Harmonic Sampling: Zero-Parameter Flows on Compact Manifolds

Isay Katsman [*1] **Anna C. Gilbert** [*2]

Abstract

Sampling from probability distributions over compact Riemannian manifolds is a central challenge in the natural sciences, with critical applications ranging from directional statistics to lattice quantum field theory (QFT). Recent approaches rely on manifold continuous normalizing flows to model these densities; however, training these manifold flows via likelihood maximization or distribution matching is notoriously unstable and computationally expensive. In this paper, we bypass the optimization paradigm entirely by introducing Harmonic Sampling, a novel, zero-parameter flow-based sampling framework. By projecting the target density onto the manifold's harmonic basis (the eigenfunctions of the Laplace-Beltrami operator), we analytically derive a deterministic, closed-form vector field that exactly transports a uniform prior to the target distribution. We provide rigorous theoretical guarantees demonstrating that this vector field is globally well-posed, singularity-free, and bounded, with approximation errors that decay super-algebraically. Empirically, we validate our framework on hyperspheres (S^2, S^3) and the special unitary groups $SU(2)$ and $SU(3)$. Across all evaluated geometries, our generated samples demonstrate rapid spectral convergence, recovering exact ground-truth densities. For lattice QFT applications, we show how one can easily adapt Harmonic Sampling to natively enforce strict gauge (conjugation) invariances. In short, our paper provides a method to obtain exact, high-fidelity samples given only a density function over a compact manifold, entirely eliminating the need for neural network training in this context.

[1]Department of Applied and Computational Mathematics, Yale University, New Haven, CT, USA [2]Department of Statistics & Data Science, Yale University, New Haven, CT, USA. Correspondence to: Anna C. Gilbert <anna.gilbert@yale.edu>.

Proceedings of the AI4Physics Workshop at the 43rd International Conference on Machine Learning (AI4Physics@ICML 2026), Seoul, South Korea. 2026. Copyright 2026 by the author(s).

1. Introduction

Learning probabilistic models for data has long been a central focus in machine learning and statistics. While immense effort has been dedicated to developing powerful generative models over Euclidean spaces (Goodfellow et al., 2014; Chen et al., 2018; Grathwohl et al., 2019), less attention has been allocated to the non-Euclidean setting, despite the fact that many real-world systems inherently possess a manifold structure (Boyda et al., 2021; Rezende et al., 2020). Density modelling and sampling over non-Euclidean spaces have critical applications ranging from quantum field theory in physics (Wirnsberger et al., 2020) and motion estimation in robotics (Feiten et al., 2013) to protein-structure prediction in computational biology (Hamelryck et al., 2006).

In particular, many of the manifolds encountered in the physical sciences are compact (e.g., hyperspheres S^n, or the special unitary groups $SU(n)$). The concrete problem we address in this paper is the fast and efficient sampling of known target densities on compact manifolds. That is, given a compact manifold $\mathcal{M}$ and a target density function $\nu : \mathcal{M} \to \mathbb{R}$, we seek to draw independent, accurate samples from ν quickly. This is a critical bottleneck in various applied and theoretical physics subfields, notably in lattice quantum field theory (QFT) (Boyda et al., 2021). Traditional sampling techniques, such as Markov Chain Monte Carlo (MCMC) (Bishop, 2006), suffer from slow mixing times and high autocorrelation, rendering them too computationally expensive for many practical physics simulations. To circumvent this, prior work (Kanwar et al., 2020; Boyda et al., 2021) has proposed constructing manifold-specific continuous normalizing flows (CNFs), training them to approximate the target density, and subsequently using the learned flows as fast, independent samplers.

Although flow-based constructions offer rapid sampling post-training, the optimization procedure itself is notoriously difficult. Training continuous normalizing flows on manifolds requires backpropagating through numerical ODE solvers, a process that is frequently numerically unstable, highly sensitive to hyperparameters, and computationally expensive (Lou et al., 2020).

To resolve this, we propose moving beyond the computational optimization paradigm entirely. Instead, we leverage the harmonic functions of the underlying compact manifold

(the eigenfunctions of the Laplace-Beltrami operator) to construct an *explicit, closed-form* transport vector field. Because our flow is defined analytically, it requires zero training. By converting the density approximation problem into a signal decomposition problem in the harmonic domain, we obtain a deterministic ordinary differential equation (ODE) that exactly transports a uniform base distribution into the desired target density. Consequently, the unpredictable instabilities of deep learning optimization are replaced by the well-understood dynamics of numerical integration and harmonic analysis: we can smoothly trade off sampling speed and accuracy simply by adjusting the step size and tolerance of our chosen ODE solver, or the number of harmonics used in our density approximation. The only approximation we make is the numerical solution of an exact, analytic problem.

We demonstrate the versatility of this framework across several fundamental manifolds. First, we tackle hyperspheres (S^2 and S^3), which are foundational for directional statistics (Fisher, 1953). Second, we address the Lie groups $SU(2)$ and $SU(3)$. In the context of lattice QFT, sampling on $SU(n)$ is strictly governed by physical symmetries; the target distributions are gauge-invariant, meaning their densities must be invariant under matrix conjugation. We show that our harmonic approach naturally and elegantly accommodates these strict symmetry constraints.

In summary, our primary contributions are as follows:

(i) We introduce a novel, zero-optimization manifold continuous normalizing flow for sampling on compact manifolds. We provide rigorous theoretical guarantees (Theorem 1), proving that our analytically derived vector field is globally well-posed and exactly yields the target density.

(ii) We demonstrate the empirical viability of our approach on S^2 and S^3, showing that it can efficiently and accurately sample from standard directional distributions (e.g., von Mises-Fisher) without the need for rejection sampling or neural network training.

(iii) We apply Harmonic Sampling to $SU(2)$ and $SU(3)$, specifically targeting the conjugation-invariant densities required by lattice QFT. We show that our method easily accommodates these physical symmetries and generates high-fidelity samples.

2. Related Work

Our work sits at the intersection of several rapidly developing subfields in machine learning: manifold density estimation, symmetry-preserving generative models, and transport-based sampling.

Normalizing Flows on Manifolds Normalizing flows have long been touted as powerful generative models over Euclidean space (Dinh et al., 2016; Chen et al., 2018; Grathwohl et al., 2019). Similar to GANs (Goodfellow et al., 2014) and VAEs (Kingma & Welling, 2013), normalizing flows learn to map samples from a tractable prior density to a complex target density. However, unlike other models, flows account for changes in volume via the change-of-variables formula, enabling exact density evaluation. Recognizing the need for theoretically principled modelling on non-Euclidean spaces, several works have extended flows to manifolds. Initial efforts focused on manifold-specific constructions, such as tori and spheres (Rezende et al., 2020) or hyperbolic spaces (Bose et al., 2020). Subsequently, Lou et al. (2020), Mathieu & Nickel (2020), and Falorsi & Forré (2020) concurrently introduced general constructions by extending Neural ODEs (Chen et al., 2018) to arbitrary Riemannian manifolds. More recently, Elamvazhuthi et al. (2026) extended this framework to manifolds with boundary by introducing projected neural ODEs, establishing universal approximation guarantees for architectures that interleave geometric updates. While our work takes inspiration from the continuous-time framework of Lou et al. (2020) and related papers, we fundamentally diverge from their methodology: rather than training a neural network to approximate a vector field, we specialize to compact manifolds to obtain exact, zero-optimization flows via harmonic analysis.

Symmetry-Preserving Models and Lattice QFT In many scientific applications, target distributions possess strict inherent symmetries that generative models must respect. Motivated by this, a rich line of work has developed equivariant normalizing flows for both Euclidean spaces (Köhler et al., 2020) and manifolds (Katsman et al., 2021). A primary driver of this research is lattice quantum field theory (QFT), where sampling from gauge-invariant target densities (often defined on Lie groups like $SU(n)$) is a severe computational bottleneck for various integrals that currently utilize traditional MCMC methods. To address this, Kanwar et al. (2020) and Boyda et al. (2021) proposed equivariant flow-based samplers for $U(1)$ and $SU(n)$, with the work by Katsman et al. (2021) generalizing further to an equivariant Riemannian manifold context. While these models successfully preserve conjugation invariance and improve sampling efficiency, they require complex, computationally expensive, frequently unstable training paradigms. Our approach circumvents these issues entirely; by utilizing the harmonic basis of the manifold, our method naturally accommodates conjugation-invariant densities on $SU(n)$ without requiring any neural network parameterization or gradient-based training.

Diffusion and Score-Based Models on Manifolds Because our method relies on reversing the heat flow to trans-

port a target density to a uniform distribution, it shares deep mathematical ties with diffusion and score-based generative models (Song et al., 2021; Ho et al., 2020). Recently, these frameworks have been generalized to Riemannian manifolds (De Bortoli et al., 2022; Huang et al., 2022), where the forward noising process is modelled as Brownian motion on the manifold. To reverse this process, these methods must learn the Stein score (the gradient of the log-density) by training a neural network via score matching. In contrast, Harmonic Sampling is designed for scenarios where the target density function is known but difficult to sample from. By analytically computing the harmonic decomposition of the target, we obtain the exact score function in closed form at all times t, allowing us to simulate the reversed diffusion ODE deterministically and exactly, thereby bypassing the need for score matching entirely.

3. Background

In this section, we provide a terse overview of the necessary concepts for understanding our paper. We address fundamental notions from Riemannian geometry, harmonic analysis, and the basic set-up of continuous normalizing flows on manifolds. For a more detailed introduction to Riemannian geometry and Lie groups, we refer the reader to textbooks such as Lee (2013), Bump (2004), and Hall (2015).

3.1. Riemannian Geometry and Harmonic Analysis

A Riemannian manifold $(\mathcal{M}, \psi)$ is an n-dimensional manifold (Lee, 2013) $\mathcal{M}$ equipped with a smooth collection of inner products $\psi_x(\cdot,\cdot)$ on the tangent space $T_x\mathcal{M}$ at each point $x \in \mathcal{M}$. The Riemannian metric ψ induces a natural volume measure $d\mathcal{M}$ and a distance metric d_ψ on the manifold. In this work, we restrict our focus to *compact* manifolds. Intuitively, these are spaces that are "closed" and of strictly finite extent (such as spheres, tori, or bounded matrix groups) as opposed to unbounded spaces like the infinite Euclidean plane[1]. Crucially, this geometric property ensures the manifold has a finite total volume V, and as a consequence, $\mathcal{M}$ naturally admits a uniform distribution $\rho_{\text{unif}}(x) = 1/V$. More generally, a *probability density* on any Riemannian manifold is defined as a non-negative scalar function $\rho : \mathcal{M} \to \mathbb{R}_{\geq 0}$ that integrates to one with respect to the volume measure ($\int_{\mathcal{M}} \rho(x) d\mathcal{M} = 1$). In our generative modelling framework, our primary objective is to draw independent samples distributed according to some complex, predefined *target density* $\nu(x)$.

For a smooth scalar function $f : \mathcal{M} \to \mathbb{R}$, the Riemannian gradient $\nabla_x f$ is the unique vector on $T_x\mathcal{M}$ such that $\psi(\nabla_x f, v) = D_x f(v)$ for all tangent vectors $v \in T_x\mathcal{M}$, where $D_x f$ is the differential of f (also called the pushforward, which is a generalization of the classical Euclidean Jacobian). The gradient ∇f then refers to the entire vector field, and the Laplace-Beltrami operator Δ is defined as the divergence of the gradient, $\Delta f = \nabla \cdot \nabla f$.

Crucially for our method, when $\mathcal{M}$ is a compact manifold, $-\Delta$ admits a discrete spectrum of non-negative eigenvalues $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots \to \infty$. The corresponding eigenfunctions $h_n(x) : \mathcal{M} \to \mathcal{M}$, satisfying $\Delta h_n(x) = -\lambda_n h_n(x)$, are called the *harmonic functions* of the manifold. They form a complete orthonormal basis for $L^2(\mathcal{M})$ (Donnelly, 2006), generalizing the standard Fourier basis to non-Euclidean geometries.

3.2. Lie Groups and Symmetries

Many applications in the physical sciences require modelling densities over Lie groups, which are manifolds that also possess a smooth group structure. In particular, lattice quantum field theory heavily relies on the Special Unitary group $SU(n)$, defined as the set of $n \times n$ complex unitary matrices with determinant equal to 1.

Compact Lie groups like $SU(n)$ possess a unique, translation-invariant probability measure called the *Haar measure* (Folland, 2016), which serves as the natural uniform distribution (the uninformative prior) over the group. Furthermore, physically meaningful target densities over these groups must often obey strict symmetries; notably, they must be *gauge-invariant*, which corresponds mathematically to being invariant under matrix conjugation. A target density ν on $SU(n)$ is conjugation-invariant if $\nu(U) = \nu(gUg^{-1})$ for all $U, g \in SU(n)$.

3.3. Normalizing Flows on Manifolds

Manifold Normalizing Flow Let $(\mathcal{M}, \psi)$ be a Riemannian manifold. A normalizing flow on $\mathcal{M}$ is a diffeomorphism $f_\theta : \mathcal{M} \to \mathcal{M}$ (parametrized by θ) that transforms a tractable prior density ρ_0 into a complex model density ρ_1. The resulting density is computed via the Riemannian change of variables[2]:

$$\rho_1(x) = \rho_0\left(f_\theta^{-1}(x)\right)\left|\det{}_\psi D_x f_\theta^{-1}\right| \tag{1}$$

Manifold Continuous Normalizing Flow (CNF) A continuous normalizing flow constructs this diffeomorphism by integrating a time-dependent vector field $w_t(x)$ over a finite time horizon. The trajectories of individual particles $\gamma : [0, T] \to \mathcal{M}$ are governed by the ordinary differential

[1] For a rigorous definition of compact manifold, see Lee (2002).

[2] Here, $\det_\psi$ is the determinant function with volume induced by the Riemannian metric ψ.

equation:

$$\frac{d\gamma(t)}{dt} = w_t(\gamma(t)), \qquad \gamma(0) = z \sim \rho_0(z) \tag{2}$$

At the population level, the temporal evolution of the probability density $\rho_t(x)$ under this deterministic vector field is strictly governed by the continuity equation (Villani, 2009):

$$\frac{\partial \rho_t}{\partial t} + \nabla \cdot (\rho_t w_t) = 0 \tag{3}$$

Standard manifold CNFs (Lou et al., 2020) parameterize $w_t(x)$ with a neural network and train it via maximum likelihood estimation to minimize distribution distances. In contrast, our work bypasses this optimization entirely by analytically deriving $w_t(x)$ to exactly satisfy Eq. 3 for a chosen target density.

4. Methodology: Exact Harmonic Sampling

The core contribution of our work is a closed-form manifold Continuous Normalizing Flow (CNF) that exactly maps the uniform distribution on a compact manifold to a desired target density. We achieve this by analyzing the heat equation in the harmonic domain, deriving its equivalent probability flow ODE, and deterministically reversing the diffusion over a finite time horizon.

4.1. Harmonic Decomposition of Densities

Let $(\mathcal{M}, \psi)$ be a compact, boundaryless Riemannian manifold of dimension d and finite volume V. Let Δ denote the Laplace-Beltrami operator on $\mathcal{M}$. Because $\mathcal{M}$ is compact, standard spectral theory (Donnelly, 2006) guarantees that $-\Delta$ has a discrete spectrum of non-negative eigenvalues $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots \to \infty$. Furthermore, there exists a corresponding sequence of smooth eigenfunctions $h_n(x)$ that form an orthonormal basis for $L^2(\mathcal{M})$.

Suppose we are given a smooth, strictly positive target probability density $\nu(x)$ on $\mathcal{M}$. Because the eigenfunctions form a complete basis, we can project $\nu(x)$ onto this basis to obtain its harmonic series representation:

$$\begin{aligned} \nu(x) &= \sum_{n=0}^{\infty} a_n h_n(x), \\ \text{where} \quad a_n &= \int_{\mathcal{M}} \nu(x) h_n(x) d\mathcal{M} \end{aligned} \tag{4}$$

Consider the $n = 0$ base term. The eigenvalue $\lambda_0 = 0$ corresponds to constant eigenfunctions, as $\Delta f = 0$ for f constant. By the orthonormality requirement $\int_{\mathcal{M}} h_0(x)^2 d\mathcal{M} = 1$, we must have $h_0(x) = 1/\sqrt{V}$. Consequently, the coefficient a_0 evaluates to:

$$a_0 = \int_{\mathcal{M}} \nu(x) \frac{1}{\sqrt{V}} d\mathcal{M} = \frac{1}{\sqrt{V}} \int_{\mathcal{M}} \nu(x) d\mathcal{M} = \frac{1}{\sqrt{V}} \tag{5}$$

where the final equality follows from the fact that $\nu(x)$ is a valid probability density integrating to 1. The projection onto the zeroth harmonic is therefore $a_0 h_0(x) = \left(1/\sqrt{V}\right) \cdot \left(1/\sqrt{V}\right) = 1/V$. We can thus explicitly separate the uniform distribution term from the higher-frequency harmonics:

$$\nu(x) = \frac{1}{V} + \sum_{n=1}^{\infty} a_n h_n(x) \tag{6}$$

4.2. The Harmonic Heat Flow

We first consider the forward diffusion process that smooths the target distribution $\nu(x)$ into the uniform distribution. Let $\rho : \mathcal{M} \times \mathbb{R}^+ \to \mathbb{R}^+$, given as $\rho(\cdot, \tau)$, be a valid density for every $\tau \geq 0$ and evolve it under the heat equation $\frac{\partial \rho}{\partial \tau} = \Delta \rho$, with the initial condition $\rho(x, 0) = \nu(x)$. By expanding this partial differential equation in the harmonic domain, we obtain the exact density evolution and its corresponding deterministic vector field in closed form.

Lemma 1 (Harmonic Probability Flow)**.** *The exact solution to the heat flow on the manifold $\mathcal{M}$ with initial condition $\rho(x, 0) = \nu(x)$ is given by:*

$$\rho(x, \tau) = \frac{1}{V} + \sum_{n=1}^{\infty} a_n e^{-\lambda_n \tau} h_n(x) \tag{7}$$

Furthermore, the evolution of the density $\rho(x, \tau)$ under the heat equation is equivalent to its evolution under the continuity equation $\frac{\partial \rho}{\partial \tau} = -\nabla \cdot (\rho v)$ guided by the deterministic vector field:

$$v(x, \tau) = -\nabla \log \rho(x, \tau) \tag{8}$$

Proof. See Appendix A.

4.3. Exact Sampling on a Finite Time Horizon

Lemma 1 defines a forward flow that transports the target distribution into the uniform distribution as $\tau \to \infty$. To construct a practical generative model, we must reverse this flow and map the infinite time horizon $\tau \in [0, \infty)$ to a finite generation time interval. We map the generative process to $t \in [0, 1]$, where $t = 0$ corresponds to the uniform distribution ($\tau = \infty$), and $t = 1$ corresponds to the target distribution ($\tau = 0$).

Theorem 1 (Exact Harmonic Sampling)**.** *Define the time transformation $t = e^{-\lambda_1 \tau/2}$. The time-varying probability density for $t \in [0, 1]$ is given by:*

$$\rho_t(x) = \frac{1}{V} + \sum_{n=1}^{\infty} a_n t^{2\lambda_n/\lambda_1} h_n(x) \tag{9}$$

Consider the ordinary differential equation governing $x(t) \in \mathcal{M}$:

$$\frac{dx(t)}{dt} = \frac{2}{\lambda_1 t} \nabla \log \rho_t(x), \quad x(0) \sim \textit{Uniform}(\mathcal{M}) \quad (10)$$

Then, at $t = 1$, the distribution of $x(1)$ is exactly the target distribution $\nu(x)$. Furthermore, the vector field governing $x(t)$ smoothly extends to $t = 0$ and is Lipschitz continuous on the entire interval $t \in [0, 1]$, guaranteeing global existence and uniqueness of the sampling trajectories.

Proof. The full proof, which demonstrates the exactness of the flow via the normalizing flow continuity equation and establishes well-posedness by rigorously resolving the apparent singularity at $t = 0$, is provided in Appendix A.

4.4. Computational Implementation

The analytic flow defined by Theorem 1 translates directly into a fast, closed-form sampling algorithm. The speed of the algorithm depends on the dimension of the manifold, the desired accuracy, and the number of samples necessary to achieve the desired accuracy of the decomposition coefficients. While high dimensional manifolds can pose computational challenges, the sampling to compute the harmonic coefficients can be parallelized. Unlike standard continuous normalizing flows that require expensive neural ODE training, our method requires zero optimization. The sampling algorithm is as follows:

1. **Initialization:** Draw $x_0 \sim \text{Uniform}(\mathcal{M})$.
2. **Harmonic Truncation:** Approximate $\rho_t(\cdot)$ by computing the harmonic decomposition of $\nu(\cdot)$ up to the first N basis functions.
3. **Riemannian Integration:** Numerically integrate the exact ODE (Eq. 10) from $t = 0$ to $t = 1$ by discretizing the continuous time horizon into small steps of size Δt. At each integration step i, the Euclidean gradients of the density must be converted to Riemannian gradients to define an instantaneous velocity vector strictly within the local tangent space $T_{x_i}\mathcal{M}$. The step size Δt is necessary to scale this instantaneous velocity into a discrete tangent displacement vector. To ensure the particle advances along the curved geometry without leaving the manifold or suffering from projection drift, this displacement is evaluated via the Riemannian exponential map (do Carmo, 1992), which transports the particle along a geodesic: $x_{i+1} = \exp_{x_i}\left(\frac{2}{\lambda_1 t_i} \nabla \log \rho_{t_i}(x_i) \Delta t\right)$.
4. **Polynomial Time Scheduling:** Standard fixed-step integration methods (e.g., Euler or Runge-Kutta) can struggle because the target density $\rho_t(x)$ may become sharply peaked near $t \to 1$. We stabilize the solver by utilizing a 4th-order Runge-Kutta (RK4) scheme adapted for manifolds via the Riemannian exponential map (Munthe-Kaas, 1999; Hairer et al., 1993) alongside a non-linear polynomial step schedule (e.g., Δt decreasing proportionally to $(1 - t)^p$). This allocates a higher integration resolution near $t = 1$ where the vector field varies most rapidly.

5. Experiments

We include experiments demonstrating how to use Harmonic Sampling to exactly generate from a target density ν on a compact manifold $\mathcal{M}$. We focus on manifolds that are particularly pertinent to the physical sciences: the spheres S^2 and S^3 (fundamental for directional statistics (Fisher, 1953)) as well as the Lie groups $SU(2)$ and $SU(3)$ (fundamental for Lattice Quantum Field Theory (Boyda et al., 2021)).

5.1. Generation over Hyperspheres $\mathbf{S^2}$ and $\mathbf{S^3}$

In our first demonstrated use cases, we sample from the hyperspheres S^2 and S^3 using von Mises-Fisher distributions as the target densities.

5.1.1. Target Distribution

We wish to sample from a given von Mises-Fisher (vMF) distribution. The probability density function of the vMF distribution for a p-dimensional unit vector $x \in S^{p-1}$ is given by:

$$f_p(x; \mu, \kappa) = C_p(\kappa) \exp(\kappa \mu^T x) \quad (11)$$

where $\kappa \geq 0$ is the concentration parameter, $||\mu|| = 1$ is the mean direction, and the normalization constant $C_p(\kappa)$ is:

$$C_p(\kappa) = \frac{\kappa^{p/2-1}}{(2\pi)^{p/2} I_{p/2-1}(\kappa)} \quad (12)$$

where I_v denotes the modified Bessel function of the first kind at order v (Pennec et al., 2020). If $p = 3$ (corresponding to S^2), the normalization constant reduces to $C_3(\kappa) = \frac{\kappa}{2\pi(e^{\kappa} - e^{-\kappa})}$. The distribution is unimodal for $\kappa > 0$ and uniform on the sphere for $\kappa = 0$.

5.1.2. Spherical Harmonics

To implement the flow described in Theorem 1, we must project the target density onto the eigenfunctions of the Laplace-Beltrami operator over S^{p-1}. For the 2-sphere ($p = 3$), the eigenfunctions are precisely the classical spherical harmonics. Following Feng & Adler (2019), the spherical harmonics for S^2 are given by:

$$Y_l^m(\theta, \phi) = (-1)^m \sqrt{\frac{(2l+1)}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta) e^{im\phi} \tag{13}$$

for $l \geq 0$ and $m \in \{-l, \ldots, l\}$, where $P_l^m(x)$ are the associated Legendre polynomials. The set of functions $\{Y_l^m\}$ forms a complete orthonormal basis for $L^2(S^2)$, and each function Y_l^m has a corresponding Laplace-Beltrami eigenvalue of $\lambda_l = l(l+1)$.

For a general hypersphere $S^{p-1} \subset \mathbb{R}^p$ (where $p \geq 3$), this basis naturally generalizes to hyperspherical harmonics (Dai & Xu, 2013). These are geometrically defined as the restriction of harmonic homogeneous polynomials of degree l in $\mathbb{R}^p$ to the unit sphere. In standard hyperspherical coordinates $(\theta_1, \ldots, \theta_{p-2}, \phi)$, the orthonormal basis functions are indexed by a sequence of integers $l = l_1 \geq l_2 \geq \cdots \geq l_{p-2} \geq |m| \geq 0$. They can be explicitly constructed using Gegenbauer polynomials $C_n^{(\alpha)}(x)$:

$$Y_{l_1,\ldots,l_{p-2},m}(\theta_1, \ldots, \theta_{p-2}, \phi) = \mathcal{N}_{l,m} e^{im\phi} \prod_{j=1}^{p-2} \sin^{l_{j+1}}(\theta_j)\, C_{l_j - l_{j+1}}^{\left(l_{j+1} + \frac{p-j-1}{2}\right)}(\cos\theta_j) \tag{14}$$

where we denote $l_{p-1} = |m|$, and $\mathcal{N}_{l,m}$ acts as the normalization constant to ensure the functions have unit norm. For any given principal degree $l \geq 0$, all corresponding hyperspherical harmonics share the degenerate Laplace-Beltrami eigenvalue $\lambda_l = l(l+p-2)$. By computing the projection onto this basis up to a chosen maximum degree, we obtain a complete, orthogonal, and analytically differentiable representation of the target density that systematically scales to higher-dimensional spheres.

We compute the harmonic decomposition coefficients over S^2 and S^3 by integrating the target density with respect to each harmonic individually using a Monte Carlo approximation. For the experiments in our paper, 10^6 Monte Carlo samples is sufficient for accurate estimates of the decomposition coefficients. This is a general approach that will work for any compact manifold, although we also wish to note that an exact quadrature rule may also be employed if the manifold structure is well known (as it is for S^n).

5.1.3. Sampling on S^2

For our first experiment, we target a von Mises-Fisher distribution on S^2 with mean direction $\mu = [0, 0, 1]^T$ and concentration $\kappa = 10.0$. To solve the generating ordinary differential equation (Eq. 10), we utilize a 4th-order Runge-Kutta integrator with 600 steps, paired with a non-linear polynomial time schedule.

Visualizations of 1,000 generated samples utilizing the first $N = 64$ spherical harmonics are presented in Figure 1, demonstrating that our method captures the geometry of the target density. To quantitatively verify the exactness of our generated samples, we compute the Empirical Mean Alignment with the target axis ($\mathbb{E}[\mu^T x]$) and the Empirical Covariance Error ($||\Sigma_{\text{true}} - \Sigma_{\text{sample}}||_F$) across 5 independent generation runs (1,000 samples each). As shown in Table 1 of Appendix C, Harmonic Sampling rigorously preserves the spatial spread and rotational symmetry of the target distribution. By systematically increasing the harmonic truncation limit N, the metrics from our flow samples converge to the exact baseline established via rejection sampling.

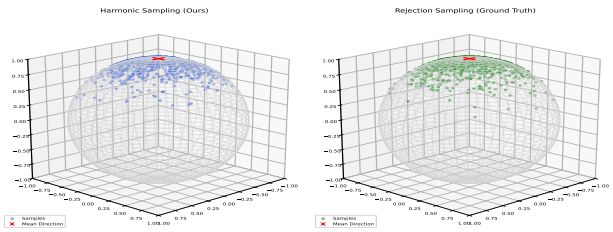

Figure 1. Comparison of sampling a von Mises-Fisher distribution on S^2 ($\mu = [0, 0, 1]^T, \kappa = 10.0$). **Left:** Harmonic Sampling utilizing the first 64 spherical harmonics and 600 integration steps. **Right:** Exact rejection sampling (baseline). Generated samples are plotted in blue, with the target mean direction μ indicated by the red cross.

5.1.4. Sampling on S^3

We naturally extend our sampling methodology to the 3-sphere ($S^3 \subset \mathbb{R}^4$). We define the target vMF density with mean direction $\mu = [0, 0, 0, 1]^T$ and concentration $\kappa = 15.0$. Visualizations of 2,000 generated samples ($N = 160$, 800 integration steps) across projected 3D cross-sections are shown in Figure 2.

Numerical evaluation confirms the exactness of our method in higher dimensions. Table 2 in Appendix C reports the convergence metrics for the 3-sphere across multiple spectral resolutions (evaluated over independent runs of 1,000 samples each). Similar to S^2, as the number of basis functions increases, the empirical metrics of the generated samples rapidly approach those of the exact ground truth. This demonstrates the capacity of Harmonic Sampling to scale gracefully to higher-dimensional compact spaces without hyperparameter tuning or model training.

5.2. Gauge-Invariant Generation over $SU(2)$ and $SU(3)$

While initially motivated by symmetries in lattice gauge theory, sampling from matrix-conjugation-invariant distribu-

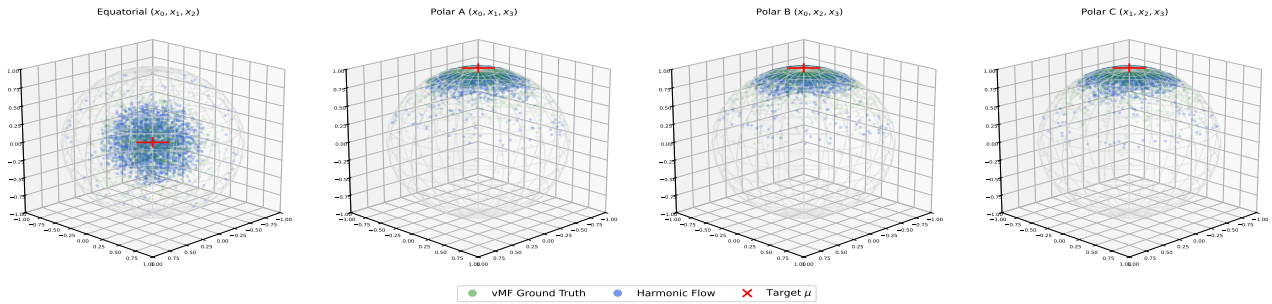

Figure 2. Comparison of sampling a von Mises-Fisher distribution on S^3 ($\mu = [0, 0, 0, 1]^T, \kappa = 15.0$) visualized via 3D cross-section projections. Generated samples using Harmonic Sampling (160 spherical harmonics, 800 integration steps) are plotted in blue, while exact rejection sampling (baseline) samples are plotted in green. The target mean direction μ is indicated by the red cross in each projected subspace.

tions on the special unitary groups $SU(n)$ serves as a highly challenging, multimodal benchmark for generative models on manifolds.

5.2.1. Target Distributions

We evaluate our method on the benchmark matrix-conjugation-invariant distributions introduced by Boyda et al. (2021). These distributions take the form:

$$\nu(U) = \frac{1}{Z} \exp\left(\frac{\beta}{n} \mathrm{ReTr} \sum_{k=1}^{n} c_k U^k\right) \tag{15}$$

where c_k and β are scalar parameters dictating the shape and multimodality of the density, and Z is the normalizing constant with respect to the Haar measure. We denote the unnormalized log-density as $\log \tilde{\nu}(U) = \frac{\beta}{n} \mathrm{ReTr} \sum c_k U^k$.

5.2.2. Conjugation-Invariant Harmonics and Exact Quadrature

The general eigenfunctions of the Laplace-Beltrami operator on $SU(n)$ are the matrix elements of its irreducible representations (e.g., the full Wigner D-matrices for $SU(2)$) (Hall, 2015; Folland, 2016). However, because our target densities are conjugation invariant, their projection onto any non-invariant harmonic is strictly zero. The sub-basis of conjugation-invariant harmonics consists precisely of the *characters* of the irreducible representations (the traces of the representation matrices).

Restricting our basis to characters yields a massive computational advantage. Because characters depend only on the conjugacy class of the matrix, they are purely functions of the matrix eigenvalues. This allows us to invoke the Weyl Integration Formula (Hall, 2015) to drastically reduce the dimensionality of the harmonic decomposition integrals. Rather than relying on noisy Monte Carlo integration over the full manifold, we can perform exact numerical integrals:

- **For $SU(2)$:** We are dealing with a 3-dimensional manifold where the eigenvalues are strictly of the form $e^{i\theta}, e^{-i\theta}$. The irreducible representations are indexed by the degree $L \geq 0$. Using the matrix trace $\mathrm{Tr}(U) = 2\cos(\theta)$, the corresponding character is given by the Weyl Character Formula (Hall, 2015):

$$\chi_L(\theta) = \frac{\sin((L+1)\theta)}{\sin(\theta)} \tag{16}$$

  Computing the harmonic decomposition coefficients over $SU(2)$ thus elegantly reduces to a highly accurate 1D numerical quadrature (via the trapezoidal rule) over the angle $\theta \in [0, \pi]$.

- **For $SU(3)$:** We are dealing with an 8-dimensional manifold. Let the eigenvalues be $z_1 = e^{i\theta_1}$, $z_2 = e^{i\theta_2}$, and $z_3 = e^{-i(\theta_1+\theta_2)}$ (due to the unit determinant constraint). The irreducible representations are indexed by a pair of non-negative integers (p, q). The character $\chi_{p,q}(z_1, z_2, z_3)$ is computed via the Weyl Character Formula (Hall, 2015) as the ratio of alternating polynomials:

$$\chi_{p,q}(z_1, z_2, z_3) = \frac{z_1^A(z_2^B - z_3^B) + z_2^A(z_3^B - z_1^B) + z_3^A(z_1^B - z_2^B)}{z_1^2(z_2 - z_3) + z_2^2(z_3 - z_1) + z_3^2(z_1 - z_2)} \tag{17}$$

  where $A = p + q + 2$, $B = q + 1$, and the denominator is the Vandermonde determinant D. Computing the harmonic decomposition coefficients over $SU(3)$ elegantly reduces to a 2D numerical quadrature over $\theta_1, \theta_2 \in [-\pi, \pi]$ weighted by D.

By utilizing these reduced-dimensional quadratures, we can compute the exact harmonic coefficients of the target density in milliseconds, entirely bypassing the need for stochastic sampling or optimization.

5.2.3. Sampling Results on $SU(2)$

Using the 1D character quadrature, we decompose the three $SU(2)$ Boyda benchmark densities (Boyda et al., 2021) and

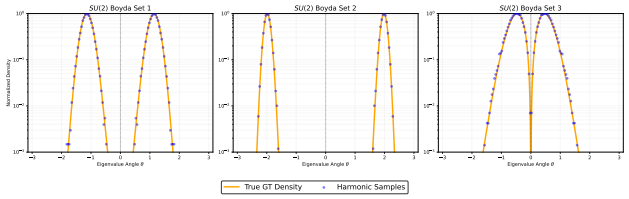

Figure 3. Learned eigenvalue angle densities on $SU(2)$ for the three Boyda benchmark target distributions: Set 1 ($c = [0.98, -0.63, -0.21]$), Set 2 ($c = [0.17, -0.65, 1.22]$), and Set 3 ($c = [1.0, 0.0, 0.0]$), all evaluated for $\beta = 9.0$. Across all landscapes, the empirical angle distribution of 20,000 samples generated via Harmonic Sampling (32 character representations, 250 integration steps) matches the exact analytical target density without any parameter optimization or training.

integrate the harmonic flow using 250 steps. As shown in Figure 3, our zero-optimization flow ($N = 32$, 20,000 generated samples) smoothly and exactly captures the target distributions across varying densities.

To quantitatively verify the exactness of our generated samples, we compute the expected unnormalized log-density under the generated measure, $\mathbb{E}[\log \tilde{\nu}(U)]$, and compare it to a high-precision Monte Carlo baseline estimate. As detailed in Table 3 of Appendix C, across 5 independent generation runs (5,000 samples each), our harmonic flow achieves high accuracy even with few harmonics, properly transporting the uniform prior into the high-density regions of the target distribution.

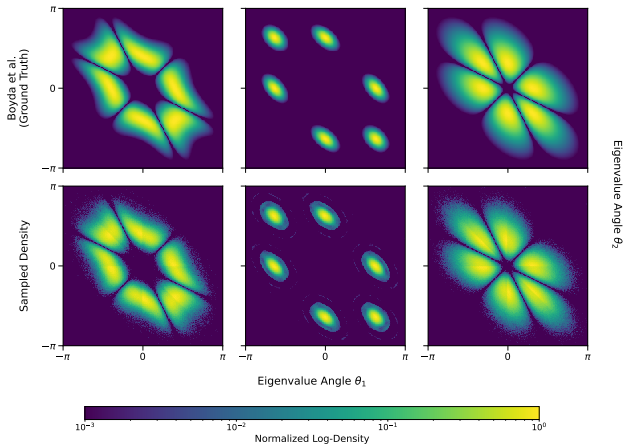

Figure 4. Learned 2D eigenvalue angle densities on $SU(3)$ for the three Boyda benchmark target distributions ($\beta = 9.0$ and c as in Figure 3). **Top Row:** The exact analytical ground truth densities. **Bottom Row:** The empirical densities of 2,000,000 samples generated via Harmonic Sampling (150 character representations, 50 integration steps). The flow naturally and exactly resolves the highly complex, multimodal geometries and symmetries of the target distributions without the discrete artifacts seen in neural coupling-layer approaches (Boyda et al., 2021).

5.2.4. Sampling Results on $SU(3)$

We extend our character basis formulation to $SU(3)$ using the 2D Weyl integration scheme as described before. For this experiment, we target the three $SU(3)$ Boyda benchmark distributions.

Visualized via the 2D eigenvalue angle density using 2,000,000 high-resolution generated samples ($N = 150$, 50 integration steps) in Figure 4, Harmonic Sampling cleanly and naturally resolves the sharp multimodal peaks and the exact symmetry of the target distributions. This entirely avoids the discrete "grid-like" artifacts commonly seen in neural coupling-layer approaches (Boyda et al., 2021).

To rigorously validate this higher-dimensional setting, we compute the expected unnormalized log-densities for 10,000 generated samples per run utilizing 150 integration steps, comparing them against an exact high-precision Monte Carlo baseline. As detailed in Table 4 of Appendix C, our method achieves tight adherence across all three challenging densities.

6. Conclusion and Future Work

We introduced Harmonic Sampling, an exact, zero-parameter continuous normalizing flow for compact manifolds. By leveraging manifold harmonics, we analytically derived a deterministic vector field that transports a uniform prior into a specified target density. We provided theoretical guarantees establishing that the flow is globally well-posed, singularity-free, and exact. Empirically, we demonstrated the versatility of our approach on hyperspheres S^2 and S^3, matching target directional distributions. Furthermore, we applied our framework to the Lie groups $SU(2)$ and $SU(3)$, natively enforcing the matrix-conjugation (gauge) invariances required in lattice quantum field theory. By utilizing the Weyl integration formula over the character basis, our method computes exact flows in milliseconds, bypassing the artifacts and computational cost of training neural manifold flows. Ultimately, by replacing the instabilities of neural network training with the exactness of harmonic analysis, Harmonic Sampling provides a rigorous and efficient tool for probabilistic modeling in the physical sciences.

Future Work. While Harmonic Sampling is highly effective for compact geometries, a natural and exciting theoretical extension is to generalize this methodology to non-compact manifolds, such as hyperbolic spaces. Because the Laplace-Beltrami spectrum on non-compact manifolds is continuous rather than discrete, the harmonic series expansion must generalize to an integral transform (e.g., the Helgason-Fourier transform), presenting unique theoretical and numerical challenges for evaluating the generative ODE.

Impact Statement

This paper presents work whose goal is to advance the field of generative modelling. There are many potential societal consequences of our work, none which we feel must be specifically highlighted here.

References

Bishop, C. M. Pattern recognition and machine learning (information science and statistics). 2006.

Bose, J., Smofsky, A., Liao, R., Panangaden, P., and Hamilton, W. Latent variable modelling with hyperbolic normalizing flows. In *Proceedings of the 37th International Conference on Machine Learning*, pp. 1045–1055, 2020.

Boyda, D., Kanwar, G., Racanière, S., Rezende, D. J., Albergo, M. S., Cranmer, K., Hackett, D. C., and Shanahan, P. E. Sampling using $su(n)$ gauge equivariant flows. *Physical Review D*, 103(7):074504, 2021.

Bump, D. *Lie groups*. Springer, 2004.

Chavel, I. *Eigenvalues in Riemannian Geometry*. Academic Press, 1984.

Chen, R. T. Q., Rubanova, Y., Bettencourt, J., and Duvenaud, D. K. Neural ordinary differential equations. In *Advances in Neural Information Processing Systems*, volume 31, pp. 6571–6583, 2018.

Dai, F. and Xu, Y. *Approximation Theory and Harmonic Analysis on Spheres and Balls*. Springer, 2013.

De Bortoli, V., Mathieu, E., Hutchinson, M., Thornton, J., Teh, Y. W., and Doucet, A. Riemannian score-based generative modelling. In *Advances in Neural Information Processing Systems*, volume 35, pp. 2406–2422, 2022.

Dinh, L., Sohl-Dickstein, J., and Bengio, S. Density estimation using real nvp. *arXiv preprint arXiv:1605.08803*, 2016.

do Carmo, M. P. *Riemannian Geometry*. Birkhäuser, 1992.

Donnelly, H. G. Eigenfunctions of the laplacian on compact riemannian manifolds. *Asian Journal of Mathematics*, 10: 115–126, 2006.

Elamvazhuthi, K., Biswal, S., Rosenblum, K., Katyal, A., Qu, T., Ma, G., and Sonthalia, R. Geometry-preserving neural architectures on manifolds with boundary. *ArXiv*, abs/2602.03082, 2026. URL https://api.semanticscholar.org/CorpusID:285276151.

Falorsi, L. and Forré, P. Neural ordinary differential equations on manifolds. *arXiv preprint arXiv:2006.06663*, 2020.

Feiten, W., Lang, M., and Hirche, S. Rigid motion estimation using mixtures of projected gaussians. *Proceedings of the 16th International Conference on Information Fusion*, pp. 1465–1472, 2013.

Feng, R. and Adler, R. J. Critical radius and supremum of random spherical harmonics. *The Annals of Probability*, 2019.

Fisher, R. A. Dispersion on a sphere. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 217:295 – 305, 1953. URL https://api.semanticscholar.org/CorpusID:123166853.

Folland, G. B. *A Course in Abstract Harmonic Analysis*. CRC Press, 2016.

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. Generative adversarial nets. In *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2*, pp. 2672–2680, 2014.

Grathwohl, W., Chen, R. T. Q., Bettencourt, J., and Duvenaud, D. Scalable reversible generative models with free-form continuous dynamics. In *International Conference on Learning Representations*, 2019.

Grigor'yan, A. *Heat Kernel and Analysis on Manifolds*. American Mathematical Society, 2009.

Hairer, E., Wanner, G., and Nørsett, S. P. *Solving ordinary differential equations I: Nonstiff problems*. Springer, 1993.

Hall, B. C. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Springer, 2015.

Hamelryck, T., Kent, J. T., and Krogh, A. Sampling realistic protein conformations using local structural bias. *PLoS Computational Biology*, 2(9), 2006.

Ho, J., Jain, A., and Abbeel, P. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33:6840–6851, 2020.

Huang, C.-W., Aghajohari, M., Bose, J., Panangaden, P., and Courville, A. Riemannian diffusion models. In *Advances in Neural Information Processing Systems*, volume 35, pp. 2750–2761, 2022.

Kanwar, G., Albergo, M. S., Boyda, D., Cranmer, K., Hackett, D. C., Racaniere, S., Rezende, D. J., and Shanahan, P. E. Equivariant flow-based sampling for lattice gauge theory. *Physical Review Letters*, 125(12):121601, 2020.

Katsman, I., Lou, A., Lim, D., Jiang, Q., Lim, S.-N., and Sa, C. D. Equivariant manifold flows. *ArXiv*, abs/2107.08596, 2021.

Kingma, D. P. and Welling, M. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

Köhler, J., Klein, L., and Noe, F. Equivariant flows: Exact likelihood generative learning for symmetric densities. In *Proceedings of the 37th International Conference on Machine Learning*, pp. 5361–5370, 2020.

Lang, S. *Fundamentals of Differential Geometry*. Springer, 1999.

Lee, J. M. Introduction to smooth manifolds. 2002.

Lee, J. M. *Introduction to Smooth Manifolds*. Graduate Texts in Mathematics. Springer New York, 2013.

Lou, A., Lim, D., Katsman, I., Huang, L., Jiang, Q., Lim, S. N., and De Sa, C. M. Neural manifold ordinary differential equations. In *Advances in Neural Information Processing Systems*, volume 33, pp. 17548–17558, 2020.

Mathieu, E. and Nickel, M. Riemannian continuous normalizing flows. In *Advances in Neural Information Processing Systems*, volume 33, pp. 2503–2515, 2020.

Munthe-Kaas, H. High order runge-kutta methods on manifolds. *Applied Numerical Mathematics*, 29(1):115–127, 1999.

Pennec, X., Sommer, S., and Fletcher, T. Riemannian geometric statistics in medical image analysis. 2020.

Rezende, D. J., Papamakarios, G., Racaniere, S., Albergo, M., Kanwar, G., Shanahan, P., and Cranmer, K. Normalizing flows on tori and spheres. In *Proceedings of the 37th International Conference on Machine Learning*, pp. 8083–8092, 2020.

Song, Y., Sohl-Dickstein, J., Kingma, D. P., Kumar, A., Ermon, S., and Poole, B. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2021.

Villani, C. *Optimal Transport: Old and New*. Springer, 2009.

Wirnsberger, P., Ballard, A. J., Papamakarios, G., Abercrombie, S., Racanière, S., Pritzel, A., Jimenez Rezende, D., and Blundell, C. Targeted free energy estimation via learned mappings. *The Journal of Chemical Physics*, 153 (14):144112, 2020.

Appendix

A. Proofs for Exact Harmonic Sampling

In this section, we provide the full mathematical proofs for the lemmas and theorems introduced in Section 4.

A.1. Proof of Lemma 1 (Harmonic Probability Flow)

Lemma 1 (Harmonic Probability Flow). *The exact solution to the heat flow on the manifold $\mathcal{M}$ with initial condition $\rho(x, 0) = \nu(x)$ is given by:*

$$\rho(x, \tau) = \frac{1}{V} + \sum_{n=1}^{\infty} a_n e^{-\lambda_n \tau} h_n(x)$$

Furthermore, the evolution of the density $\rho(x, \tau)$ under the heat equation is equivalent to its evolution under the continuity equation $\frac{\partial \rho}{\partial \tau} = -\nabla \cdot (\rho v)$ guided by the deterministic vector field:

$$v(x, \tau) = -\nabla \log \rho(x, \tau)$$

Proof. To solve the heat equation $\frac{\partial \rho}{\partial \tau} = \Delta \rho$ analytically using standard spectral methods (Grigor'yan, 2009), we express the time-dependent density in the harmonic basis as $\rho(x, \tau) = \sum_{n=0}^{\infty} c_n(\tau) h_n(x)$. Substituting this expansion into the heat equation and exploiting the linearity of the Laplace-Beltrami operator yields:

$$\frac{\partial}{\partial \tau} \sum_{n=0}^{\infty} c_n(\tau) h_n(x) = \Delta \left(\sum_{n=0}^{\infty} c_n(\tau) h_n(x) \right) = \sum_{n=0}^{\infty} c_n(\tau) \Delta h_n(x) \tag{18}$$

Recall that by definition, the eigenfunctions satisfy $\Delta h_n(x) = -\lambda_n h_n(x)$. Substituting this into the right-hand side gives:

$$\sum_{n=0}^{\infty} \frac{dc_n(\tau)}{d\tau} h_n(x) = \sum_{n=0}^{\infty} -\lambda_n c_n(\tau) h_n(x) \tag{19}$$

Because the eigenfunctions $h_n(x)$ form an orthogonal basis for $L^2(\mathcal{M})$, we can equate the coefficients for each n independently. This reduces the partial differential equation to a system of decoupled, first-order ordinary differential equations:

$$\frac{dc_n(\tau)}{d\tau} = -\lambda_n c_n(\tau) \quad \Rightarrow \quad c_n(\tau) = c_n(0) e^{-\lambda_n \tau} \tag{20}$$

The initial coefficients $c_n(0)$ are strictly given by the harmonic projection of the target density $\nu(x)$. From Eq. 4, we know $c_0(\tau) h_0(x) = 1/V$ (where the eigenvalue $\lambda_0 = 0$ implies $e^0 = 1$) and $c_n(0) = a_n$ for $n \geq 1$. Therefore, the exact solution to the heat flow on the manifold is exactly Eq. 7.

To demonstrate the equivalence to the continuity equation, we use the fundamental identity $\Delta \rho = \nabla \cdot (\nabla \rho)$ and the logarithmic gradient property $\nabla \rho = \rho \nabla \log \rho$. We can rewrite the heat equation directly as:

$$\frac{\partial \rho}{\partial \tau} = \nabla \cdot (\rho \nabla \log \rho) \tag{21}$$

Substituting the proposed vector field $v(x, \tau) = -\nabla \log \rho(x, \tau)$ immediately yields $\frac{\partial \rho}{\partial \tau} = -\nabla \cdot (\rho v)$, fulfilling the continuity equation. □

A.2. Proof of Theorem 1 (Exact Harmonic Sampling)

Theorem 1 (Exact Harmonic Sampling). *Define the time transformation $t = e^{-\lambda_1 \tau / 2}$. The time-varying probability density for $t \in [0, 1]$ is given by:*

$$\rho_t(x) = \frac{1}{V} + \sum_{n=1}^{\infty} a_n t^{2\lambda_n / \lambda_1} h_n(x).$$

Consider the ordinary differential equation governing $x(t) \in \mathcal{M}$:

$$\frac{dx(t)}{dt} = \frac{2}{\lambda_1 t} \nabla \log \rho_t(x), \quad x(0) \sim \text{Uniform}(\mathcal{M}).$$

Then, at $t = 1$, the distribution of $x(1)$ is exactly the target distribution $\nu(x)$. Furthermore, the vector field governing $x(t)$ smoothly extends to $t = 0$ and is Lipschitz continuous on the entire interval $t \in [0, 1]$, guaranteeing global existence and uniqueness of the sampling trajectories.

Proof. By inverting the time transformation, we obtain $\tau(t) = -\frac{2}{\lambda_1} \log t$. Substituting this into the forward heat solution (Eq. 7) yields $e^{-\lambda_n \tau(t)} = e^{(\lambda_n/\lambda_1) \cdot 2 \log t} = t^{2\lambda_n/\lambda_1}$, which immediately recovers Eq. 9. This confirms that $\rho_t(x) \equiv \rho(x, \tau(t))$.

Next, we derive the ordinary differential equation that generates this density evolution. We evaluate the time derivative of the time-scaled density $\rho_t(x)$ using the chain rule. From Lemma 1, the continuity equation for the heat flow can be written as $\frac{\partial \rho}{\partial \tau} = \nabla \cdot (\rho \nabla \log \rho)$. Evaluating this at $\tau(t)$ and immediately substituting $\rho_t(x)$ yields:

$$\frac{\partial \rho_t(x)}{\partial t} = \frac{\partial \rho(x, \tau)}{\partial \tau}\bigg|_{\tau=\tau(t)} \frac{d\tau(t)}{dt} = \nabla \cdot \left[\rho_t(x) \nabla \log \rho_t(x)\right] \left(-\frac{2}{\lambda_1 t}\right). \tag{22}$$

Because the scalar factor $\left(-\frac{2}{\lambda_1 t}\right)$ depends only on time and is independent of the spatial coordinates x, it can be brought freely inside the spatial divergence operator $\nabla \cdot$:

$$\frac{\partial \rho_t(x)}{\partial t} + \nabla \cdot \left(\rho_t(x) \left[\frac{2}{\lambda_1 t} \nabla \log \rho_t(x)\right]\right) = 0. \tag{23}$$

By the fundamental relationship between deterministic particle trajectories and density transport (often referred to as the continuity equation for Continuous Normalizing Flows (Villani, 2009)), if a time-varying probability density $\mu_t(x)$ satisfies $\frac{\partial \mu_t}{\partial t} + \nabla \cdot (\mu_t w_t) = 0$, then $\mu_t(x)$ is exactly the marginal density of particles evolving according to the ordinary differential equation $\frac{dx(t)}{dt} = w_t(x)$. Matching our derived Eq. 23 to this standard form, the required vector field is exactly $w_t(x) = \frac{2}{\lambda_1 t} \nabla \log \rho_t(x)$. This recovers precisely the ordinary differential equation given in Eq. 10. Integrating this ODE forward from $t = 0$ to $t = 1$ transports the initial uniform distribution (ρ_0) to the exact target distribution ($\rho_1 = \nu$).

To establish global existence and uniqueness, we must rigorously address the apparent singularity at $t = 0$ caused by the $1/t$ scaling term. We analyze the behavior of the vector field $\frac{dx}{dt} = \frac{2}{\lambda_1 t} \frac{\nabla \rho_t(x)}{\rho_t(x)}$ as $t \to 0$. Let $\mathcal{E}_1 = \{n \mid \lambda_n = \lambda_1\}$ be the index set corresponding to the first non-zero eigenvalue. For infinitesimally small t, the time-scaled density and its gradient expand as:

$$\rho_t(x) = \frac{1}{V} + t^2 \sum_{n \in \mathcal{E}_1} a_n h_n(x) + \mathcal{O}(t^p), \tag{24}$$

$$\nabla \rho_t(x) = t^2 \sum_{n \in \mathcal{E}_1} a_n \nabla h_n(x) + \mathcal{O}(t^p), \tag{25}$$

where $p = 2\lambda_k/\lambda_1 > 2$ is dictated by the next distinct eigenvalue $\lambda_k > \lambda_1$. We can now evaluate the limit of the vector field by applying the standard limit laws for quotients. For the numerator, the $1/t$ factor reduces the leading t^2 term to a linear t dependence:

$$\lim_{t \to 0} \frac{2}{\lambda_1 t} \nabla \rho_t(x) = \lim_{t \to 0} \frac{2}{\lambda_1} \left(t \sum_{n \in \mathcal{E}_1} a_n \nabla h_n(x) + \mathcal{O}(t^{p-1})\right) = 0. \tag{26}$$

For the denominator, we clearly have $\lim_{t \to 0} \rho_t(x) = 1/V$. Because the limit of the numerator is 0 and the limit of the denominator is strictly positive ($1/V \neq 0$), the limit of the quotient is precisely 0:

$$\lim_{t \to 0} \frac{dx(t)}{dt} = \frac{\lim_{t \to 0} \frac{2}{\lambda_1 t} \nabla \rho_t(x)}{\lim_{t \to 0} \rho_t(x)} = \frac{0}{1/V} = 0. \tag{27}$$

The singularity is cleanly canceled by the leading t^2 factor of the density gradient. The vector field smoothly extends to 0 at $t = 0$. Because $\nu(x)$ is smooth and $\mathcal{M}$ is compact, Weyl's Law (Chavel, 1984) guarantees rapid decay of the spectral coefficients a_n, ensuring the vector field is smooth and globally Lipschitz on $\mathcal{M} \times [0, 1]$. By the Picard-Lindelöf theorem on manifolds (Lang, 1999), the initial value problem is well-posed, and the flow mapping the uniform distribution to $\nu(x)$ is exact and unique. □

B. Extended Theory and Convergence Guarantees

In this section, we provide extended theoretical results for Harmonic Sampling. First, we generalize our continuous normalizing flow to transport between any two arbitrary smooth distributions (rather than restricting the base distribution to be uniform). Second, we provide rigorous convergence bounds on the harmonic truncation error, giving a practical criterion for choosing the number of basis functions N.

B.1. Generalization to Arbitrary Source Distributions

While Theorem 1 assumes a uniform base distribution $x(0) \sim \text{Uniform}(\mathcal{M})$ for simplicity and numerical convenience, the Harmonic Sampling framework trivially extends to tasks where we seek to map samples between two arbitrary, unstructured distributions.

Theorem 2 (General Harmonic Flow)**.** *Let $\mathcal{M}$ be a closed (compact and boundaryless) Riemannian manifold. Let $\mu(x)$ and $\nu(x)$ be any smooth, strictly positive target probability densities on $\mathcal{M}$. There exists an explicitly defined, piece-wise continuous vector field $w_t(x)$ that exactly transports $\mu(x)$ to $\nu(x)$ over the time interval $t \in [-1, 1]$.*

Proof. By Theorem 1, there exists a density evolution $\rho_t^\mu(x)$ and a corresponding vector field $w_t^\mu(x) = \frac{2}{\lambda_1 t}\nabla \log \rho_t^\mu(x)$ that exactly transports the uniform distribution to $\mu(x)$ over $t \in [0, 1]$. Similarly, there exists $\rho_t^\nu(x)$ and $w_t^\nu(x)$ that exactly transports the uniform distribution to $\nu(x)$ over $t \in [0, 1]$.

To construct a flow from $\mu(x)$ to $\nu(x)$, we concatenate the reverse flow of μ with the forward flow of ν. We define the generalized time interval $t \in [-1, 1]$ and construct the global vector field as follows:

$$w_t(x) = \begin{cases} -w_{|t|}^\mu(x) & \text{for } t \in [-1, 0) \\ 0 & \text{for } t = 0 \\ w_t^\nu(x) & \text{for } t \in (0, 1] \end{cases} \tag{28}$$

Integrating the ODE $\frac{dx}{dt} = w_t(x)$ starting from $x(-1) \sim \mu(x)$ will transport the particles along the reversed paths of μ's harmonic flow, arriving exactly at the uniform distribution at $t = 0$. From $t = 0$ to $t = 1$, the ODE integrates the forward harmonic flow of ν, transporting the uniform intermediate distribution exactly to $\nu(x)$. Because both constituent vector fields are globally well-posed and smoothly extend to 0 at the uniform limit, the concatenated flow is exact, bounded, and unique. □

B.2. Spectral Truncation and Practical Convergence

In practice (as described in Section 4.4), we cannot compute the infinite sum in Eq. 9. We must truncate the harmonic decomposition at the first N basis functions, yielding the approximated density $\rho_t^N(x)$ and the approximated vector field $w_t^N(x)$.

The following theorem guarantees that this truncation is safe, showing that the error decays incredibly fast as N increases. It relies on standard results from harmonic analysis bounding the growth of Laplace-Beltrami eigenfunctions (Donnelly, 2006).

Theorem 3 (Harmonic Truncation Error Bound)**.** *Let $\nu(x) \in C^\infty(\mathcal{M})$ be a smooth, strictly positive target density. Let $w_t(x)$ be the exact harmonic vector field from Theorem 1, and $w_t^N(x) = \frac{2}{\lambda_1 t}\nabla \log \rho_t^N(x)$ be the truncated vector field. The supremum error of the truncated vector field decays super-algebraically with N. That is, for any positive integer $k > 0$, there exists a constant C_k and a threshold N^* such that for all $N \geq N^*$:*

$$\sup_{x \in \mathcal{M},\ t \in [0,1]} ||w_t(x) - w_t^N(x)||_\infty \leq C_k \lambda_{N+1}^{-k}. \tag{29}$$

Proof. Let the absolute truncation error in the density be $e_N(x,t) = \rho_t(x) - \rho_t^N(x) = \sum_{n=N+1}^{\infty} a_n t^{2\lambda_n/\lambda_1} h_n(x)$. Because the target density $\nu(x)$ is infinitely differentiable (C^∞), the harmonic coefficients $a_n = \int_{\mathcal{M}} \nu(x) h_n(x) d\mathcal{M}$ decay faster than any polynomial in λ_n (Chavel, 1984). That is, $|a_n| = \mathcal{O}(\lambda_n^{-m})$ for any $m > 0$.

Furthermore, by standard elliptic regularity and Weyl's Law (Chavel, 1984), the supremum norms $||h_n||_\infty$ and $||\nabla h_n||_\infty$ grow at most polynomially with λ_n.

Because the exact heat flow preserves strict positivity and the initial density $\nu(x)$ is strictly positive, the exact density is bounded away from zero everywhere: $\inf_{x\in\mathcal{M}, t\in[0,1]} \rho_t(x) = \rho_{\min} > 0$. Since $\sup_{x,t} |e_N(x,t)|$ converges uniformly to 0, there exists an integer N^* such that for all $N \geq N^*$, $\sup_{x,t} |e_N(x,t)| \leq \rho_{\min}/2$. Thus, the truncated density approximation is uniformly bounded below: $\rho_t^N(x) \geq \rho_{\min}/2 > 0$. We note that this observation means that, in practice, we never see `NaN` values in the truncated density, even for small N, because the density is strictly positive and bounded away from zero. While we cannot determine N^* a priori, it does depend on the smoothness of the target density $\nu(x)$ and the geometry of the manifold $\mathcal{M}$, which we assume to be smooth as well.

We now bound the difference in the vector fields:

$$\begin{aligned} w_t(x) - w_t^N(x) &= \frac{2}{\lambda_1 t}\left(\frac{\nabla\rho_t(x)}{\rho_t(x)} - \frac{\nabla\rho_t^N(x)}{\rho_t^N(x)}\right) \\ &= \frac{2}{\lambda_1}\frac{\rho_t^N(x)\frac{\nabla e_N(x,t)}{t} - \frac{e_N(x,t)}{t}\nabla\rho_t^N(x)}{\rho_t(x)\rho_t^N(x)}. \end{aligned}$$

To resolve the $1/t$ singularity near $t = 0$, observe that the time-scaled error terms expand as:

$$\frac{e_N(x,t)}{t} = \sum_{n=N+1}^{\infty} a_n t^{2\lambda_n/\lambda_1 - 1} h_n(x), \quad \frac{\nabla e_N(x,t)}{t} = \sum_{n=N+1}^{\infty} a_n t^{2\lambda_n/\lambda_1 - 1} \nabla h_n(x). \tag{30}$$

Because the sequence of eigenvalues is non-decreasing, for any truncation $N \geq 1$, the remaining terms are indexed by $n \geq 2$, meaning $\lambda_n \geq \lambda_1$. Therefore, the exponent satisfies $(2\lambda_n/\lambda_1 - 1) \geq 1$, ensuring that $t^{2\lambda_n/\lambda_1 - 1} \leq t^1 \leq 1$ for all $t \in [0,1]$.

Consequently, the time-scaled error tails converge absolutely and do not diverge as $t \to 0$. By the super-algebraic decay of a_n, both $\sup_{x,t} ||\frac{e_N}{t}||$ and $\sup_{x,t} ||\frac{\nabla e_N}{t}||$ are bounded by $\mathcal{O}(\lambda_{N+1}^{-k})$ for any chosen k.

Since the denominator $\rho_t(x)\rho_t^N(x) \geq \rho_{\min}^2/2$ is strictly positive for $N \geq N^*$, and the numerator decays as $\mathcal{O}(\lambda_{N+1}^{-k})$, the supremum error of the entire vector field obeys $||w_t(x) - w_t^N(x)||_\infty = \mathcal{O}(\lambda_{N+1}^{-k})$, completing the proof. □

Practical Criterion for Choosing N. Theorem 3 provides a highly practical stopping criterion for the decomposition step. Because the error is strongly dominated by the first omitted term $n = N + 1$, the ℓ_2 integration error of the sampling trajectories is directly proportional to the magnitude of the harmonic coefficients.

To achieve a target error tolerance ϵ for the sampler, one simply computes the coefficients a_n iteratively (e.g., using the exact Weyl quadratures described in Section 5) and stops when $|a_N| < \epsilon$. Unlike deep learning flows where approximation error is opaque and bounds are loose, Harmonic Sampling offers a transparent, monitorable, and monotonically decreasing error profile.

C. Detailed Quantitative Experimental Results

In this section, we provide the full quantitative breakdown of the convergence results for the spheres S^2 and S^3, as well as for the special unitary groups $SU(2)$ and $SU(3)$. These tables supplement the qualitative visualizations provided in the main text.

C.1. Convergence on S^2 and S^3

Table 1 details the empirical mean alignment and covariance error as a function of the harmonic truncation limit N for the von Mises-Fisher distribution on the 2-sphere. Table 2 details the same for the 3-sphere. We see rapid convergence in the number of harmonics used in both cases.

Table 1. Quantitative evaluation of Harmonic Sampling on the S^2 von Mises-Fisher target distribution ($\kappa = 10.0$). We report the Empirical Mean Alignment and Empirical Covariance Error (Mean $\pm$ Std over 5 independent runs) as a function of the harmonic truncation limit N. As the spectral resolution increases, the generated samples converge to the exact ground truth metrics.

Number of Harmonics (N)	**Empirical Mean Alignment**	**Empirical Covariance Error**
9	0.795 ± 0.007	0.099 ± 0.006
16	0.847 ± 0.007	0.050 ± 0.007
25	0.873 ± 0.006	0.024 ± 0.007
64	0.895 ± 0.002	0.007 ± 0.002
Exact (Ground Truth)	$\mathbf{0.900}$	$\mathbf{0.000}$

Table 2. Quantitative evaluation of Harmonic Sampling on the S^3 von Mises-Fisher target distribution ($\kappa = 15.0$). We report the Empirical Mean Alignment and Empirical Covariance Error (Mean $\pm$ Std over 5 independent runs) as a function of the harmonic truncation limit N. As the spectral resolution increases, the generated samples converge to the exact ground truth metrics.

Number of Harmonics (N)	**Empirical Mean Alignment**	**Empirical Covariance Error**
5	0.617 ± 0.002	0.184 ± 0.002
10	0.637 ± 0.004	0.186 ± 0.004
20	0.766 ± 0.008	0.064 ± 0.006
40	0.830 ± 0.005	0.023 ± 0.003
Exact (Ground Truth)	$\mathbf{0.838}$	$\mathbf{0.000}$

C.2. Convergence on $SU(n)$ Groups

Tables 3 and 4 provide the expected log-density and absolute error for the Boyda benchmark distributions on $SU(2)$ and $SU(3)$. Note that the harmonics considered here are the conjugation invariant harmonics, as specified in Section 5.2.2, which is valid (and in fact desirable) since the target distributions are invariant to conjugation.

Table 3. Quantitative evaluation of Harmonic Sampling on the $SU(2)$ Boyda benchmark distributions ($\beta = 9.0$ and c as in Figure 3). We report the Expected Log-Density (Mean $\pm$ Std over 5 independent runs) and the Absolute Error relative to the high-precision Monte Carlo baseline. As the character basis spectral resolution (N) increases, the generated manifold samples rapidly converge to the exact target measure across all densities.

Target Density	Number of Harmonics (N)	Expected Log-Density	Absolute Error
Boyda Set 1	2	6.962 ± 0.034	1.751
	4	8.156 ± 0.023	0.557
	8	8.637 ± 0.015	0.076
	16	8.698 ± 0.008	0.015
	Exact (Ground Truth)	**8.713**	**0.000**
Boyda Set 2	2	6.787 ± 0.068	6.444
	4	11.364 ± 0.023	1.867
	8	12.540 ± 0.012	0.690
	16	13.141 ± 0.012	0.089
	Exact (Ground Truth)	**13.230**	**0.000**
Boyda Set 3	2	5.425 ± 0.021	2.118
	4	7.097 ± 0.024	0.446
	8	7.481 ± 0.011	0.062
	16	7.498 ± 0.017	0.045
	Exact (Ground Truth)	**7.543**	**0.000**

Table 4. Quantitative evaluation of Harmonic Sampling on the $SU(3)$ Boyda benchmark distributions ($\beta = 9.0$ and c as in Figure 3). We report the Expected Log-Density (Mean $\pm$ Std over 5 independent runs) and the Absolute Error relative to the high-precision Haar Monte Carlo baseline. As the character basis spectral resolution (N) increases, the generated manifold samples steadily converge to the exact target measure across all multimodal densities.

Target Density	Number of Harmonics (N)	Expected Log-Density	Absolute Error
Boyda Set 1	10	5.408 ± 0.021	0.452
	20	5.748 ± 0.010	0.112
	40	5.812 ± 0.008	0.047
	80	5.822 ± 0.009	0.038
	Exact (Ground Truth)	**5.860**	**0.000**
Boyda Set 2	10	8.890 ± 0.003	1.512
	20	9.512 ± 0.022	0.890
	40	9.994 ± 0.018	0.408
	80	10.241 ± 0.008	0.161
	Exact (Ground Truth)	**10.402**	**0.000**
Boyda Set 3	10	4.934 ± 0.023	0.273
	20	5.149 ± 0.018	0.057
	40	5.187 ± 0.014	0.020
	80	5.180 ± 0.015	0.026
	Exact (Ground Truth)	**5.206**	**0.000**