and P.-S. Huang, "An all-atom protein generative model," *Proc. Natl. Acad. Sci. U. S. A.* **121**, e2311500121 (2024).
- ⁴⁰C. K. Joshi, X. Fu, Y.-L. Liao, V. Gharakhanyan, B. K. Miller, A. Sriram, and Z. W. Ulissi, "All-atom diffusion transformers: Unified generative modelling of molecules and materials," in *Forty-Second International Conference on Machine Learning*, 2025.
- ⁴¹S. Yang, K. Cho, A. Merchant, P. Abbeel, D. Schuurmans, I. Mordatch, and E. Cubuk, "Scalable diffusion for materials generation," in *NeurIPS 2023 AI for Science Workshop*, 2023.
- ⁴²R. Jiao, W. Huang, P. Lin, J. Han, P. Chen, Y. Lu, and Y. Liu, "Crystal structure prediction by joint equivariant diffusion," in *Thirty-Seventh Conference on Neural Information Processing Systems*, 2023.
- ⁴³R. Jiao, W. Huang, Y. Liu, D. Zhao, and Y. Liu, "Space group constrained crystal generation," in *The Twelfth International Conference on Learning Representations*, 2024.
- ⁴⁴C. Zeni, R. Pinsler, D. Zügner, A. Fowler, M. Horton, X. Fu, Z. Wang, A. Shysheya, J. Crabbé, S. Ueda, R. Sordillo, L. Sun, J. Smith, B. Nguyen, H. Schulz, S. Lewis, C.-W. Huang, Z. Lu, Y. Zhou, H. Yang, H. Hao, J. Li, C. Yang, W. Li, R. Tomioka, and T. Xie, "A generative model for inorganic materials design," *Nature* **639**, 624 (2025).
- ⁴⁵D. Levy, S. S. Panigrahi, S.-O. Kaba, Q. Zhu, K. L. K. Lee, M. Galkin, S. Miret, and S. Ravanbakhsh, "SymmCD: Symmetry-preserving crystal generation with diffusion models," in *The Thirteenth International Conference on Learning Representations*, 2025.
- ⁴⁶A. Sriram, B. K. Miller, R. T. Q. Chen, and B. M. Wood, "FlowLLM: Flow matching for material generation with large language models as base distributions," in *Advances in Neural Information Processing Systems*, edited by A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang (Curran Associates, Inc., 2024), Vol. 37, pp. 46025–46046.
- ⁴⁷K. A. Nicoli, S. Nakajima, N. Strodthoff, W. Samek, K.-R. Müller, and P. Kessel, "Asymptotically unbiased estimation of physical observables with neural samplers," *Phys. Rev. E* **101**, 023304 (2020); *arXiv:1910.13496*.
- ⁴⁸K. A. Nicoli, C. J. Anders, L. Funcke, T. Hartung, K. Jansen, P. Kessel, S. Nakajima, and P. Stornati, "Estimation of thermodynamic observables in lattice field theories with deep generative models," *Phys. Rev. Lett.* **126**, 032001 (2021).
- ⁴⁹M. Gabrié, G. M. Rotskoff, and E. Vanden-Eijnden, "Adaptive Monte Carlo augmented with normalizing flows," *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2109420119 (2022).

- ⁵⁰W. Grathwohl, R. T. Q. Chen, J. Bettencourt, and D. Duvenaud, “Scalable reversible generative models with free-form continuous dynamics,” in International Conference on Learning Representations, 2019.
- ⁵¹A. Gholaminejad, K. Keutzer, and G. Biroli, “ANODE: Unconditionally accurate memory-efficient gradients for neural ODEs,” in International Joint Conferences on Artificial Intelligence, 2019.
- ⁵²X. Liu, C. Gong, and Q. Liu, “Flow straight and fast: Learning to generate and transfer data with rectified flow,” in The Eleventh International Conference on Learning Representations, 2023.
- ⁵³B. Bernu, J. P. Hansen, Y. Hiwatari, and G. Pastore, “Soft-sphere model for the glass transition in binary alloys: Pair structure and self-diffusion,” *Phys. Rev. A* **36**, 4891 (1987).
- ⁵⁴D. N. Perera and P. Harrowell, “Stability and structure of a supercooled liquid mixture in two dimensions,” *Phys. Rev. E* **59**, 5721 (1999).
- ⁵⁵G. Jung, G. Biroli, and L. Berthier, “Predicting dynamic heterogeneity in glass-forming liquids by physics-inspired machine learning,” *Phys. Rev. Lett.* **130**, 238202 (2023).
- ⁵⁶G. Jung, M. Ozawa, G. Biroli, and L. Berthier, “Numerical investigation of the equilibrium Kauzmann transition in a two-dimensional atomistic glass,” *Proc. Natl. Acad. Sci. U. S. A.* **123**, e2612355123 (2026).
- ⁵⁷W. Kob and H. C. Andersen, “Testing mode-coupling theory for a supercooled binary Lennard-Jones mixture I: The van Hove correlation function,” *Phys. Rev. E* **51**, 4626 (1995).
- ⁵⁸N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, “Equation of state calculations by fast computing machines,” *J. Chem. Phys.* **21**, 1087 (1953).
- ⁵⁹D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed., *Computational Science Series Vol. 1* (Academic Press, San Diego, 2002).
- ⁶⁰L. Grenioux, L. Galliano, L. Berthier, G. Biroli, and M. Gabrié (2026). “Dataset for ‘Boltzmann generators for amorphous particle systems,’” Zenodo. <https://doi.org/10.5281/zenodo.23388791>
- ⁶¹A. Tong, K. Fatras, N. Malkin, G. Huguet, Y. Zhang, J. Rector-Brooks, G. Wolf, and Y. Bengio, “Improving and generalizing flow-based generative models with minibatch optimal transport,” *Trans. Mach. Learn. Res.* (published online) (2024), available online at <https://openreview.net/forum?id=CD9Snc73AW>.
- ⁶²M. S. Albergo, M. Goldstein, N. M. Boffi, R. Ranganath, and E. Vanden-Eijnden, “Stochastic interpolants with data-dependent couplings,” in *Proceedings of the 41st International Conference on Machine Learning, Proceedings of Machine Learning Research Vol. 235*, edited by R. Salakhutdinov, Z. Kolter, K. Heller, A. Weller, N. Oliver, J. Scarlett, and F. Berkenkamp (PMLR, 2024), pp. 921–937.
- ⁶³K. Fatras, T. Sejourne, R. Flamary, and N. Courty, “Unbalanced minibatch optimal transport; applications to domain adaptation,” in *Proceedings of the 38th International Conference on Machine Learning, Proceedings of Machine Learning Research Vol. 139*, edited by M. Meila and T. Zhang (PMLR, 2021), pp. 3186–3197.
- ⁶⁴Y. Song, J. Gong, M. Xu, Z. Cao, Y. Lan, S. Ermon, H. Zhou, and W.-Y. Ma, “Equivariant flow matching with hybrid probability transport for 3D molecule generation,” in *Advances in Neural Information Processing Systems*, edited by A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine (Curran Associates, Inc., 2023), Vol. 36, pp. 549–568.
- ⁶⁵R. Irwin, A. Tibo, J. P. Janet, and S. Olsson, “Efficient 3D molecular generation with flow matching and scale optimal transport,” in ICML 2024 AI for Science Workshop, 2024.
- ⁶⁶E. Flenner and G. Szamel, “Hybrid Monte Carlo simulation of a glass-forming binary mixture,” *Phys. Rev. E* **73**, 061505 (2006).
- ⁶⁷J.-L. Barrat and J.-P. Hansen, *Basic Concepts for Simple and Complex Liquids* (Cambridge University Press, 2003).
- ⁶⁸S. Agapiou, O. Papaspiliopoulos, D. Sanz-Alonso, and A. M. Stuart, “Importance sampling: Intrinsic dimension and computational cost,” *Stat. Sci.* **32**, 405 (2017).
- ⁶⁹E. Hoffmann, L. Klein, F. Noe, J. Rogal, and M. Schebek, “Boltzmann generators for condensed matter via Riemannian flow matching,” in AI for Accelerated Materials Design - ICLR 2026, 2026.
- ⁷⁰C. B. Tan, J. Bose, C. Lin, L. Klein, M. M. Bronstein, and A. Tong, “Scalable equilibrium sampling with sequential Boltzmann generators,” in *Proceedings of the 42nd International Conference on Machine Learning, Proceedings of Machine Learning Research Vol. 267*, edited by A. Singh, M. Fazel, D. Hsu, S. Lacoste-Julien, F. Berkenkamp, T. Maharaj, K. Wagstaff, and J. Zhu (PMLR, 2025), pp. 58467–58498.
- ⁷¹L. Grenioux and M. Noble, “Diffusion-based annealed Boltzmann generators: Benefits, pitfalls and hopes,” *arXiv:2601.21026* (2026).
- ⁷²A. Phillips, H.-D. Dau, M. J. Hutchinson, V. De Bortoli, G. Deligiannidis, and A. Doucet, “Particle denoising diffusion sampler,” in *Proceedings of the 41st International Conference on Machine Learning, Proceedings of Machine Learning Research Vol. 235*, edited by R. Salakhutdinov, Z. Kolter, K. Heller, A. Weller, N. Oliver, J. Scarlett, and F. Berkenkamp (PMLR, 2024), pp. 40688–40724.
- ⁷³F. Zhang, L. I. Midgley, and J. M. Hernández-Lobato, “Efficient and unbiased sampling from Boltzmann distributions via variance-tuned diffusion models,” *Trans. Mach. Learn. Res.* (published online) (2025), available online at <https://openreview.net/forum?id=Jq2dcMCS5R>.
- ⁷⁴B. McNaughton, M. V. Milošević, A. Perali, and S. Pilati, “Boosting Monte Carlo simulations of spin glasses using autoregressive neural networks,” *Phys. Rev. E* **101**, 053312 (2020).
- ⁷⁵L. M. D. Bono, F. Ricci-Tersenghi, and F. Zamponi, “On the performance of machine-learning-assisted Monte Carlo in sampling from simple statistical physics models,” *arXiv:2505.22598* (2025).
- ⁷⁶R. T. Q. Chen (2018). “torchdiffeq,” GitHub. <https://github.com/rtqichen/torchdiffeq>