Noname manuscript No.
(will be inserted by the editor)

Generating an equidistributed net on a sphere using random rotations

Somnath Chakraborty · **Hariharan Narayanan**

Received: date / Accepted: date

Abstract We develop a randomized algorithm (that succeeds with high probability) for generating an ε-net in a sphere of dimension n. The basic scheme is to pick an alphabet consisting of $O(n\ln(1/\varepsilon)+\ln(1/\delta))$ random rotations, form all possible words of length $O(n\ln(1/\varepsilon))$ from this alphabet, and require these words act on a fixed point. We show the set of points so generated is equidistributed at a scale of ε.

Keywords ε-net · equidistributed net · random rotations · Wasserstein distance

Mathematics Subject Classification (2010) Primary 22D40

1 Introduction

In this article, we develop a randomized algorithm (with high success probability) for the generation of an ε-net in a unit sphere of dimension n. The basic scheme is to pick an alphabet consisting of $k := O(n\ln(1/\varepsilon)+\ln(1/\delta))$ random rotations, take all possible words of length $\ell := O(n\ln(1/\varepsilon))$ from this alphabet, and make them act on a fixed point on the sphere. We show that this set of points is equidistributed at a scale of ε. The ε-net so generated, requires, in a sense explained in subsection 1.1,

School of Technology and Computer Science,
Tata Institute of Fundamental Research,
Mumbai 400005, India
Tel: +91-22-2278-2988
Fax: +91-22-2278-2299
E-mail: somnath.chakraborty@alumni.iu.edu

School of Technology and Computer Science,
Tata Institute of Fundamental Research,
Mumbai 400005, India
Tel: +91-22-2278-2455
Fax: +91-22-2278-2299
E-mail: hariharan.narayanan@tifr.res.in

less randomness than the natural Monte Carlo method of taking a large number of i.i.d random samples from the sphere. Quantitatively, our result is most interesting when we let $\varepsilon \to 0$ for a fixed the dimension n. We are motivated in part by the possibility of using equidistributed points for the integration of Lipschitz functions over an n-sphere. There is a large body of related research that has to do with minimum energy configurations on the sphere; see for example [17]. The net we produce, is ε-close in Hausdorff distance to the n-sphere, and is also equidistributed (with high probability) in the following sense: the uniform counting measure ν over the net is close to the uniform measure σ on the n-sphere in 1-Wasserstein distance. This implies that the integral of every 1-Lipschitz function on the n-sphere $\mathbb{S}^n$ with respect to ν is "ε-close" to its integral with respect to σ.

In order to put this work in context, we briefly survey earlier work. In [2], Alon and Roichman proved that, *given any $\delta > 0$, there exists a $c(\delta) > 0$ such that for any finite group G, and a random subset $S \subset G$ of order at least $c(\delta)\log|G|$, the induced Cayley graph $\chi(G,S)$ has small normalized second largest eigenvalue (in absolute value)*:

$$\mathbb{E}\left(|\lambda_2^*(\chi(G,S))|\right) < \delta. \tag{1}$$

Considering random walk on expander multigraphs, it follows that every element $g \in G$ is an S-word of length at most $\log|G|$. In [13], Russell and Landau deviced a short proof (with better constants) of this result while rephrasing the question using representation theory as follows. For an irreducible representation $\rho \in \hat{G}$, let d_ρ be its dimension; let R be the regular representation of G, and $D = \sum_{\rho\in\hat{G}} d_\rho$. Russel and Landau ([13]) proved that (1) holds for random subsets $S \subset G$ of order at least

$$\left(\tfrac{2\ln 2}{\delta} + o(1)\right)^2 \log|D|$$

This was obtained via an application of *tail bounds for operator-valued random variables*, as in Ahlswede and Winter [1], building upon the following observation: *the normalized adjacency matrix of $\chi(G,S)$ is the operator*

$$(2|S|)^{-1}\sum_{s\in S}(R(s)+R(s^{-1})), \tag{2}$$

presented in terms of the standard basis of $\mathbb{C}[G]$.

In equivalent terms, the quoted result from [2] means that random Cayley graph $\chi(G,S)$ is an expander and also, the operator in (2) has a spectral gap satisfying (1). Now let G be a compact Lie group, and μ a left-invariant Borel probability measure on G. One considers the averaging operator $z_\mu : L^2(G) \to L^2(G)$, given by

$$z_\mu(f)(x) = \int_G f(xg)\, d\mu(g)$$

In [5], Bourgain and Gamburd established that if $G = SU_d$ (group of d-dimensional special unitary matrices), then z_μ has spectral radius < 1 when $\mathrm{supp}(\mu)$ is finite algebraic subset generating a nonabelian free subgroup of G. This has since been extended to all compact connected simple Lie groups by Benoist and de Saxcé in [3], where it was shown that z_μ has a spectral gap if and only if μ is *almost diophantine*. A corollary to the main result in [3] is the following: if μ is finitely-supported almost diophantine then the set of words in $\mathrm{supp}(\mu)$ of fixed length approaches G in Hausdorff distance.

1.1 Reduced use of randomness

In [15], a quantitative version of the spectral gap question was considered. It was shown that the Hausdorff distance between G, a compact connected Lie group, and the subset of fixed length words on a random essentially small finite alphabet $S \subset G$ decays exponentially in the length of the words, with high probability. This was done via an analysis of the heat kernel with respect to a suitable finite dimensional subspace of $L^2(G)$ and an application of *tail bounds for operator-valued random variables* as in [1]. We note that the results of the present article are not implied by the results of [15], because the dimension of the Lie group SO_n (the group of n-dimensional orthogonal matrices with determinant 1) is $n(n+1)/2$, and so the bounds from [15] for the length of the words and the number of generators, that apply for general compact Lie groups would be quadratic in n rather than linear in n. In the special case of the unitary group U_n, such a result with a quadratic dependence on dimension for the length of the words was previously obtained by Hastings and Harrow in Theorem 5 of [10], however in their result the number of generators is specified in a indirect manner, whose dependence on n is not obvious. On the other hand, the bounds obtained in the present work are linear in n, both for the number of generators and the length of the words. In fact, for these parameters, the value of $(2k)^\ell$ is close to the volumetric lower bound of $(1/\varepsilon)^{\Omega(n)}$ on the size of an ε-net of $\mathbb{S}^n$.

1.2 Main results

The two main results of this paper are stated below. In the versions below, we have simplified the expressions as they appear in the main body. In the following statements, C denotes a universal constant. For a finite set S, the ℓ-fold product S^ℓ consists of all S-words of length ℓ inside the free group generated by S. As a consequence of Theorem 3.15, we have the following.

Theorem 1:
Let $S \subset SO_{n+1}$ consist of k iid random elements, drawn from the Haar measure on SO_{n+1}, where

$$k \geq C\left(\ln\left(\tfrac{1}{\delta}\right) + n\ln\left(\tfrac{1}{\varepsilon}\right)\right),$$

for some given $\varepsilon, \delta \in (0,1)$. Let $\hat{S} := S \sqcup S^{-1}$ be the (multi)set of all elements in S and their inverses. Let

$$\ell \geq 9n\log_2\left(\tfrac{1}{\varepsilon}\right).$$

If ε is sufficiently small then the probability that $\hat{S}^\ell x_0 \subseteq \mathbb{S}^n$ is an r-net in $\mathbb{S}^n$, where

$$r := 4\varepsilon\sqrt{n\ln\frac{1}{\varepsilon}} = \tilde{\mathscr{O}}(\varepsilon),$$

is at least $1-\delta$.

As a consequence of Theorem 4.14, we have the following.

Theorem 2:
For $n>1$, let σ be the uniform probability measure on $\mathbb{S}^n$. Let ℓ, k and $\hat{S}$ be as defined in Theorem 1. Let $x_0 \in \mathbb{S}^n$ and let ν be the probability measure on $\mathbb{S}^n$, uniformly supported on $\hat{S}^\ell x_0$. If ε is sufficiently small, then the following holds with probability at least $1-\delta$:

$$W_1(\sigma,\nu) := \sup_{\phi\in Lip_1} (\textstyle\int)_{\mathbb{S}^n} \phi \; d\sigma - \int_{\mathbb{S}^n} \phi \; d\nu \leq 2\sqrt{n}\varepsilon,$$

Here W_1 denotes the 1-Wasserstein distance between two measures supported on $\mathbb{S}^n$.

For the remainder of this section, we assume given a real number model of computation, in which only standard algebraic operations are allowed on Gaussian random vectors, but bits are not manipulated. Thus, for k and ℓ as described in Theorem 6, we choose a set S consisting of k orthogonal matrices, each picked up independently from the Haar measure of SO_{n+1}. Consider all $(2k)^\ell$ words of length ℓ in these generators and their inverses. Apply the resulting matrices to the vector $\mathbf{e}_{n+1} = (0,0,\cdots,0,1)^T$. Then these $(2k)^\ell$ points form an equidistributed net, that can be used for integrating a 1-Lipschitz function to within an additive error of ε. If we assume an oracle that outputs independent n-dimensional Gaussian random vectors when queried, then the whole process requires only kn queries to this oracle. Note that the obvious procedure of producing an equidistributed net, would require $\varepsilon^{-\Omega(n)}$ calls to the Gaussian oracle (which would then be normalized to lie on the sphere). The latter method uses exponentially more randomness than our procedure using random rotations.

2 Application of Spherical Harmonics

This section briefly reviews the basics of harmonic analysis on the unit n-sphere $\mathbb{S}^n \subset \mathbb{R}^{n+1}$. The two lemmas in this section will be used in an essential manner in the later sections.

Let g be the standard Riemannian metric tensor on $\mathbb{S}^n$. We denote by $\Pi(x,x_0)$ the space of smooth paths in $\mathbb{S}^n$, between the points $x, x_0 \in \mathbb{S}^n$. The Riemannian distance between points $x, x_0 \in \mathbb{S}^n$ is given by

$$d(x,x_0) = \inf_{\gamma\in\Pi(x,x_0)} \int_0^1 |\gamma'(t)|_g \; dt \tag{3}$$

Considering the standard inclusion $\mathbb{S}^n \subset \mathbb{R}^{n+1}$, one has

$$d(x,x_0) = \arccos\langle x,x_0\rangle$$

There is a unique geodesic path between any $x,x_0\mathbb{S}^n$. Let σ denote the standard euclidean surface probability measure on $\mathbb{S}^n$. For any Borel subset $B \subset \mathbb{S}^n$, if $\hat{B} := \{\alpha x : x \in B, \alpha \in [0,1]\}$ then

$$\sigma(B) = \frac{\lambda(\hat{B})}{\lambda(D)},$$

where λ is the standard Lebesgue measure on $\mathbb{R}^{n+1}$ and $D \subseteq \mathbb{R}^{n+1}$ is the unit disk centered at origin, so that $\mathbb{S}^n = \partial D$. We recall that the Lebesgue measure of the unit n-sphere in $\mathbb{R}^n$ is

$$\Omega_n := \lambda(\mathbb{S}^n) = \frac{2\pi^{\frac{n+1}{2}}}{\Gamma\left(\frac{n+1}{2}\right)} \tag{4}$$

Now, the fact that σ is SO_{n+1}- invariant follows from usual rotation invariance of Lebesgue measure on $\mathbb{R}^{n+1}$. Hence (by unimodularity of the compact Lie group SO_{n+1} of rotations of $\mathbb{S}^n$), the measure σ is the unique probability measure on $\mathbb{S}^n$ induced by the haar measure on SO_{n+1}. For $n > 1$, let $\Delta := \Delta_{\mathbb{S}^n}$ be the negative of the Laplace-Beltrami operator on $\mathbb{S}^n$. Thus, given $g \in C^2(\mathbb{S}^n)$, one has

$$-\Delta(g) = \Delta_{\mathbb{R}^n}(\tilde{g})\mid_{\mathbb{S}^n}$$

where $\tilde{g} : \mathbb{R}^{n+1} - \{0\} \to \mathbb{R}$ is defined by $\tilde{g}(x) = g(|x|^{-1}x)$. A rather more intrinsic definition is $\Delta = -\delta \cdot d$, where $d : C^\infty(\mathbb{S}^n) \to \Lambda(\mathbb{S}^n)$ is the differential map, and $\delta = d^*$ is the adjoint. It is well-known that the Hilbert-product space $L^2(\mathbb{S}^n)$ decomposes into a direct sum of the eigenspaces of Δ, in the sense that the L^2-closure of the direct sum is $L^2(\mathbb{S}^n)$:

$$L^2(\mathbb{S}^n) = \bigoplus_{k=0}^{\infty} H_k(\mathbb{S}^n) \tag{5}$$

Recall that $H_k(\mathbb{S}^n)$ is the space of degree-k homogeneous harmonic polynomials in $n+1$ variables, restricted to $\mathbb{S}^n$; the dimension of $H_k(\mathbb{S}^n)$ is

$$h_k := \binom{n+k}{n} - \binom{n+k-2}{n} \tag{6}$$

and the corresponding eigenvalue is $\lambda_k := k(n+k-1)$. Note that, for any $n,k > 0$, one has

$$\begin{aligned}
\sum_{a=0}^{k} \dim H_a(\mathbb{S}^n) &= \sum_{a=0}^{k} \left(\tbinom{n+a}{a} - \tbinom{n+a-2}{a-2} \right) \\
&= \binom{n+k}{k} + \binom{n+k-1}{k-1} \\
&= \dim H_k(\mathbb{S}^{n+1})
\end{aligned}$$

One has

$$\begin{aligned}
\lim_{k\to\infty}\frac{\sum\limits_{k\le a\le k\sqrt[n]{2}}\dim H_a(\mathbb{S}^n)}{k^n} &= 2\lim_{k\to\infty}\frac{\sum\limits_{0\le a\le\lfloor k\sqrt[n]{2}\rfloor}\dim H_a}{2k^n} - \lim_{k\to\infty}\frac{\sum\limits_{0\le a\le k}\dim H_a}{k^n}\\
&= 2\lim_{l\to\infty}\frac{\sum\limits_{0\le a\le l}\dim H_a}{l^n} - \lim_{k\to\infty}\frac{\sum\limits_{0\le a\le k}\dim H_a}{k^n}\\
&= \lim_{k\to\infty}\frac{\sum\limits_{0\le a\le k}\dim H_a}{k^n}\\
&= \lim_{k\to\infty}\frac{\dim H_k(\mathbb{S}^{n+1})}{k^n}\\
&= \lim_{k\to\infty}k^{-n}\binom{n+k}{k} + \lim_{k\to\infty}k^{-n}\binom{n+k-1}{k-1}\\
&= \frac{2}{n!}
\end{aligned}$$

Moreover, notice that, when $k > 2n+1$, one has

$$\begin{aligned}
\binom{n+k}{k}^2 &= \frac{1}{(n!)^2}\prod_{i=1}^{n}(k+i)(k+n-i+1)\\
\text{AM} \ge \text{GM} \;\Rightarrow\qquad &< \frac{1}{(n!)^2}\prod_{i=1}^{n}2k^2\\
&= \frac{2^n k^{2n}}{(n!)^2}
\end{aligned}$$

Equivalently, writing H_λ for the eigenspace corresponding to eigenvalue λ, one gets

$$\lim_{\lambda\to\infty}\frac{\sum\limits_{\lambda\le a\le\lambda\sqrt[n]{4}}\dim H_a}{\lambda^{\frac{n}{2}}} = \frac{2}{n!}$$

and for all $\lambda > 6n^2+3n$, the inequality $\lambda_{ak} < a^2\lambda_k$ implies

$$\frac{\sum\limits_{\lambda\le a\le\lambda\sqrt[n]{4}}\dim H_a}{\lambda^{\frac{n}{2}}} < \frac{2(2^{\frac{n}{2}})}{n!}$$

where, the sum ranges over all eigenvalues in $[\lambda,\lambda\sqrt[n]{4}]$. Fix a point $x_0\in\mathbb{S}^n$. Let $H_t(x)$ be the heat kernel on $\mathbb{S}^n$, corresponding to Brownian motion started at x_0. That is, $H_t(x)$ is the fundamental solution to the problem

$$\begin{aligned}
\frac{\partial u}{\partial t} &= -\Delta_{\mathbb{S}^n}u\\
\lim_{t\to 0^+}u(t,x) &= \delta_{x_0}(x)
\end{aligned}$$

where the convergence is with respect to the weak* topology. Here $\delta_{x_0}(x)$ is the Diract-δ distribution. A Brownian motion on $\mathbb{S}^n$, started at $x_0 \in \mathbb{S}^n$ has infinitesimal generator $H_{\frac{t}{2}}(x)$. Fixing orthonormal basis $\phi_{k,1}, \cdots, \phi_{k,h_k}$ of $H_k := H_{\lambda_k}$ for each $k \geq 0$, one has

$$\begin{aligned} H_t(x) &= \sum_{k=0}^{\infty} e^{-\lambda_k t} \sum_{i=1}^{h_k} \phi_{k,i}(x_0)\phi_{k,i}(x) \\ &= \sum_{k=0}^{\infty} e^{-\lambda_k t} h_k P_{k,n}(x \cdot x_0) \end{aligned}$$

where the last equality is due to *addition theorem* of spherical harmonics (see theorem 2.26 of [14]) that correspond to the usual addition formula for trigonometric functions when $n = 1$; here $P_{k,n}(t)$ denotes the Legendre polynomial of degree k and dimension $n+1$; in explicit terms, this polynomial is

$$P_{k,n}(t) = \sum_{j=0}^{\lfloor \frac{k}{2} \rfloor} C_{2j} t^{k-2j} (1-t^2)^j \tag{7}$$

where the coefficients are given by

$$C_0 = 1, \qquad C_{2j} = (-1)^j \frac{k(k-1)\cdots(k-2j+1)}{(2\cdot 4\cdots 2j)(n(n+2)\cdots(n+2j-2))}$$

One has

$$\begin{aligned} \int_{\mathbb{S}^n} H_t(x)^2 \, d\sigma(x) &= \sum_{k=0}^{\infty} e^{-2\lambda_k t} (h_k)^2 \int_{\mathbb{S}^n} (P_{k,n}(x \cdot x_0))^2 \, d\sigma(x) \\ &= \sum_{k=0}^{\infty} e^{-2\lambda_k t} h_k \end{aligned}$$

where the last equation is a well-known properties of Legendre polynomials (see theorem 2.29 of [14]). We note that by theorem 2.29 of [14], one has

$$H(t,x,x) = \sum_{k=0}^{\infty} e^{-\lambda_k t} h_k$$

for all $t > 0$. It is known that $H_t(x) > 0$ for all $t > 0$.

For $M > 0$, let $H_{t,M}(x)$ be defined by

$$H_{t,M}(x) := \sum_{\lambda_k \leq M} e^{-\lambda_k t} \sum_{i=1}^{h_k} \phi_{k,i}(x_0)\phi_{k,i}(x) \tag{8}$$

Lemma 1 *Suppose that $t \in (0, 6^{-1})$ and for any $\eta > 0$, let $M \geq 4^{\frac{k_0}{n}}$ where*

$$k_0 > \max\left(\tfrac{1}{2}\sqrt{\log_2 \tfrac{1}{\eta}}, n\log_2\left(\tfrac{n}{t}\right) + 2n\log_2\log_2\left(\tfrac{n}{t}\right) \right); \tag{9}$$

then the following inequality holds:

$$||H_t - H_{t,M}||_{L^2}^2 \leq \eta^2$$

Proof When $k > n\log_2(3n)$, one has $4^{\frac{k}{n}} > 6n^2+3n$. For $k \geq 0$, write

$$I_k = (4^{\frac{k}{n}}, 4^{\frac{k+1}{n}}]$$

For $k_0 > n\log_2(3n)$, one has

$$\begin{aligned}\sum_{\lambda \geq 4^{\frac{k_0}{n}}} e^{-2\lambda t}\dim H_\lambda &= \sum_{k\geq k_0}\sum_{\lambda\in I_k} e^{-2\lambda t}\dim H_\lambda \\ &\leq \sum_{k\geq k_0}(\sup)_{\lambda\in I_k} e^{-2\lambda t}\,(\Sigma)_{\lambda\in I_k}\dim H_\lambda \\ &\leq \frac{2(2^{\frac{n}{2}})}{n!}\sum_{k\geq k_0} 2^{1+k} e^{-(2^{\frac{2k+n}{n}})t}\end{aligned}$$

Suppose that an integer $k > k_0$ satisfies

$$k \geq n\log_2\left(\tfrac{k}{t}\right)$$

Then

$$e^{-(2^{\frac{2k+n}{n}})t} = (e^{-2^{\frac{2k}{n}}})^{2t} < e^{-\frac{2k^2}{t}} < 2^{-17k^2}$$

Consider the inequality

$$\frac{k}{\log_2\left(\frac{k}{t}\right)} \geq n \tag{10}$$

By monotone property of the logarithm function, the following inequality is equivalent to (10) above:

$$k\left(1 - \frac{\log_2\log_2\left(\frac{k}{t}\right)}{\log_2\left(\frac{k}{t}\right)}\right) \geq n\log_2\left(\tfrac{n}{t}\right) \tag{11}$$

For $x \in (2^e, +\infty)$, the function

$$g(x) = 1 - \frac{\log_2\log_2 x}{\log_2 x}$$

satisfies $0 < g(x) < 1$, has global minima $g(2^e) = 1 - e^{-1}\log_2 e > 0.46$, and is increasing. Since $t \in (0, 6^{-1})$ and $n > 1$, the condition $n/t > 2^e$ is satisfied; because $k \geq n\log_2(3n)$, the following inequality implies (11):

$$k \geq n\log_2\left(\tfrac{n}{t}\right)\left(1 - \frac{\log_2\log_2\left(\frac{n}{t}\right)}{\log_2\left(\frac{n}{t}\right)}\right)^{-1} \tag{12}$$

We claim that, for $n > 1$ and $t \in (0, 6^{-1})$, the following inequality holds:

$$\begin{aligned}\frac{\log_2\left(\frac{n}{t}\right) + 2\log_2\log_2\left(\frac{n}{t}\right)}{\log_2\left(\frac{n}{t}\right)} &= 1 + \frac{2\log_2\log_2\left(\frac{n}{t}\right)}{\log_2\left(\frac{n}{t}\right)} \\ &\geq \left(1 - \frac{\log_2\log_2\left(\frac{n}{t}\right)}{\log_2\left(\frac{n}{t}\right)}\right)^{-1} \\ &= \frac{\log_2\left(\frac{n}{t}\right)}{\log_2\left(\frac{n}{t}\right) - \log_2\log_2\left(\frac{n}{t}\right)}\end{aligned}$$

This is equivalent to

$$\log_2\left(\tfrac{n}{t}\right) \geq 2\log_2\log_2\left(\tfrac{n}{t}\right)$$

Writing $x = \frac{n}{t}$, this is equivalent to $\sqrt{x} \geq \log_2 x$. The function $h(x) = \sqrt{x} - \log_2 x$ has derivative $h'(x) = \frac{1}{2\sqrt{x}} - \frac{1}{x\ln 2}$, which is increasing for $x > \frac{4}{(\ln 2)^2}$, and $h(12) > 0$. This proves the claim.

Thus, $k_0 = n\log_2\left(\frac{n}{t}\right) + 2n\log_2\log_2\left(\frac{n}{t}\right)$ implies

$$\begin{aligned}
||H_t - H_{t,M}||_{L^2}^2 &= \sum_{\lambda \geq 4^{\frac{k_0}{n}}} e^{-2\lambda t} \dim H_\lambda \\
&\leq \frac{2(2^{\frac{n}{2}})}{n!} \sum_{k \geq k_0} 2^{1+k} e^{-(2^{\frac{2k+n}{n}})t} \\
&\leq \frac{4(2^{\frac{n}{2}})}{n!} \sum_{k \geq k_0} 2^{1+k-17k^2} \\
&\leq \frac{2^{\frac{n}{2}}}{n!} \sum_{k \geq k_0} 2^{-16k^2} \\
&\leq \frac{2^{\frac{n}{2} - 16k_0^2 + 1}}{n!} \\
&\leq \eta^2
\end{aligned}$$

Remark 1 If $\varepsilon \in (0, \frac{1}{3n})$ and $t = \varepsilon^2$, one has

$$n\log_2\left(\tfrac{n}{t}\right) + 2n\log_2\log_2\left(\tfrac{n}{t}\right) < \frac{3n}{2}\log_2\left(\tfrac{1}{t}\right)\left(1 + \tfrac{2\log_2\log_2(5n)}{\log_2(5n)}\right) \tag{13}$$

We write

$$a_n := \frac{2\log_2\log_2(5n)}{\log_2(5n)}$$

Then the lemma above implies

$$||H_t - H_{t,M}||_{L^2}^2 \leq \eta^2$$

for $M = 4^{\frac{k_0}{n}}$, where $k_0 > \max\left(\log_2\left(\frac{1}{\eta}\right), \frac{3n}{2}\left(1 + a_n\right)\log_2\left(\frac{1}{t}\right)\right)$.

Lemma 2 *Let $c = \min\{\ln\sqrt{2}, n^{-1}\}$. For all $t \in (0, c)$ and all $M > 0$, following inequality holds:*

$$||1_{\mathbb{S}^n} - H_{t,M}||_{L^2}^2 < \frac{t^{-n}}{(n-1)!2^{n-2}} \tag{14}$$

Proof For any $M > 0$, addition theorem implies

$$||1_{\mathbb{S}^n} - H_{t,M}||_{L^2}^2 := \sum_{0 < \lambda \leq M} e^{-2\lambda t} \dim H_\lambda$$

The function $\phi(x) = (x+a)te^{-2xt}$, for $x \in (0,\infty)$, satisfies

$$\phi'(x) = te^{-2xt}(1-2(x+a)t), \quad \phi''(x) = -4t^2e^{-2x}(1-(x+a)t)$$

which shows that $\phi(x) \le (2e)^{-1}e^{2at}$. This yields

$$\begin{aligned}
||1_{S^n} - H_{t,M}||_{L^2}^2 &= \sum_{0<\lambda_k\le M} e^{-2k(n+k-1)t}\dim H_k \\
&\le 2\sum_{0<\lambda_k\le M} \frac{1}{(n-1)!}e^{-2k(n+k-1)t}\prod_{i=1}^{n-1}(k+i) \\
&\le 2\sum_{0<\lambda_k\le M} \frac{e^{-2k^2t}}{(n-1)!t^{n-1}}\prod_{i=1}^{n-1}(k+i)te^{-2kt} \\
&\le 2\sum_{0<\lambda_k\le M} \frac{e^{-2k^2t}}{(n-1)!(2et)^{n-1}}\prod_{i=1}^{n-1}e^{2it} \\
&\le \frac{t^{-(n-1)}e^{(n-1)(nt-1)}}{(n-1)!2^{n-2}}\sum_{k>0}e^{-2k^2t} \\
&< \frac{t^{-(n-1)}}{(n-1)!2^{n-2}}\sum_{k\ge 0}e^{-2kt} \\
&< \frac{t^{-(n-1)}}{(1-e^{-2t})2^{n-2}(n-1)!}
\end{aligned}$$

If $t \in (0, \ln\sqrt{2})$ then $1-e^{-2t} > t$, which implies that, for $t \in (0,c)$ where $c = \min\{\ln\sqrt{2}, n^{-1}\}$, and any $M > 0$, one has

$$||1_{\mathbb{S}^n} - H_{t,M}||_{L^2}^2 < \frac{t^{-n}}{(n-1)!2^{n-2}}$$

3 Hausdorff distance

We recall the following definition:

Definition 1 Given a subset $\hat{S} \subset \mathbb{S}^n$, and $\varepsilon \ge 0$, let $\hat{S}_\varepsilon$ be the union of all ε-neighbourhoods of points in $\hat{S}$; the Hausdorff distance $d_H(\hat{S},\mathbb{S}^n)$ is defined to be

$$\begin{aligned}
d_H(\hat{S},\mathbb{S}^n) :&= \max\{\sup_{x\in\mathbb{S}^n}\inf_{y\in\hat{S}} d(x,y), \sup_{y\in\hat{S}}\inf_{x\in\mathbb{S}^n} d(x,y)\} \\
&= \inf\{\varepsilon \ge 0 : \mathbb{S}^n \subseteq \hat{S}_\varepsilon\}
\end{aligned}$$

where $\hat{S}_\varepsilon = \{x \in \mathbb{S}^n : d(x,x_0) < \varepsilon \text{ for some } x_0 \in S\}$.

Our analysis in this section will be based on an application of the following theorem, first appeared in [1].

Theorem 1 (Ahlswede-Winter) *Let V be a finite dimensional Hilbert space, with $\dim V = D$. Let $A_1, \cdots, A_k$ be independent identically distributed random variables taking values in the cone of positive semidefinite operators on V, such that $\mathbb{E}[A_i] = A \geq \mu I$ for some $\mu \geq 0$, and $A_i \leq I$. Then, for all $\varepsilon \in [0, 0.5]$, the following holds:*

$$\mathbb{P}\left(\tfrac{1}{k} \textstyle\sum_{i=1}^{k} A_i \notin [(1-\varepsilon)A, (1+\varepsilon)A] \right) \leq 2D \exp\left(\tfrac{-\varepsilon^2 \mu k}{2 \ln 2} \right) \tag{15}$$

Let $S \subset SO_{n+1}$ be a non-empty subset, with $|S| = k$. For $M > 6n^2 + 3n$, let

$$E_M := \bigoplus_{0<\lambda\leq M} H_\lambda(S^n)$$

Recall (inequality 7) that $\dim E_M \leq \frac{2(2M)^{\frac{n}{2}}}{n!}$. Because $\Delta := \Delta_{\mathbb{S}^n}$ is SO_{n+1}-invariant, the subspace E_M is invariant under the operators

$$A_s(f)(x) := \frac{1}{2} f(x) + \frac{1}{4}(f(xs) + f(xs^{-1})), \qquad s \in SO_{n+1} \tag{16}$$

Due to rotation invariance of the surface probability measure σ, the operators $A_s : E_M \to E_M$ turns out to be self-adjoint. Positive semidefiniteness of A_s follows from the identity

$$\langle A_s f, f \rangle = \frac{1}{4} \int_{\mathbb{S}^n} (f(x) + f(xs))^2 \, d\sigma(x)$$

Moreover, writing μ for the unique right-invariant Haar (probability) measure on SO_{n+1}, one has

$$(\mathbb{E}_{s\sim\mu}(A_s)) f(x) = \frac{1}{2} f(x) + \frac{1}{4} \int_{SO_{n+1}} f(xs) \, d\mu(s) + \frac{1}{4} \int_{SO_{n+1}} f(xs^{-1}) \, d\mu(s)$$

Writing $\tau : SO_{n+1} \to \mathbb{S}^n$ for the map $s \mapsto xs$, one has

$$\begin{aligned} \int_{SO_{n+1}} f(xs) \, d\mu(s) &= \int_{SO_{n+1}} (f \circ \tau)(s) \, d\mu(s) \\ &= \int_{\mathbb{S}^n} f(y) \, d(\tau_* \mu) \end{aligned}$$

where $\tau_* \mu(E) = \mu(\tau^{-1}(E))$. Since $\tau_* \mu$ is rotation-invariant Borel measure on $\mathbb{S}^n$, one has $\sigma = \tau_* \mu$. Because $f \in E_M \subset L_0^2(\mathbb{S}^n)$, one has

$$\int_{SO_{n+1}} f(xs) \, d\mu(s) = \int_{\mathbb{S}^n} f(y) \, d(\tau_* \mu) = 0$$

Therefore, $(\mathbb{E})_{s\sim\mu} [A_s] f(x) = \frac{1}{2} f(x)$, which makes

$$\mathbb{E}_{s\sim\mu}[A_s] = \frac{1}{2} I, \qquad s \in SO_{n+1}$$

Furthermore, the operators $I - A_s$ are positive semidefinite for all $s \in SO_{n+1}$, because

$$\langle (I - A_s) f, f \rangle = \frac{1}{4} \int_{\mathbb{S}^n} (f(x) - f(xs))^2 \, d\sigma(x).$$

Theorem 2 *Let $S \subset SO_{n+1}$ be a set of order $|S| = k$, chosen independently, and uniformly at random from the Haar measure on SO_{n+1} and let $\hat{S} := S \sqcup S^{-1}$ be the (multi)set of all elements in S and their inverses. Let $\eta > 0$ satisfy*

$$\log_2 \frac{1}{\eta} \geq \frac{3n}{2}(1+a_n)\log_2\left(\tfrac{1}{t}\right) \tag{17}$$

where $a_n := \frac{2\log_2\log_2(5n)}{\log_2(5n)}$. Let $t \in (0,c)$ where $c = \min\{\frac{1}{6}, \frac{1}{n}\}$. For

$$\delta = \frac{2^{\frac{n+4}{2}}}{n!\eta}\exp\left(\tfrac{-k}{16\ln 2}\right)$$

and any integer $\ell > 0$ satisfying $2^\ell \geq \frac{t^{-\frac{n}{2}}}{\eta}$, the following inequality holds:

$$\mathbb{P}\Big(||1_{\mathbb{S}^n} - \tfrac{1}{(2k)^\ell}\sum_{s\in\hat{S}^\ell} H_t(x_0 s, x)||_{L^2} \leq 2\eta\Big) \geq 1-\delta \tag{18}$$

Proof Since $t \in (0,c)$ and $c = \min\{\frac{1}{6}, \frac{1}{n}\}$, one has (via 17)

$$\eta^{-\frac{2}{n}} \geq t^{-3(1+a_n)} > \begin{cases} 6^{3(1+a_n)} & \text{if } n \leq 6 \\ n^{3(1+a_n)} & \text{if } n > 6 \end{cases}$$

Evaluating a_n for $n \leq 6$, we find $\eta^{-\frac{2}{n}} > 6^{3(1+a_n)} > 6n^2 + 3n$ for all $n \leq 6$. On the other hand, since $\eta^{-\frac{2}{n}} > n^3$ for all $n > 6$, and since $n^3 - 6n^2 - 3n > 0$ for all $n > 6$, we have $\eta^{-\frac{2}{n}} > 6n^2 + 3n$ for all integer n. Setting $\varepsilon = 0.5$ and $M = \eta^{-\frac{2}{n}}$ in (15) yields

$$\mathbb{P}\left(\tfrac{1}{k}\textstyle\sum_{s\in S} A_s \notin [\tfrac{1}{4}I, \tfrac{3}{4}I]\right) \leq \frac{2^{\frac{n+4}{2}}}{n!\eta}\exp\left(\tfrac{-k}{16\ln 2}\right)$$

Therefore, for all $f \in E_M$, the following inequality holds:

$$\mathbb{P}\left(||\tfrac{1}{k}\textstyle\sum_{s\in\hat{S}} \tfrac{1}{4} f\cdot\tau(s)||_{L^2} \leq \tfrac{1}{4}||f||_{L^2}\right) \geq 1 - \frac{2^{\frac{n+4}{2}}}{n!\eta}\exp\left(\tfrac{-k}{16\ln 2}\right)$$

In particular, writing $\tilde{H}_{t,M} := 1_{\mathbb{S}^n} - H_{t,M}$, one has

$$\mathbb{P}\left(||\tfrac{1}{2k}\textstyle\sum_{s\in\hat{S}} \tilde{H}_{t,M}(xs)||_{L^2} \leq \tfrac{1}{2}||\tilde{H}_{t,M}||_{L^2}\right) \geq 1 - \frac{2^{\frac{n+4}{2}}}{n!\eta}\exp\left(\tfrac{-k}{16\ln 2}\right)$$

Iterating this inequality $\ell > 0$ times fetches

$$\mathbb{P}\Big(||\tfrac{1}{(2k)^l}\textstyle\sum_{s\in\hat{S}^l} \tilde{H}_{t,M}(xs)||_{L^2} \leq \tfrac{1}{2^l}||\tilde{H}_{t,M}||_{L^2}\Big) \geq 1 - \frac{2^{\frac{n+4}{2}}}{n!\eta}\exp\left(\tfrac{-k}{16\ln 2}\right) \tag{19}$$

Let $\tilde{H}_t = H_t - 1_{\mathbb{S}^n}$; then

$$\begin{aligned} ||\tilde{H}_t||_{L^2}^2 &= ||\tilde{H}_{t,M} - \tilde{H}_t||_{L^2}^2 + ||\tilde{H}_{t,M}||_{L^2}^2 \\ &= ||H_{t,M} - H_t||_{L^2}^2 + ||\tilde{H}_{t,M}||_{L^2}^2 \end{aligned}$$

Hence, using lemma 1 and 2 in inequality (19), one derives

$$\mathbb{P}\left(||\tfrac{1}{(2k)^\ell}\sum_{s\in\hat{S}^\ell}\tilde{H}_t(xs)||_{L^2}\leq(2^{-\ell}t^{-\frac{n}{2}}+\eta)\right)\geq 1-\frac{2^{\frac{n+4}{2}}}{n!\eta}\exp\left(\tfrac{-k}{16\ln 2}\right) \quad (20)$$

Since $H_t(xs)=H_t(x_0s^{-1},x)$ and $\hat{S}$ is inverse-symmetric, this produces (18).

The following theorem has appeared in [16]:

Theorem 3 (Nowak-Sjögren-Szarek) *Let $H(t,x_0,x)$ be the heat kernel on the sphere $\mathbb{S}^n$, corresponding to Brownian motion initiated at $x_0\in\mathbb{S}^n$. Let $n\geq 1$ and fix $T>0$. Let $\phi(x):=\arccos\langle x,x_0\rangle$ be the Riemannian distance, so that $\phi(x)\in[0,\pi]$. Then, for all $0<t\leq T$, the inequality*

$$\frac{c}{(t+\pi-\phi)^{\frac{n-1}{2}}t^{\frac{n}{2}}}\exp\left(-\frac{\phi(x)^2}{4t}\right)\leq H(t,x_0,x)\leq\frac{C}{(t+\pi-\phi)^{\frac{n-1}{2}}t^{\frac{n}{2}}}\exp\left(-\frac{\phi(x)^2}{4t}\right)$$

holds for some constants $c,C>0$ depending only on n and T. □

Let $\varepsilon,\eta\in(0,1)$ and let $x\in\mathbb{S}^n$ be such that $\phi(x)>2\varepsilon\sqrt{\ln\frac{1}{\eta\varepsilon^{2n-1}}}$. Then, taking $T=1$ in the above upperbound, one has

$$\begin{aligned}H_t(x)&\leq C_{S^n}t^{-n+\frac{1}{2}}\exp\left(-\tfrac{\phi(x)^2}{4\varepsilon^2}\right)\\&\leq C_{\mathbb{S}^n}\eta\end{aligned}$$

Notice that the constant $C_{\mathbb{S}^n}$ is independent of the initial point $x_0\in\mathbb{S}^n$. Letting $C_n:=1+C_{\mathbb{S}^n}$, we have

$$\begin{gathered}H_{\varepsilon^2}(x)<C_n\eta \qquad \forall\,\varepsilon\in(0,1),\\ \forall\,x\in\mathbb{S}^n \text{ s.t. } r(\varepsilon,\eta):=2\varepsilon\sqrt{\ln\frac{1}{\eta\varepsilon^{2n-1}}}<\phi(x)\end{gathered} \quad (21)$$

Lemma 3 *Let $\varepsilon\in(0,c)$ where $c=(n+4)^{-1}$. If $r=2\varepsilon\sqrt{\ln\frac{3C_n}{\varepsilon^{2n-1}}}$ is sufficiently small, then the following inequality implies that $x_0\hat{S}^l\subseteq\mathbb{S}^n$ is an r-net:*

$$||1_{\mathbb{S}^n}-\frac{1}{(2k)^l}\sum_{s\in\hat{S}^l}H_t(x_0s,x)||_{L^2}\leq\frac{r^{\frac{n}{2}}}{3}\sqrt{\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}+1\right)}} \quad (22)$$

Proof Let $0<\eta<(3C_n)^{-1}$; then, for any $\varepsilon\in(0,1)$, and all $x\in\mathbb{S}^n$ satisfying the inequality

$$d(x,x_0s)>r(\varepsilon,\eta),$$

it follows from (21), and positivity of the heat kernel, that $0<H_t(x_0s,x)<\frac{1_{\mathbb{S}^n}}{3}$, and (hence)

$$\frac{2_{\mathbb{S}^n}}{3}\leq 1_{\mathbb{S}^n}-H_t(x_0s,x)\leq 1_{S^n} \quad (23)$$

Write $B(x_0 s, r) \subseteq \mathbb{S}^n$ for the Riemannian disk of radius r, centered at $x_0 s \in \mathbb{S}^n$. Let $B_n \subseteq \mathbb{R}^n$ denote the unit euclidean ball; one has (see [8])

$$\lim_{r\to 0} \frac{\sigma(B(x_0 s, r))}{r^n} = \frac{\mathrm{vol}(B_n)}{\mathrm{vol}(\mathbb{S}^n)} = \frac{1}{2\sqrt{\pi}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)+1} \tag{24}$$

Here "vol" denotes the standard Lebesgue volume. Now suppose, if possible, that (22) is satisfied, and yet, $x_0\hat{S}^l \subseteq \mathbb{S}^n$ is not an r-net, so that there is $x \in \mathbb{S}^n$ such that $d(x, x_0\hat{S}^l) > r$. Writing

$$\alpha_n := \sqrt{\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)+1}},$$

we derive from (22) and (23)

$$\begin{aligned}
\frac{r^{\frac{n}{2}}\alpha_n}{3} &\geq ||1_{\mathbb{S}^n} - \frac{1}{(2k)^l}\sum_{s\in\hat{S}} H_t(x_0 s, x)||_{L^2} \\
&= ||\frac{1}{(2k)^l}\sum_{s\in\hat{S}} \left(1_{\mathbb{S}^n} - H_t(x_0 s, x)\right)||_{L^2} \\
&\geq \frac{2}{3}\left(\int_{B(x_0 s, r)} d\sigma(x)\right)^{\frac{1}{2}}
\end{aligned}$$

which produces

$$\begin{aligned}
\frac{\sigma(B(x_0 s, r))}{r^n} &= \frac{1}{r^n}\int_{B(x_0 s, r)} d\sigma(x) \\
&\leq \frac{\alpha_n^2}{4}
\end{aligned}$$

Considering (24), this is impossible if $r > 0$ is sufficiently small.

Theorem 4 *Let $\varepsilon \in (0, \frac{1}{3n})$ be small, and $r = 2\varepsilon\sqrt{\ln\frac{3C_n}{\varepsilon^{2n-1}}}$. Let $S \subset SO_{n+1}$ consist of k iid random points, drawn from the Haar measure on SO_{n+1}, where*

$$k \geq 8\ln 2\left((n+4) + 2\ln\left(\frac{1}{\delta}\right) + 6n(1+a_n)\ln\left(\frac{1}{\varepsilon}\right) - \ln(n!)\right),$$

with $a_n := \frac{2\log_2\log_2(5n)}{\log_2(5n)}$. Let $\ell = \frac{n}{2}\log_2\left(\frac{1}{r\varepsilon}\right) + (4+3a_n)n\log_2\left(\frac{1}{\varepsilon}\right)$; if r is sufficiently small then the probability that $x_0\hat{S}^\ell \subseteq \mathbb{S}^n$ is an r-net in $\mathbb{S}^n$ is at least $1-\delta$.

Proof One sees by the remark 1 (following lemma 1) that for any $\eta > 0$, if

$$k_0 > \max\left(\log_2\frac{1}{\eta}, \left(1+a_n\right)\frac{3n}{2}\log_2\left(\frac{1}{t}\right)\right)$$

and $M = 4^{\frac{k_0}{n}}$, then the following inequality holds:

$$||H_t - H_{t,M}||_{L^2}^2 \leq \eta^2$$

Let $\eta := \varepsilon^{3n(1+a_n)}$, so that for sufficiently large $\ell > 0$ (to be determined *à la* theorem 1) the parameter $M = \eta^{-\frac{2}{n}}$ ensures

$$\delta = \frac{2^{\frac{n+4}{2}}}{n!\eta} \exp\left(\tfrac{-k}{16\ln 2} \right)$$

Taking logarithm of (25), we find that it suffices to take

$$k \geq 8\ln 2\left((n+4) + 2\ln\left(\tfrac{1}{\delta} \right) + 2\ln\left(\tfrac{1}{\eta} \right) - \ln(n!) \right) \tag{25}$$

Suppose that $\ell > 0$ is large enough so that $2^{-\ell} t^{-\frac{n}{2}} \leq r^{\frac{n}{2}}$; for this to be true, we require $\ell \geq \frac{n}{2} \log_2\left(\frac{1}{r\varepsilon^2} \right)$. We enforce the inequality $2^{-\ell} t^{-\frac{n}{2}} \leq \varepsilon^{3n(1+a_n)}$ by requiring

$$\ell \geq (4 + 3a_n) n \log_2\left(\tfrac{1}{\varepsilon} \right).$$

Threfore, if $\ell = (4+3a_n) n \log_2\left(\frac{1}{\varepsilon} \right) + \frac{n}{2} \log_2\left(\frac{1}{r\varepsilon} \right)$, then $2^{-\ell} t^{-\frac{n}{2}} \leq \min\{r^{\frac{n}{2}}, \eta\}$ holds. Since $\varepsilon > 0$ is small, one has $\alpha_n r^{\frac{n}{2}} > 6\eta = 6\varepsilon^{3n(1+a_n)}$. Thus, by theorem 2, the following inequality holds:

$$\mathbb{P}\left(||1_{\mathbb{S}^n} - \tfrac{1}{(2k)^\ell} \sum_{s \in \hat{S}^\ell} H_t(x_0 s, x)||_{L^2} \right) \leq \frac{r^{\frac{n}{2}} \alpha_n}{3} \geq 1 - \delta. \tag{26}$$

The proof is complete by lemma 3.

4 Equidistribution and Wasserstein Distance

Let (Y,d) be a compact connected Riemannian manifold. Let $C(Y)$ be the Banach space of continuous functions on Y, and $\mathscr{M}(Y)$ its dual — consisting of linear functionals on $C(Y)$ — equipped with weak* topology; recall that, by compactness of Y, every linear functional is bounded, and hence, continuous. Let $\mathscr{M}(Y)$ be the space of all finite Borel measures on Y. By *Reisz-Markov theorem*, there is a bijection $\mathscr{M}(Y) \cong \mathscr{M}(Y)$, defined by

$$\mu \mapsto (f) \mapsto \int_Y f \, d\mu$$

that is closed under addition and scalar multiplication. The space $\mathscr{M}(Y)$ inherits the sequential (weak) topology on $\mathscr{M}(Y)$ via this bijection. Thus, one says $\mu_n \Rightarrow \mu$ if and only if

$$\int_Y f \, d\mu_n \to \int_Y f \, d\mu$$

for every $f \in C(Y)$. Since the Lipschitz functions are dense in $C(Y)$, it suffices to consider only the 1-Lipschitz functions (*i.e.*, functions $\phi \in C(Y)$ satisfying $|\phi(x) -$

$\phi(x_0)| \leq d(x, x_0)$ for all $x, x_0 \in \mathbb{S}^n$) in the above limit.

For probability measures $\mu, \nu \in \mathscr{M}(Y)$, the *Prokhorov distance* $d_P(\mu, \nu) \geq 0$ is defined to be

$$d_P(\mu, \nu) = \inf\{\varepsilon > 0 : \mu(B) \leq \nu(B_\varepsilon) + \varepsilon \ \forall\, B \in \mathscr{B}(Y)\}$$

where $B_\varepsilon := \{y \in Y : \exists\, b \in B,\ d(b, y) < \varepsilon\}$. This gives a metric on the convex subspace $\mathscr{P}(Y) \subset \mathscr{M}(Y)$ of probability measures on Y, and — by *Prokhorov's theorem* — the induced metric topology on $\mathscr{P}(Y)$ is the subspace of the weak topology on $\mathscr{M}(Y)$; moreover, the space $\mathscr{P}(Y)$ is compact.

We recall that, in a Riemannian manifold (Y, d), the 1-*Wasserstein distance* between two regular Borel probability measures μ and ν on X is defined to be

$$W_1(\mu, \nu) := \inf_{\lambda \in \Pi(\mu, \nu)} \int_{Y \times Y} d(x, y)\, d\lambda$$

where $\Pi(\mu, \nu)$ is the space of couplings of μ and ν; that is, $\Pi(\mu, \nu)$ is the space of all regular Borel probability measures on $Y \times Y$ such that the following holds: $\lambda \in \Pi(\mu, \nu)$ if and only if for every Borel set $B \in \mathscr{B}(Y)$, one has

$$\lambda(Y \times B) = \mu(B) \quad \text{and} \quad \lambda(B \times Y) = \nu(B)$$

Let $\mathrm{Lip}_1(Y)$ be the space of all 1-Lipschitz functions on Y; we recall that, for any $c > 0$, one says $f \in \mathrm{Lip}_c(Y)$ if and only if $|f(x) - f(y)| \leq c \cdot d(x, y)$ for all $x, y \in Y$. The following duality theorem first appeared in [12].

Theorem 5 (Kantorovič - Rubinšteín) *For any $\mu, \nu \in \mathscr{P}(Y)$, the following equality holds:*

$$W_1(\mu, \nu) = \sup_{\phi \in Lip_1(Y)} \left(\textstyle\int_Y \phi \, d\mu - \int_Y \phi \, d\nu \right) \tag{27}$$

Definition 2 Let $\varepsilon > 0$. Let μ be a Borel probability measure on (Y, d). A finite nonempty subset $U \subset Y$ is said to be *strongly* (μ, ε)*-equidistributed* if the following inequality holds:

$$\sup_{\phi \in C(Y)} \left(\textstyle\int_Y \phi \, d\mu - \frac{1}{|U|} \sum_{y \in U} \phi(y) \right) < \varepsilon ||\phi||_{C(Y)}$$

As mentioned before, for $U \subset Y$ to be strongly (μ, ε)-equidistributed, it suffices to have a constant $c := c(\mu)$ such that

$$\sup_{\phi \in \mathrm{Lip}_1(Y)} \left(\textstyle\int_Y \phi \, d\mu - \frac{1}{|U|} \sum_{y \in U} \phi(y) \right) < c\varepsilon \tag{28}$$

Below we show *strong* (μ, ε)-equidistribution of a subset of $\mathbb{S}^n$ of appropriate size and low degree of randomness.

The following lemma will be useful in course of proving the main theorem of this subsection.

Lemma 4 (Fourier convergence) *Fix $y \in \mathbb{S}^n$, and let $0 < a < b$; then the Fourier-Laplace expansion of $H_t(y,x)$ converges uniformly to $H_t(y,x)$ in $[a,b] \times \mathbb{S}^n$.*

Proof Fix an orthonormal basis $\phi_{1,k}, \cdots, \phi_{h_k,k}$ for the eigenspace $H_k(\mathbb{S}^n)$. Then the Fourier-Laplace expansion of the heat kernel based at $y \in \mathbb{S}^n$ is

$$H_t(y,x) = \sum_{k=0}^{\infty} e^{-\lambda_k t} \sum_{i=1}^{h_k} \phi_{i,k}(x)\phi_{i,k}(y).$$

Write

$$\alpha_k(x) := \sum_{i=1}^{h_k} \phi_{i,k}(x)\phi_{i,k}(y)$$

so that

$$H_t(y,x) = \sum_{k=0}^{\infty} e^{-\lambda_k t} \alpha_k(x).$$

One has

$$\begin{aligned} ||\alpha_k(x)||_{L^2}^2 &= (\textstyle\int)_{\mathbb{S}^n} \alpha_k^2(x)\, d\sigma(x) \\ &= \sum_{i=1}^{h_k} \phi_{i,k}(y)^2 \int_{\mathbb{S}^n} \phi_{i,k}(x)^2\, d\sigma(x) \\ &= \sum_{i=1}^{h_k} \phi_{i,k}(y)^2 \\ &= h_k \end{aligned}$$

by rotation invariance of the sum (see *Lemma 2.19 and Theorem 2.29*, [14])

$$\sum_{i=1}^{h_k} \phi_{i,k}(y)^2$$

Therefore,

$$\begin{aligned} ||\alpha_k(x)||_{C^0(\mathbb{S}^n)} &\leq \sqrt{h_k}||\alpha_k||_{L^2} \\ &= h_k \end{aligned}$$

and since

$$\begin{aligned} h_k &= \binom{n+k}{n} - \binom{n+k-2}{n} \\ &\leq 2\binom{n+k-1}{n} \\ &\leq 2e^n \left(\tfrac{n+k-1}{n}\right)^n \\ &\leq 2e^{n+k-1}, \end{aligned}$$

this forces

$$\begin{aligned}\sup_{[a,b]\times\mathbb{S}^n}\sum_{k=0}^{\infty}\left(e^{-\lambda_k t}\alpha_k(x)\right) &\leq \sum_{k=0}^{\infty} e^{-\lambda_k a}h_k\\ &\leq \sum_{k=0}^{\infty} 2e^{(1-ka)(n+k-1)}\\ &< \infty.\end{aligned}$$

The Weierstrass' $M-$test implies uniform convergence of $H_t(x,y)$, to a continous function on $[a,b]\times\mathbb{S}^n$; the claim follows by uniqueness of the continuous limit.

Lemma 5 *Let $d(\cdot,\cdot)$ be the Riemannian distance on $\mathbb{S}^n$. Let σ be the uniform surface probability measure on $\mathbb{S}^n$. For all $t>0$, one has*

$$\int_{\mathbb{S}^n} d(x_0,x)^2\, H_t(y,x)\, d\sigma(x) \leq \frac{\pi^2 nt}{4} \tag{29}$$

Proof We write $H_t(x) := H_t(x_0,x)$, and let ν_t^* be the Borel measure whose Radon-Nikodym derivative is

$$\frac{d\nu_t^*}{d\sigma} = H_t(x)$$

Note that, on standard $\mathbb{S}^n$, for any $x\neq x_0$ one has

$$\frac{d(x_0,x)}{||x-x_0||} = \frac{\arccos\langle x_0,x\rangle}{||x_0-x||} = \frac{\frac{\theta}{2}}{\sin(\frac{\theta}{2})} \leq \frac{\pi}{2} \tag{30}$$

Therefore, it suffices to show that $\mathbb{E}||X-x_0||^2 \leq nt$ for random variable $X\sim\nu_t^*$. Without loss of generality we now assume that $\mathbb{S}^n$ is embedded in $\mathbb{R}^{n+1}$ as the unit sphere with center at $-\mathbf{e}_{n+1} = (0,\cdots,0,-1)$, and $x_0 = \mathbf{0}$.

Let $\{X_u \mid u\in[0,t]\}$ be a Brownian motion on $\mathbb{S}^n$, starting at $x_0\in\mathbb{S}^n$, with infinitesimal generator $H_{\frac{t}{2}}(x)$. For each positive integer $m>0$, consider the equi-partition

$$0 = t_0 < t_1 < \cdots < t_m = t$$

where $t_{i+1}-t_i = m^{-1}t$ for $i = 0,1,\cdots,m-1$. Now define $\{X_i^{(m)}\}_{i=0}^m$ as follows:

$$X_i^{(m)} = X_{\frac{it}{m}}$$

By linearity of expectation, for any integer $m>0$ one has

$$\begin{aligned}\mathbb{E}(||X_m^{(m)}||^2) &= \int_{S^n}||x||^2\, H_t(x)\, d\sigma(x)\\ &= \mathbb{E}(||X_{m-1}^{(m)}||^2) + 2\mathbb{E}(\langle X_m^{(m)}-X_{m-1}^{(m)}, X_{n-1}^{(m)}\rangle) + \mathbb{E}(||X_m^{(m)}-X_{m-1}^{(m)}||^2)\\ &= 2\sum_{i=1}^{m}\mathbb{E}(\langle X_i^{(m)}-X_{i-1}^{(m)}, X_{i-1}^{(m)}\rangle) + \sum_{i=1}^{m}\mathbb{E}(||X_i^{(m)}-X_{i-1}^{(m)}||^2)\end{aligned} \tag{31}$$

Fix a realization of the Brownian motion $\{X_u \mid u \in [0,t]\}$. For integer $1 \le i \le m$, we consider the tangent space $T_{X_{i-1}^{(m)}}(\mathbb{S}^n)$. Write

$$Z_{i-1}^{(m)} := \operatorname{argmin}_{z \in T_{X_{i-1}^{(m)}}(\mathbb{S}^n)} ||z||, \qquad Y_i^{(m)} := \operatorname{argmin}_{z \in T_{X_{i-1}^{(m)}}(\mathbb{S}^n)} ||z - X_i^{(m)}||$$

In explicit terms, one has

$$\begin{aligned} Y_i^{(m)} &= X_i^{(m)} - \langle \mathbf{e}_{n+1} + X_{i-1}^{(m)}, X_i^{(m)} - X_{i-1}^{(m)} \rangle (\mathbf{e}_{n+1} + X_{i-1}^{(m)}) \\ Z_{i-1}^{(m)} &= \langle \mathbf{e}_{n+1} + X_{i-1}^{(m)}, X_{i-1}^{(m)} \rangle (\mathbf{e}_{n+1} + X_{i-1}^{(m)}) \end{aligned}$$

Orthogonality relations such as

$$Y_i^{(m)} - X_{i-1}^{(m)} \perp Z_{i-1}^{(m)}, \qquad \text{and} \qquad X_i^{(m)} - Y_i^{(m)} \perp X_{i-1}^{(m)} - Z_{i-1}^{(m)}$$

are immediate; moreover, one has

$$\begin{aligned} \langle X_i^{(m)} - Y_i^{(m)}, Z_{i-1}^{(m)} \rangle &= \langle \mathbf{e}_{n+1} + X_{i+1}^{(m)}, X_{i-1}^{(m)} \rangle \langle \mathbf{e}_{n+1} + X_{i+1}^{(m)}, X_i^{(m)} - X_{i-1}^{(m)} \rangle \\ &= \langle \mathbf{n}, X_{i-1}^{(m)} \rangle \langle \mathbf{n}, X_i^{(m)} - X_{i-1}^{(m)} \rangle \end{aligned}$$

where $\mathbf{n} = -\mathbf{e}_{n+1} - X_{i+1}^{(m)}$ is the unit normal to $T_{X_{i-1}^{(m)}(\mathbb{S}^n)}$ pointing inward. From the inequalities

$$\langle \mathbf{n}, X_{i-1}^{(m)} \rangle \le 0 \le \langle \mathbf{n}, X_i^{(m)} - X_{i-1}^{(m)} \rangle,$$

one has $\langle X_i^{(m)} - Y_i^{(m)}, Z_{i-1}^{(m)} \rangle \le 0$. Hence,

$$\begin{aligned} \langle X_i^{(m)} - X_{i-1}^{(m)}, X_{i-1}^{(m)} \rangle &= \langle Y_i^{(m)} - X_{i-1}^{(m)}, Z_{i-1}^{(m)} \rangle + \langle X_i^{(m)} - Y_i^{(m)}, X_{i-1}^{(m)} - Z_{i-1}^{(m)} \rangle \\ &\quad + \langle X_i^{(m)} - Y_i^{(m)}, Z_{i-1}^{(m)} \rangle + \langle Y_i^{(m)} - X_{i-1}^{(m)}, X_{i-1}^{(m)} - Z_{i-1}^{(m)} \rangle \\ &\le \langle Y_i^{(m)} - X_{i-1}^{(m)}, X_{i-1}^{(m)} - Z_{i-1}^{(m)} \rangle \end{aligned}$$

Suppose $X_{i-1}^{(m)} = z$ and $X_i = z'$ in $\mathbb{S}^n$. Let $z'' := z''(z,z') \in \mathbb{S}^n$ be such that

$$z'' + z' - 2z = \langle \mathbf{n}, z'' + z' - 2z \rangle \mathbf{n}. \tag{32}$$

Since the function

$$g(z'') := ||z'' + z' - 2z - \langle \mathbf{n}, z'' + z' - 2z \rangle \mathbf{n}||$$

takes arbitrarily small positive values, such a point $z'' \in \mathbb{S}^n$ — that satisfies (32) — exists by continuity of $g(z'')$ and compactness of $\mathbb{S}^n$. Note that

$$\begin{aligned} &\mathbb{P}\{Y_i^{(m)} - X_{i-1}^{(m)} = z'' - z - \langle \mathbf{n}, z'' - z \rangle \mathbf{n} \mid X_{i-1}^{(m)} = z\} \\ = \;&\mathbb{P}\{Y_i^{(m)} - X_{i-1}^{(m)} = z' - z - \langle \mathbf{n}, z' - z \rangle \mathbf{n} \mid X_{i-1}^{(m)} = z\} \end{aligned}$$

by independence of increments for Brownian motion on euclidean space. From

$$\begin{aligned} & \langle Y_i^{(m)} - X_{i-1}^{(m)}, X_{i-1}^{(m)} - Z_{i-1}^{(m)} \rangle \,|_{X_{i-1}^{(m)}=z, X_i^{(m)}=z'} \\ = & \ -\langle Y_i^{(m)} - X_{i-1}^{(m)}, X_{i-1}^{(m)} - Z_{i-1}^{(m)} \rangle \,|_{X_{i-1}^{(m)}=z, X_i^{(m)}=z''} \end{aligned}$$

we derive

$$\mathbb{E}(\langle Y_i^{(m)} - X_{i-1}^{(m)}, X_{i-1}^{(m)} - Z_{i-1}^{(m)} \rangle) = 0$$

It thus suffices to prove

Lemma 6

$$\lim_{m\to\infty} \sum_{i=1}^{m} \mathbb{E}(||X_i^{(m)} - X_{i-1}^{(m)}||^2) = nt.$$

Proof We will use the stereographic projection of $\mathbb{S}^n \setminus \{0\}$ onto $\mathbb{R}^n$. It can be shown (see for example [7]) that the image Y_t of a standard Brownian motion on $\mathbb{S}^n$ via the stereographic projection onto $\mathbb{R}^n$, where $r = |Y_t|$, with $\Delta_{\mathbb{R}^n}$ being the Laplacian on $\mathbb{R}^n$, has an infinitesimal generator $(1/2)\Delta_{\mathbb{S}^n}$ that satisfies

$$\Delta_{\mathbb{S}^n} = \left(\frac{1+r^2}{2}\right)^2 \Delta_{\mathbb{R}^n} - (n-2)\left(\frac{r(1+r^2)}{2}\right)\frac{\partial}{\partial r}.$$

Applying this to the function $f(x) = \|x\|^2$, we see that

$$\begin{aligned} \lim_{t\to 0} \mathbb{E}^0 Y_t^2/t &= (1/2)\Delta_{\mathbb{S}^n} r^2|_{r=0} \\ &= (1/2)\left(\frac{1+r^2}{2}\right)^2 \Delta_{\mathbb{R}^n}(r^2)|_{r=0} - (n-2)\left(\frac{r(1+r^2)}{4}\right)\frac{\partial}{\partial r}(r^2)|_{r=0} \\ &= n. \end{aligned}$$

It follows that for any i,

$$m\mathbb{E}(||X_i^{(m)} - X_{i-1}^{(m)}||^2),$$

converges as $m \to \infty$ to nt, proving the lemma.

Lemma 7 *Let $n > 1$ and for integers $k \geq 0$, let $P_{k,n}(t)$ be the Legendre polynomial of degree k and dimension $n+1$. Let $h_k = \dim H_k(\mathbb{S}^n)$ and*

$$\gamma_k = \int_{-1}^{1} (1-t)^{\frac{1}{2}} (1-t^2)^{\frac{n-2}{2}} P_{k,n}(t)\, dt$$

Then the following inequality holds for all $n \geq 4$:

$$\sum_{k=0}^{\infty} e^{-\lambda_k t} \gamma_k h_k \leq \sqrt{nt}. \tag{33}$$

Proof We recall the *Hecke-Funk formula*: for any function $\chi : [-1,1] \to \mathbb{R}$, which satisfies the inequality

$$\int_{[-1,1]} |\chi(t)|(1-t^2)^{\frac{n-2}{2}}\, dt < \infty,$$

and any eigenfunction $\phi \in \mathscr{H}_k(\mathbb{S}^n)$ and point $y \in \mathbb{S}^n$ one has

$$\int_{\mathbb{S}^n} \chi(y\cdot x)\phi(x)\, d\sigma(x) = \frac{\Omega_{n-1}}{\Omega_n}\phi(y)\int_{-1}^{1} \chi(t)(1-t^2)^{\frac{n-2}{2}} P_k^n(t)\, dt$$

Consider the function $\chi(t) = \sqrt{2}(1-t)^{\frac{1}{2}}$, taking values in $[0,2]$; this satisfies the hypothesis in Hecke-Funk formula, and since $||y-x|| = \sqrt{2}(1-\langle y,x\rangle)^{\frac{1}{2}}$, one gets

$$\int_{\mathbb{S}^n} ||y-x||\phi(x)\, d\sigma(x) = \frac{\sqrt{2}\Omega_{n-1}}{\Omega_n}\phi(y)\int_{-1}^{1} (1-t)^{\frac{1}{2}}(1-t^2)^{\frac{n-2}{2}} P_{k,n}(t)\, dt$$

Consider the Fourier- Laplace expansion of the heat kernel, as in lemma 4 above. By uniform convergence (lemma 4) of the Fourier-Laplace expansion of the heat kernel, one has

$$\begin{aligned}
\int_{\mathbb{S}^n} d(y,x)\; H_t(y,x)\; d\sigma(x) &= \sum_{k=0}^{\infty} e^{-\lambda_k t}\,(\textstyle\sum)_{i=1}^{h_k}\, \phi_{i,k}(y)\int_{\mathbb{S}^n} d(y,x)\; \phi_{i,k}(x)\; d\sigma(x) \\
&= \frac{\sqrt{2}\Omega_{n-1}}{\Omega_n}\sum_{k=0}^{\infty} e^{-\lambda_k t}\gamma_k\,(\textstyle\sum)_{i=1}^{h_k}\, \phi_{i,k}^2(y) \\
&= \frac{\sqrt{2}\Omega_{n-1}}{\Omega_n}\sum_{k=0}^{\infty} e^{-\lambda_k t}\gamma_k h_k
\end{aligned}$$

Note that

$$\begin{aligned}
\frac{\sqrt{2}\Omega_{n-1}}{\Omega_n} &= \sqrt{\frac{2}{\pi}}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \\
&\geq 1
\end{aligned}$$

for all $n \geq 4$. Hence, Lemma 5 together with Hölder inequality implies

$$\begin{aligned}
\sqrt{nt} &\geq \int_{\mathbb{S}^n} d(y,x)\; H_t(y,x)\; d\sigma(x) \\
&\geq \sum_{k=0}^{\infty} e^{-\lambda_k t}\gamma_k h_k
\end{aligned}$$

Theorem 6 *For $n > 1$, let σ be the probability measure on $\mathbb{S}^n$ corresponding to Haar probability measure on the group of rotations SO_{n+1}. Let $\varepsilon, \delta > 0$ be sufficiently small and $r = 2\varepsilon\sqrt{\ln\frac{3C_n}{\varepsilon^{2n-1}}}$. Let $S \subseteq SO_{n+1}$ be a random subset such that $|S| = k$ satisfies the inequality in theorem 4, namely*

$$k > 8\ln 2\left(\left(\right)n+4\right) + 2\ln\left(\tfrac{1}{\delta}\right) + 6n(1+a_n)\ln\left(\tfrac{1}{\varepsilon}\right) - \ln(n!),$$

where $a_n := \frac{2\log_2\log_2(5n)}{\log_2(5n)}$*. Let* $x_0 \in \mathbb{S}^n$ *and let* ν *be the uniform probability measure on* $\mathbb{S}^n$*, supported on* $\hat{S}^\ell x_0$*, where* $\hat{S} = S \cup S^{-1}$ *as before and*

$$\ell = \frac{n}{2}\log_2\left(\tfrac{1}{r\varepsilon}\right) + (4+3a_n)\log_2\left(\tfrac{1}{\varepsilon}\right).$$

Then, with probability at least $1-\delta$*, the following inequality holds:*

$$W_1(\sigma,\nu) \le 2\sqrt{n}\varepsilon$$

Proof Let $Lip_{1,0}(\mathbb{S}^n)$ be the set of mean-zero Lip_1-functions on $\mathbb{S}^n$. By theorem 5, it suffices to show that

$$\sup_{\phi\in Lip_{1,0}(\mathbb{S}^n)} \left(\textstyle\int\right)_Y \phi\, d\sigma - \int_Y \phi\, d\nu < \varepsilon \tag{34}$$

We note that, for any such function $\phi \in Lip_{1,0}(\mathbb{S}^n)$, if $\phi(x_0) = ||\phi||_{L^\infty}$ then

$$\begin{aligned}
0 &= \int_{\mathbb{S}^n} \phi(x)\, d\mu(x) \\
&= \int_{\mathbb{S}^n} \phi(x_0)\, d\mu(x) + \int_{\mathbb{S}^n} \left(\phi(x)-\phi(x_0)\right)\, d\mu(x) \\
&= \phi(x_0) + \int_{\mathbb{S}^n} \left(\phi(x)-\phi(x_0)\right)\, d\mu(x) \\
\Rightarrow \qquad \phi(x_0) &\le \int_{\mathbb{S}^n} |\phi(x)-\phi(x_0)|\, d\mu(x) \\
&\le \int_{\mathbb{S}^n} d(x,x_0)\, d\mu(x) \\
&\le \pi
\end{aligned}$$

For sufficiently small $t>0$, we let ν_t^* be the Borel probability measure on $\mathbb{S}^n$ whose density is

$$\frac{d\nu_t^*(x)}{d\sigma} = \frac{1}{|\hat{S}^l|}\sum_{y\in\hat{S}^l x_0} H_t(y,x)$$

Then, for $t=\varepsilon^2$, one has

$$\begin{aligned}
W_1(\sigma,\nu_t^*) &= \sup_{\phi\in Lip_{1,0}(\mathbb{S}^n)} \left(\textstyle\int_{\mathbb{S}^n}\phi(x)\,d\sigma(x) - \int_{\mathbb{S}^n}\phi(x)\,d\nu_t^*(x)\right) \\
&\le \sup_{\phi\in Lip_{1,0}(\mathbb{S}^n)} \int_{\mathbb{S}^n} |\phi(x)|\cdot\left(1_{\mathbb{S}^n} - \tfrac{1}{(2k)^l}\textstyle\sum_{y\in\hat{S}^l x_0} H_t(y,x)\right) d\sigma(x) \\
&\le \sup_{\phi\in Lip_{1,0}(\mathbb{S}^n)} ||\phi||_{L^\infty}\cdot\int_{\mathbb{S}^n} \left(1_{\mathbb{S}^n} - \tfrac{1}{(2k)^l}\textstyle\sum_{y\in\hat{S}^l x_0} H_t(y,x)\right) d\sigma(x) \\
&\le \pi\varepsilon^{3n} \qquad\qquad \text{(see Theorem 4)}
\end{aligned} \tag{35}$$

with probability at least $1-\delta$.

For any function $\phi \in Lip_{1,0}(\mathbb{S}^n)$, define $\tilde{\phi}_t : \mathbb{S}^n \to \mathbb{R}$ to be

$$\tilde{\phi}_t(x) = \frac{1}{|\hat{S}^\ell|} \sum_{y \in \hat{S}^\ell x_0} \phi(y) H_t(y,x)$$

From uniform convergence of the Fourier-Laplace expansion of heat-kernel, it follows that $\int_{\mathbb{S}^n} H_t(y,x)\, d\sigma(x) = 1$; hence, putting $t = \varepsilon^2$, one has

$$\begin{aligned}\int_{\mathbb{S}^n} \tilde{\phi}_{\varepsilon^2}(x)\, d\sigma(x) &= \frac{1}{|\hat{S}^\ell|} \sum_{y \in \hat{S}^\ell x_0} \phi(y) \int_{\mathbb{S}^n} H_t(y,x)\, d\sigma(x) \\ &= \int_{\mathbb{S}^n} \phi(x)\, d\nu(x). \end{aligned} \tag{36}$$

Moreover,

$$\begin{aligned} & |\int_{\mathbb{S}^n} \phi(x)\, d\nu_t^*(x) - \int_{\mathbb{S}^n} \tilde{\phi}_t(x)\, d\sigma(x)| \\ = \; & \frac{1}{(2k)^\ell} | \sum_{y \in \hat{S}^l x_0} \int_{\mathbb{S}^n} (\, \phi(x) - \phi(y)\,) H_t(y,x)\, d\sigma(x)| \\ \leq \; & \frac{1}{(2k)^\ell} \sum_{y \in \hat{S}^l x_0} \int_{\mathbb{S}^n} |\phi(x) - \phi(y)|\, H_t(y,x)\, d\sigma(x) \\ \leq \; & \frac{1}{(2k)^\ell} \sum_{y \in \hat{S}^l x_0} \int_{\mathbb{S}^n} d(y,x)\, H_t(y,x)\, d\sigma(x) \\ \leq \; & \sqrt{nt} \end{aligned} \tag{37}$$

by lemma 5 and Hölder inequality applied to $d(y,x) = d(y,x) \cdot 1_{\mathbb{S}^n}$ while integrating with respect to $H_t(y,x)\, d\sigma(x)$. Therefore, for $t = \varepsilon^2 > 0$ sufficiently small, equations (35), (36), and (37) yield

$$\begin{aligned} W_1(\sigma, \nu) &\leq W_1(\mu, \nu_t^*) + W_1(\nu_t^*, \nu) \\ &\leq \pi \varepsilon^{3n} + \sqrt{n}\varepsilon \\ &\leq 2\sqrt{n}\varepsilon \end{aligned}$$

Lemma 8 *With assumptions as in theorem 6, the following inequality holds:*

$$W_p(\sigma, \nu) \leq 4\pi \sqrt[2p]{n} \varepsilon^{\frac{1}{p}}$$

Proof We write ν_t^* for the probability measure on $\mathbb{S}^n$ whose Radon-Nikodym derivative is

$$\frac{d\nu_t^*}{d\sigma} = \frac{1}{|\hat{S}|^l} \sum_{y \in \hat{S}^l x_0} H_t(y,x)$$

Let $(X,Y) \sim \gamma$, where $\gamma \in \Gamma(\sigma, \nu_t^*)$ is a coupling that realizes $W_1(\sigma, \nu_t^*)$. Such a coupling exists, for instance, because

a) the set $\mathscr{P}(\mathbb{S}^n \times \mathbb{S}^n)$ of Borel probability measures on $\mathbb{S}^n \times \mathbb{S}^n$ is sequentially compact in Prohorov metric,

b) $\Gamma(\sigma, \nu_t^*) \subset \mathscr{P}(\mathbb{S}^n \times \mathbb{S}^n)$ is closed (being inverse image of a point under the continuous map that takes a measure to its marginals),
c) the Prohorov metric metrizes the weak topology on $\mathscr{P}(\mathbb{S}^n \times \mathbb{S}^n)$.

One has

$$\begin{aligned} W_p^p(\sigma, \nu_t^*) &\leq \mathbb{E}[||X-Y||^p] \\ &\leq \int_0^{\pi^p} \mathbb{P}[||X-Y||^p \geq t]\, dt \\ &\leq \int_0^{\pi} \mathbb{P}[||X-Y|| \geq s]\, ps^{p-1}\, ds \end{aligned}$$

Applying Hölder's inequality, one obtains

$$\begin{aligned} W_p^p(\sigma, \nu_t^*) &\leq \pi^{p-1} p \left(\textstyle\int\right)_0^\pi \mathbb{P}[||X-Y|| \geq s]\, ds \\ &= \pi^{p-1} p W_1(\sigma, \nu_t^*) \\ \Rightarrow \qquad W_p(\sigma, \nu_t^*) &\leq (\pi^{p-1}p)^{\frac{1}{p}} W_1(\sigma, \nu_t^*)^{\frac{1}{p}} \\ &\leq 2\pi W_1(\sigma, \nu_t^*)^{\frac{1}{p}} \end{aligned}$$

Similarly, one has $W_p(\nu_t^*, \nu) \leq 2\pi W_1(\nu_t^*, \nu)^{\frac{1}{p}}$, forcing

$$\begin{aligned} W_p(\sigma, \nu) &\leq W_p(\sigma, \nu_t^*) + W_p(\nu_t^*, \nu) \\ &\leq 2\pi \left(W_1(\sigma, \nu_t^*)^{\frac{1}{p}} + W_1(\nu_t^*, \nu)^{\frac{1}{p}} \right) \\ &\leq 4\pi n^{\frac{1}{2p}} \varepsilon^{\frac{1}{p}} \end{aligned}$$

5 Conclusion

We proved two results about the finite time behavior of a random Markov Chain on the sphere S^n whose transitions correspond to rotations chosen uniformly at random. The first result states that for $k = O(n \ln \frac{1}{\varepsilon} + \ln \frac{1}{\delta})$ random rotations and $\ell = O(n \ln 1/\varepsilon)$, if one takes the image of the north pole on the sphere under all possible words of length ℓ in the k alphabets and their inverses, one obtains an ε-net with high probability. For these parameters, the value of $(2k)^\ell$ is close to the volumetric lower bound of $(1/\varepsilon)^{\Omega(n)}$ on the size of an ε-net of S^n. Secondly, we show that this ε-net is equidistributed with probability at least $1-\delta$ in the sense that the 1-Wasserstein distance of the uniform measure on the net is within ε of the uniform measure on S^n.

These results can respectively be applied to approximately minimize any given 1- Lipschitz function on the sphere (by evaluation on the ε-net) and to approximately integrate a 1-Lipschitz function on the sphere. In both cases the approximation is within an additive ε of the true value.

Acknowledgements SC would like to thank Sandeep Juneja and Jaikumar Radhakrishnan for helpful conversations. HN was partially supported by a Ramanujan fellowship.

References

1. R. Ahlswede and A. Winter, *Strong converse for identification via quantum channels*, IEEE Trans. Inform. Theory 48 (2002), no. 3, 569–579.
2. N. Alon, and Y. Roichman, *Random Cayley Graphs and Expanders*, Random Structures Algorithms 5 (1994), no. 2, 271–284.
3. Y. Benoist, and N. de Saxcé, *A Spectral Gap Theorem in Simple Lie Groups*, Invent. Math. 205 (2016), no. 2, 337–361.
4. J. Bourgain, *On random walks in large compact Lie groups*, Geometric aspects of functional analysis, 55–63, Lecture Notes in Math. (2169), Springer, Cham, 2017
5. J. Bourgain, and A. Gamburd, *A spectral gap theorem in $SU(d)$*, J. Eur. Math. Soc. (JEMS) 14 (2012), no. 5, 1455–1511.
6. ———, *On the spectral gap for finitely-generated subgroups of $SU(2)$*, Invent. Math. 171 (2008), no. 1, 83–121.
7. T. Carne, *Brownian motion and stereographic projection*, Ann. Inst. H. Poincaré Probab. Statist. 21 (1985), no. 2, 187–196.
8. A. Gray, *The volume of a small geodesic ball of a Riemannian manifold*, Michigan Math. J. 20 (1973), 329–344 (1974).
9. A. Grigor'yan, *Heat kernel and analysis on manifolds*, AMS/IP Studies in Advanced Mathematics (47), American Mathematical Society, Providence; International Press, Boston, ISBN: 978-0-8218-4935-4
10. M. Hastings, and A. Harrow, *Classical and quantum tensor product expanders*, Quantum Info. Comput. 9, 3 (March 2009), 336-360.
11. E. Hsu, *Stochastic analysis on manifolds*, Graduate Studies in Mathematics (38) AMS, Providence, ISBN: 0-8218-0802-8.
12. L. Kantorovič, and G. Rubinšteín, *On a space of completely additive functions* (Russian), Vestnik Leningrad Univ, 13 1958 no 7, 52–59.
13. Z. Landau, and A. Russell, *Random cayley graph expanders: A simple proof of the Alon-Roichman Theorem*, Electron. J. Combin. 11 (2004), no. 1, Research Paper 62, 6 pp.
14. M. Morimoto, *Analytic functionals on the sphere*, Translations of Mathematical Monographs (178), AMS, Providence, ISBN: 0-8218-0585-1.
15. H. Narayanan, *On The Distribution Of Random Words In A Compact Lie Group*, preprint arXiv:1804.07146v1.
16. A. Nowak, P. Sjögren, and T. Szarek, *Sharp estimates of the spherical heat kernel*, to appear in J. Math. Pures Appl., available online 19 October 2018.
17. Ed. Saff, `https://my.vanderbilt.edu/edsaff/spheres-manifolds/`
18. C. Villani, *Optimal transport*, Grundlehren der Mathematischen Wissenschaften (338), Springer-Verlag, Berlin, ISBN: 978-3-540-71049-3.