Kac's random walk on the special orthogonal group mixes in polynomial time

Article *in* Proceedings of the American Mathematical Society · November 2016

DOI: 10.1090/proc/13598

Yunjiang Jiang
Google Inc.

PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 00, Number 0, Pages 000–000
S 0002-9939(XX)0000-0

KAC'S RANDOM WALK ON THE SPECIAL ORTHOGONAL GROUP MIXES IN POLYNOMIAL TIME

YUNJIANG JIANG

(Communicated by Professor Walter Van Assche)

Dedicated to Professor Theodore Shifrin

Abstract. We give the first proof of polynomial total variation mixing time bound for the Kac random walk on the special orthogonal groups. The proof relies on the exact spectral gap computation by E. A. Carlen et al [2] and D. Maslen [7], and hinges on two novel ingredients: a multi-dimensional generalization of Turans's lemma for polynomials on the unit circle, proved by F. L. Nazarov, and a Morse-theoretic result due to J. Milnor. The techniques are robust in the sense that the step distribution of the walk does not have to be uniformly supported on the circles, and that the model can be generalized to higher dimensional particles or other compact Lie groups, provided the corresponding relaxation time is polynomial in the dimension.

1. Introduction

Consider the following probabilistic toy model of Boltzmann gas dynamics. Assume we have n gas particles living in $\mathbb{R}^1$, whose only physically relevant measurements are their velocities, which are represented by the components of an n-dimensional vector $\mathbf{v} = (v_1, \dots, v_n)$. When two particles collide, their velocities are transformed from an old pair to a new pair; the only constraint is that the total kinetic energy is preserved. Thus $v_i(old)^2 + v_j(old)^2 = v_i(new)^2 + v_j(new)^2$, for all $i \neq j$. In particular, after normalization, the n-vector $\mathbf{v}$ lives on the unit sphere in $\mathbb{R}^n$.

To capture the randomness of the collision pairs as well as the transformation of velocity pairs, we introduce the following discrete time Markov model for the evolution of the above collision process: the state of the gas ensemble is represented by a point $\mathbf{p} \in S^{n-1}$. At each step, two distinct coordinates $i \neq j$ are uniformly chosen and $\mathbf{p}$ is rotated along the oriented $i \wedge j$ plane by a uniform angle $\theta \in [0, 2\pi)$. All the random choices made are independent of one another. This is known as the Kac random walk on the sphere. One can also define the corresponding walk (1.1) on the Lie group $SO(n)$ by applying the same random rotation to each of the standard basis vectors in an orthonormal frame; the walk on the sphere then becomes a projected random walk by virtue of the relation $S^{n-1} \cong SO(n)/SO(n-1)$.

Received by the editors December 15, 2013 and, in revised form, November 22, 2016.
2000 *Mathematics Subject Classification.* Primary 60J05; Secondary 22C05.
Key words and phrases. Kac walk, total variation mixing time, Nazarov's inequality.
Research partially supported by NSF graduate fellowship.

By standard Markov chain theory, this random walk converges in distribution to the uniform measure on S^{n-1} resp. $SO(n)$. In fact even the rates of convergence under various norms are fairly well understood. To summarize the current state of affair, the spectral gap of the walk has been determined to be $\frac{1}{2n} + \frac{3}{2n(n-1)}$ in [2] as well as [7] (see also [5] for a Martingale approach to obtain the right order of magnitude). The convergence rate, under the L^2 *Wasserstein distance* induced from the natural Riemmanian metric on $SO(n)$, has been determined to be $\mathcal{O}(n^2 \log n)$ in [10], improving upon previous work on weak-* convergence rate in [3] and [15]. Finally total variation mixing time has been especially difficult to analyze due to the presence of singularities in the density of the running distribution. A bound of order $\mathcal{O}(n^{n^2})$ was initially obtained in [3], using the Doeblin coupling technique. Subsequently, it was improved to $\mathcal{O}(n^5 \log n)$ in [6], but only for the walk on S^{n-1}. Here we present a short argument for polynomial mixing time of the latter walk under the total variation distance.

Let K be the Kac kernel, defined by

$$(Kf)(A) = \frac{1}{\binom{n}{2}} \sum_{i \neq j}^{n} \int_0^{2\pi} f(R(i,j;\theta)A) \frac{1}{2\pi} d\theta, \tag{1.1}$$

for $A \in SO(n)$. Also let ν be the Haar measure on $SO(n)$ and denote by δ_g the point mass at an element $g \in SO(n)$. The main result of this paper is

Theorem 1.1. *For $t = 20n^5(\log n)^2$, the following total variation distance bound can be established:*

$$\|\delta_g K^t - \nu\|_{\mathrm{TV}} \leq \frac{1}{4}.$$

Here g denotes an arbitrary element in $SO(n)$. Also define the relaxation time by $t_{rel} := -\limsup_t \frac{1}{t} \log \|\delta_g K^t - \nu\|_{\mathrm{TV}}$. Then $t_{rel} \leq 2n^5$.

Clearly it suffices to take g in the statement above to be the identity element. Furthermore, by log subadditivity of total variation distance to stationarity, the distance drops by a factor of 2^k when the running time is extended by a factor of 2^k, that is,

$$\|\delta_g K^{2^k t} - \pi\|_{\mathrm{TV}} \leq \frac{1}{4} 2^{-k}.$$

Unlike the spectral gap and Wasserstein distance mixing time, the total variation mixing time of this walk has been especially difficult to analyze due to the presence of singularities in the density of the running distribution. The $\mathcal{O}(n^{n^2})$ bound obtained in [3] is currently the best in the literature. The conjectured mixing time, however, lies somewhere between n^2 and $n^2 \log n$. The lower bound of n^2 is easy to see by dimension consideration: the state space $SO(n)$ has dimension $\binom{n}{2}$ whereas the step distribution of the Kac walk is supported on a circle of dimension 1. The $n^2 \log n$ conjecture is motivated by coupon collector's phenomenon, although the noncommutative nature of the state space could render the latter irrelevant. While the Kac walk on the *sphere* can be approached in much the same way as the simplex walk [16], [11] (up to a measure 0 set, $(x_1, \ldots, x_n) \mapsto (x_1^2, \ldots, x_n^2)$ defines a 2-to-1 map from S^{n-1} to the n-simplex), the walk on $SO(n)$ is more difficult to handle since the state space cannot be identified easily with an open subset of some Euclidean space. The argument below hinges on two novel ingredients: a Morse theoretic result of J. Milnor [8] stating that the number of connected components of a real algebraic variety can be bounded in terms of its dimension and maximum degree, as well as a Remez-type inequality by F. Nazarov [9] stating that a polynomial on the unit circle of bounded degree cannot visit any neighborhood of 0 too often.

After the present paper was submitted, Aaron Smith informed me that he and Natesh Pillai obtained much sharper estimates than the ones presented in this paper, namely that the total variation mixing time is in fact $O(n^2 \log n)$ ([12]). Their proof also relies on the fact that the Jacobian of the structural map (2.1) for the Kac walk on $SO(n)$ is bounded away from zero with high probability, but establishes the latter in a much more careful manner by applying concentration inequality techniques to the smallest singular value of the (random) Jacobian matrix. The lower bound of $\Omega(n^2)$ still stands, and there is strong evidence that it is sharp at least under the Wasserstein distance. It will be interesting to see whether one can push the geometric and harmonic analytic techniques presented here to match their results.

2. Proof of the main theorem

Throughout this section, we use $\|\mu\| = \|\mu\|_{\mathrm{TV}}$ interchangeably to denote the total variation norm of the (possibly signed) measure μ; it is defined by

$$\|\mu\| = \sup_{S \text{ measurable}} |\mu(S)|.$$

Thus if μ is a probability measure, $\|\mu\| = 1$. Also we will denote by $\|\mu - \nu\|_2$ the L^2 norm of the density $\frac{d\mu}{d\nu} - 1$ of μ with respect to the measure ν. It is defined to be ∞ if the density does not have an L^2 norm.

Fix a sequence of rotation indices $(I, J) = (i_1, j_1), \dots, (i_m, j_m)$ and consider the structural map $\phi = \phi_{I,J} : \mathbb{T}^m \to SO(n)$ defined by

$$\phi(\vec{\theta}) = \gamma_{i_1,j_1}(\theta_1) \dots \gamma_{i_m,j_m}(\theta_m), \tag{2.1}$$

where $\gamma_{ij}(\theta)$ is the rotation matrix by angle θ along the $i \wedge j$ plane.

By translating the differential of ϕ to identity, we can compute the generalized Jacobian of ϕ via the following: define $v_1(\theta) = \gamma'_{i_1,j_1}(0)$,

$$v_2(\theta) = \mathrm{Ad}_{(\gamma_{\mathrm{i}_1,\mathrm{j}_1}(\theta_1))}\gamma'_{\mathrm{i}_2,\mathrm{j}_2}(0) := \gamma_{\mathrm{i}_1,\mathrm{j}_1}(\theta_1)\gamma'_{\mathrm{i}_2,\mathrm{j}_2}(0)\gamma_{\mathrm{i}_1,\mathrm{j}_1}(-\theta_1),$$

and in general,

$$v_k(\theta) = \mathrm{Ad}_{(\prod_{\mathrm{s}=1}^{\mathrm{k}-1} \gamma_{\mathrm{i}_\mathrm{s},\mathrm{j}_\mathrm{s}}(\theta_\mathrm{s}))}\gamma'_{\mathrm{i}_\mathrm{k},\mathrm{j}_\mathrm{k}}(0).$$

Then clearly $v_k = R_{\phi(\theta)}[\frac{\partial}{\partial \theta_k}\phi(\theta)]$, where $R_X v$ denotes the right translation of the vector v by $X \in SO(n)$. Since $v_k \in \mathfrak{so}(n)$, they are antisymmetric matrices under the standard basis on $\mathbb{R}^{n^2}$. So we will denote by $v_k[i,j]$, the (i,j)-th entry of v_k for $0 < i < j < n+1$. We also let $u_{i,j} = (v_1[i,j], \dots, v_m[i,j])$, i.e., the row vector in the matrix assembled from the v_k's. The generalized Jacobian function associated with ϕ is defined by

$$J(\theta) := \mathrm{Jac}\phi_\theta = \sqrt{\det(\langle \mathrm{u}_{\mathrm{i},\mathrm{j}}(\theta), \mathrm{u}_{\mathrm{i}',\mathrm{j}'}(\theta)\rangle)}.$$

We will decompose the space $\mathbb{T}^m$ into two parts:

$$A := \{\theta \in \mathbb{T}^m : |J(\theta)| > c\} \qquad \text{and} \qquad B := \mathbb{T}^m \setminus A \text{ (B stands for bad)}. \tag{2.2}$$

The size of $c > 0$ will be chosen later; it will decay super-exponentially with n.

Now consider the probability measure $\mu := \mu_{I,J}$ induced by ϕ on $SO(n)$, in other words, the pushforward measure $\phi_* \pi^m$, where π^m denotes the product uniform probability measure on $\mathbb{T}^m$. We decompose μ into two measures

$$\mu = \mu_a + \mu_b, \tag{2.3}$$

where μ_b is induced by ϕ restricted to B and μ_a induced by ϕ restricted to A. The goal is to choose c so that $\|\mu_b\|$ is small while the density of μ_a with respect to the Haar measure ν on $SO(n)$ is not too big.

Consider the second objective first: by Sard's theorem, the set of critical values of J in $\mathbb{R}$ has measure 0, so there exists an arbitrarily small regular value c^2 of J^2, which means that $(J^2)^{-1}(c^2)$ will be a smooth submanifold of codimension 1 in $\mathbb{T}^m$. This further implies that A is an open smooth submanifold of $\mathbb{T}^m$ and B is a compact smooth submanifold with boundary of $\mathbb{T}^m$. Now consider the map $\phi : A \to SO(n)$ between smooth manifolds. Let $q \in \phi(A)$, which is a regular value of $\phi|_A$ since the Jacobian map is nonvanishing at every preimage point in $\phi^{-1}(q) \cap A$. We will bound the density of $\mu_a := \phi_*(\pi^m|_A)$ with respect to ν at q. By definition of induced measure and density, this is given by

$$\frac{d\mu_a}{d\nu}(q) = \lim_{\epsilon \downarrow 0} \pi^m(\phi^{-1}(B_\epsilon(q)) \cap A)\mathrm{Vol}(\mathrm{SO(n)})/\epsilon^{\binom{\mathrm{n}}{2}}\mathrm{K}_{\binom{\mathrm{n}}{2}},$$

where $K_{\binom{n}{2}}$ is the Euclidean volume of the $\binom{n}{2}$-dimensional unit ball in $\mathbb{R}^{\binom{n}{2}}$ and $\mathrm{Vol(SO(n))}$ is the Euclidean $\binom{n}{2}$-volume of $SO(n)$; the latter can be bounded by the volume of the hypersphere in $\mathbb{R}^{n^2}$ with radius $\sqrt{n}$ which is certainly bounded by $n^{\frac{n^2-1}{2}}$.

By standard arguments in integral geometry, it is clear that the density of μ_a can be computed by the following integral:

$$\frac{d\mu_a}{d\nu}(q) = \int_{\phi^{-1}(q)\cap A} \mathrm{Vol(SO(n))}|\mathrm{J}(\theta)|^{-1}\mathrm{dVol}_\ell(\theta),$$

where $d\mathrm{Vol}_\ell$ denotes the ℓ-dimensional volume element; $\ell = m - \binom{n}{2}$.

Intuitively, the smaller the Jacobian $|J(\theta)|$ is, more mass will go into the point q from a neighborhood of θ under the map ϕ. Now by our choice of A, $|J(\theta)| > c$. Hence we can bound the induced density by

$$\frac{d\mu_a}{d\nu}(q) \le c^{-1}\mathrm{Vol(SO(n))}\mathrm{Vol}_{\mathrm{m}-\binom{\mathrm{n}}{2}}(\phi^{-1}(\mathrm{q}) \cap \mathrm{A})$$

It remains to bound the volume of the preimage hypersurface $A_q := \phi^{-1}(q) \cap A$.

We first embed $\mathbb{T}^m$ into $[-1,1]^{2m}$ in the standard way. Also let $\tilde{\phi}$ be the natural algebraic extension of ϕ to $\mathbb{R}^{2m}$, and similarly all other maps originally defined on $\mathbb{T}^m$ will be extended to $\mathbb{R}^{2m}$ in the algebraically natural way below. Next disassemble $A_q \subset \mathbb{T}^m$ into disjoint union of graphs over subspaces in $\mathbb{R}^{2m}$ of the form $\mathrm{Span}\{x_{s_1}, \ldots, x_{s_\ell}\}$, where $S := \{x_{s_1}, \ldots, x_{s_\ell}\}$ is a subset of the standard basis vectors $x_1, \ldots, x_{2m}$. The way we decompose A_q is by examining the Jacobian of the projection map $P_S : \mathbb{R}^{2m} \to \mathrm{Span}\, S$ restricted to A_q. More precisely, let $u_{i,j}[S^c] = (v_{t_1}(i,j), \ldots, v_{t_{\binom{n}{2}}}(i,j))$ be the row vectors of the matrix consisting of the upper triangular matrix elements of the derivatives of ϕ translated to the identity element, in the directions of $\partial_{t_1}, \ldots, \partial_{t_{m+\binom{n}{2}}}$, the orthogonal complement directions of $\partial_{s_1}, \ldots, \partial_{s_\ell}$ in the tangent space $T_q\mathbb{R}^{2m}$. Define the following function on $[-1,1]^{2m}$:

$$g_\eta^S(x) = \det(\langle u_{i,j}, u_{i',j'}\rangle)^2 - \eta^2 (\prod_{0<i<j\le n} \|u_{i,j}\|^2)^2.$$

The set $A_{q,\eta}^S := \{g_\eta > 0\} \cap A_q$ consists of parts of A_q that project down to $\operatorname{Span} S$ by a local-diffeomorphism with Jacobian greater than η^2. Thus $A_{q,\eta}^S$ is a disjoint union of open submanifolds of A, with number of connected components bounded by the number of components of their boundaries, namely the set $g_\eta^{-1}(0)$. For the latter we quote the following general upper bound on the number of connected components of an algebraic set, in terms of the number of real variables in the ambient space and the maximum degree of the polynomials defining the set:

Theorem 2.1. *([8]; see also [17]) Let $P : \mathbb{R}^k \to \mathbb{R}$ be a polynomial of degree d, then the sum of the Betti numbers of the level set $P^{-1}(0)$ is bounded by*

$$\sum_i H_i(P^{-1}(0)) \le \frac{1}{2} d(2d-1)^{k-1}.$$

In particular, the number of connected components, $H_0(P^{-1}(0))$, is bounded by the same RHS, since the Betti numbers count the dimensions of the homology groups, which are nonnegative.

To apply this result, take $k = 2m$, $d = [4m\frac{n^2}{2}]^2 = 4m^2n^4$, and let

$$P(x_1, \ldots, x_{2m}) = (g_\eta(x))^2 + (x_1^2 + x_2^2 - 1)^2 + \ldots (x_{2m-1}^2 + x_{2m}^2 - 1)^2.$$

To see that the degree of P is bounded by d, simply observe that each $v_{t_k}(i,j)$ is obtained by extracting the (i,j)th matrix element of the product of at most $2m$ matrices, whose entries are linear in the variables $x_1, \ldots, x_{2m}$.

Thus we can bound the volume of $A_{q,\eta}^S$ by

$$\mathrm{Vol}_\ell(\mathrm{A_{q,\eta}}) \le \mathrm{H_0(A_{q,\eta})} \max_{\mathrm{x \in A_{q,\eta}}} \mathrm{Jac_x(P|_{A_q})^{-1}} \le \mathrm{(8m^2n^4)^{2m}}\eta^{-1} \tag{2.4}$$

Furthermore, we have the following result regarding the size of η:

Lemma 2.2. *Given ℓ orthonormal vectors $v_1, \ldots, v_\ell$ in $\mathbb{R}^m$, there is some subset $S \subset [m]$ of size ℓ such that the projection of $v_1, \ldots, v_\ell$ onto the subspace spanned by e_j, $j \in S$, spans a ℓ-dimensional parallelopiped with ℓ-volume at least $\binom{m}{\ell}^{-1/2}$.*

Proof. We first claim the following generalized Pythagorean identity

$$\det\langle v_i, v_j\rangle_{i,j=1}^\ell = \sum_{S\subset[m],|S|=\ell} \det\langle v_i[S], v_j[S]\rangle_{i,j=1}^\ell, \tag{2.5}$$

where $v[S]$ stands for the projection of v onto the span of $\{e_j : j \in S\}$. This follows by expanding out the determinant on both sides using Cayley's formula and interchanging product and summation. On the left hand side we have

$$\sum_{\sigma\in S_\ell} \prod_{i=1}^\ell \sum_{j=1}^n v_{ij}v_{\sigma(i)j} \operatorname{sgn}(\sigma) = \sum_\sigma \operatorname{sgn}(\sigma) \sum_{j_1,\ldots,j_\ell\in[n]} \prod_{i=1}^n v_{ij_i}v_{\sigma(i)j_i}$$

$$= \sum_\sigma \operatorname{sgn}(\sigma) [\sum_{j_1\in[n]} \sum_{j_2\in[n]\setminus\{j_1\}} \cdots \sum_{j_\ell\in[n]\setminus\{j_1,\ldots,j_{\ell-1}\}} \prod_{i=1}^n v_{ij_i}v_{\sigma(i)j_i}],$$

where we used the fact that if $j_s = j_t$ for $1 \le s < t \le \ell$, then the summation over $\sigma \in S_\ell$ will be zero, due to the fact that multiplication by the transposition (s,t) partitions S_ℓ into sets of pairs, each of which cancel since $\operatorname{sgn}(\sigma) = -\operatorname{sgn}(\sigma(s,t))$. On the right hand side of (2.5), we similarly have

$$\sum_{\sigma \in S_\ell} \operatorname{sgn}(\sigma) \sum_{S \subset [m]; |S| = \ell} \sum_{j_1 \in S} \sum_{j_2 \in S \setminus \{j_1\}} \cdots \sum_{j_\ell \in S \setminus \{j_1, \dots, j_{\ell-1}\}} \prod_{i=1}^{\ell} v_{i j_i} v_{\sigma(i) j_i}.$$

This equals the left hand side since each collection of ℓ distinct $j_1, \dots, j_\ell$ lies in a unique subset S of size ℓ.

The conclusion of the lemma now follows from the Pigeon-hole principle and the fact that $\det \langle v_i[S], v_j[S] \rangle_{i,j=1}^{\ell}$ equals the square of the volume of the parallelopiped spanned by $v_1[S], \dots, v_\ell[S]$. $\square$

Applying the lemma, we see that for $\eta := m^{-n^2/4}$, which is less than $\binom{m}{m-\binom{n}{2}}^{-1/2}$, every point in A_q lies in some $A^S_{q,\eta}$, where $S \subset [2m]$ is a subset of size ℓ. Therefore we have the following union bound on the volume of A_q:

$$\begin{aligned}
\mathrm{Vol}_\ell(\mathrm{A_q}) &\le \mathrm{Vol}_\ell(\mathrm{A}^{[\ell]}_{\mathrm{q},\eta}) \binom{2\mathrm{m}}{\ell} \\
&\le (8m^2 n^4)^{2m} \eta^{-1} (2m)^\ell \\
&\le 2^{6m+\ell} m^{4m+\ell+\frac{n^2}{4}} n^{8m} \\
&\le 2^{7m} m^{5m} n^{8m},
\end{aligned}$$

where we recall $\ell = m - \binom{n}{2}$.

Thus we have the following uniform density estimate

$$\frac{d\mu_a}{d\nu} \le 2^{7m} m^{5m} n^{8m} c^{-1}, \tag{2.6}$$

where recall c is the infimum of the Jacobian for the map $\phi : \mathbb{T}^m \to SO(n)$ restricted to A.

We estimate the size of c next. Consider the map $J^2 : \mathbb{T}^m \to \mathbb{R}$ as a multivariate trigonometric polynomial (in particular a function on $\mathbb{C}^m$ restricted to the m-dimensional standard torus), we have the following estimate generalizing the classical Chebyshev's inequality (for polynomials), Remez's inequality, and Turan's lemma. It bounds the ratio between the maximum of a trigonometric polynomial on the circle and the maximum over some small subset in terms of the degree of the polynomial.

Theorem 2.3. *([9] Turan's lemma for polynomials on the unit circumference) For a polynomial $p(z) = \sum_{k=1}^{T} c_k z^{m_k}$ defined on $\mathbb{T} \subset \mathbb{C}$, and any measurable subset $E \subset \mathbb{T}$, we have*

$$\sum_k |c_k| \le \left(\frac{14}{\mu(E)}\right)^{T-1} \sup_{z \in E} |p(z)| \le \left(\frac{5}{\mathbb{P}(E)}\right)^{T-1} \sup_{z \in E} |p(z)|,$$

where μ denotes Lebesgue measure and $\mathbb{P}$ the uniform probability measure on $\mathbb{T}$.

The exact constant 5 is not important. Notice that the left hand side is greater than or equal to $\sup_{z\in\mathbb{T}}|p(z)|$. So we have the following easy consequence:

$$\mathbb{P}(\{z \in \mathbb{T} : |p(z)| < \epsilon\}) \leq 5(\sup_{z\in\mathbb{T}} p(z))^{-\frac{1}{T-1}}\epsilon^{\frac{1}{T-1}}. \tag{2.7}$$

Based on this crucial lemma, we can derive the following multidimensional generalization:

Corollary 2.4. Suppose $p(z_1,\ldots,z_m)$ is a trigonometric polynomial in m variables with $\sup|p(z)| \geq 1$, and $\deg_{z_j} p \leq w$, for each j, then

$$\mathbb{P}(\{z \in \mathbb{T}^m : |p(z)| < \epsilon\}) \leq 5m\epsilon^{\frac{1}{wm}}. \tag{2.8}$$

This is essentially proved by Fontes-Merz in [4]; we include a short proof here for completeness.

Proof. We use induction on the index k in z_k. Without loss of generality, assume $p(0,\ldots,0) = 1$. Fix $\epsilon_1,\epsilon_2,\ldots,\epsilon_m > 0$ to be determined later, and let $A_1 = \{\theta : |p(\theta,0,\ldots,0)| < \epsilon_1\}$. Then by Nazarov's Theorem 2.3, $\mathbb{P}(A_1) \leq 5\epsilon_1^{\frac{1}{w}}$. Now for every $\theta_1 \in A_1^c$, let

$$\begin{aligned} A_2(\theta_1) &:= \{\theta : |p(\theta_1,\theta,0,\ldots,0)| < \epsilon_1\epsilon_2\} \\ &= \{\theta : |p(\theta_1,\theta,0,\ldots,0)/\epsilon_1| < \epsilon_2\}. \end{aligned}$$

Then again we have

$$\mathbb{P}(A_2(\theta_1)) \leq 5\epsilon_2^{\frac{1}{w}}.$$

Continuing this process, we can define inductively for each k, $A_k(\theta_1,\ldots,\theta_{k-1})$, where $\theta_j \in A_j(\theta_1,\ldots,\theta_{j-1})$ for each $j < k$. Also let

$$B_k = \{\theta : \theta_1 \in A_1^c, \theta_2 \in A_2^c(\theta_1), \ldots, \theta_k \in A_k(\theta_1,\ldots,\theta_{k-1})\}.$$

Then by Fubini's theorem, $\mathbb{P}(\cup_k B_k) \leq 5\sum_{k=1}^m \epsilon_k^{\frac{1}{w}}$. If we take $\epsilon_k \equiv (\frac{\delta}{5m})^w$, and $\epsilon := \epsilon_1^m$, then

$$\mathbb{P}(\cup_k B_k) \leq \delta = 5m\epsilon^{1/(wm)}.$$

Furthermore,

$$\{\theta \in \mathbb{T}^m : |p(\theta)| \leq \epsilon\} \subset \cup_k B_k.$$

□

Now we apply the above corollary to $p = J$. Here $w = 2\binom{n}{2}$. If we choose $m = 2n^2\log n$, then with high probability ($\geq 1 - n^{-2}/2$ by coupon collector's problem) we would have chosen every pair $i < j$ in the Kac rotation sequences. This implies that $J(0,\ldots,0) = \prod_{i<j}|\{k \in [m] : (i_k,j_k) = (i,j)\}| \geq 1$ with high probability:

The derivative with respect to each θ_k is simply the skew symmetric matrix $\partial_{i_k} \wedge \partial_{j_k}$, which when restricted to the upper triangular entries is just e_{i_k,j_k}. Any two rows in the matrix $dJ_{(0,\ldots,0)}$ do not share any entry of 1's, so the matrix $(\langle u_i,u_j\rangle)_{i,j=1}^{\binom{n}{2}}$ is a diagonal matrix with integer entries.

Thus we have $\mathbb{P}(|J| < (\delta/5m)^{n^2m}) \leq \delta$. So we can choose the generalized determinant threshold

$$c = (\delta/5m)^{n^2m} = (\frac{\delta}{10}n^{-2}(\log n)^{-1})^{2n^4\log n}. \tag{2.9}$$

Indeed, δ bounds the total mass of μ_b, i.e., $\|\mu_b\| \leq \delta$.

Now recall the spectral gap of the Kac walk on $SO(n)$ is less than $1/n$ ([7], [1]). By linearity of the Markov kernel, we can write $\mu P^t = \mu_a P^t + \mu_b P^t$ and $\nu = \nu_a + \nu_b$ where $\nu_a = \nu\|\mu_a\|$. Furthermore, for $\delta < 1/2$ and $n > 2$,

$$\begin{aligned}
&\|\mu_a P^t - \nu_a P^t\|_{\mathrm{TV}}^2 \leq \|\mu_a\|^2 \|\frac{\mu_a}{\|\mu_a\|}P^t - \frac{\nu_a}{\|\nu_a\|}P^t\|_{\mathrm{TV}}^2 \\
\leq& \|\frac{d\mu_a}{d\nu}\frac{1}{\|\nu_a\|} - 1\|_2^2(1-1/n)^{2t} \leq (\frac{1}{1-\delta}2^{7m}m^{5m}n^{8m}c^{-1})^2 e^{-2t/n} \\
\leq& e^{2(n^4\log n(\log(5\delta^{-1})+2\log n+\log\log n)+n^2\log n(7\log 2+18\log n+5\log\log n)-t/n)} \\
\leq& e^{2(2n^4\log n(\log(5\delta^{-1})+2\log n)-t/n)},
\end{aligned}$$

where we used the fact $\|\mu_a\| \leq 1$, the definition of $\delta = 1 - \|\nu_a\|$, the density upper bound (2.6), the threshold value (2.9) for c, and the choice of $m = n^2\log n$. Also we threw away the lower order terms in the exponent of the last expression by doubling the leading terms. The last expression is less than ϵ^2 if

$$t \geq T(n,\delta,\epsilon) := 2n^5\log n(\log(5\delta^{-1}) + 2\log n) - \log\epsilon.$$

On the other hand, $\|\mu_b - \nu_b\| \leq \|\mu_b\| + \|\nu_b\| = 2\delta$. So by monotonicity of total variation distance under Markov transformation, for $t \geq T(n,\delta,\epsilon)$,

$$\begin{aligned}
\|\mu P^t - \nu P^t\|_{\mathrm{TV}} &\leq \|(\mu_a - \nu_a)P^t\|_{\mathrm{TV}} + \|(\mu_b - \nu_b)P^t\|_{\mathrm{TV}} \\
&\leq \epsilon + \|(\mu_b - \nu_b)P^t\|_{\mathrm{TV}} \leq \epsilon + \|(\mu_b - \nu_b)\|_{\mathrm{TV}} \\
&\leq \epsilon + 2\delta.
\end{aligned}$$

Choosing $\delta = 1/16$ and $\epsilon = 1/8$ gives that $T(n,\delta,\epsilon) < 10n^5(\log n)^2$ for $n > 1$ and $T(n,\delta,\epsilon) < 5n^5(\log n)^2$ for $n > 80$. Therefore the total variation mixing time $\tau := \min\{t > 0 : \|\mu P^t - \nu\| < 1/4\}$ is bounded by $10n^5(\log n)^2$. In fact, it is clear from the form of $T(n,\delta,\epsilon)$ that the total variation distance to stationary decays at the rate of $\mathcal{O}(e^{-\frac{t}{2n^5\log n}})$ for t large, as opposed to $e^{-\frac{t}{2n^5(\log n)^2}}$.

3. Acknowledgement

Foremost, I want to thank Rafael Potrie ([13]) and Igor Rivin ([14]) on mathoverflow for pointing me to the relevant literature in algebraic complexity theory. I am grateful to my advisor Persi Diaconis for constant encouragement and numerous discussions regarding the Kac walk. I also thank Bob Hough and Aaron Smith for useful conversations.

References

1. E. A. Carlen, M. C. Carvalho, and M. Loss, *Determination of the spectral gap for Kac's master equation and related stochastic evolution*, Acta Math. **191** (2003), no. 1, 1–54. MR 2020418 (2004j:60186)
2. Eric A. Carlen, Maria C. Carvalho, Jonathan Le Roux, Michael Loss, and Cédric Villani, *Entropy and chaos in the Kac model*, Kinet. Relat. Models **3** (2010), no. 1, 85–122. MR 2580955 (2010m:76124)

3. Persi Diaconis and Laurent Saloff-Coste, *Bounds for Kac's master equation*, Comm. Math. Phys. **209** (2000), no. 3, 729–755. MR 1743614 (2002e:60107)
4. Natacha Fontes-Merz, *A multidimensional version of Turán's lemma*, J. Approx. Theory **140** (2006), no. 1, 27–30. MR 2226674 (2007a:41006)
5. Elise Janvresse, *Spectral gap for Kac's model of Boltzmann equation*, Ann. Probab. **29** (2001), no. 1, 288–304. MR 1825150 (2002c:82068)
6. Yunjiang Jiang, *Total variation bounds of Kac random walk*, Ann. Appl. Probab. (2011).
7. David K. Maslen, *The eigenvalues of Kac's master equation*, Math. Z. **243** (2003), no. 2, 291–331. MR 1961868 (2004b:60012)
8. J. Milnor, *On the Betti numbers of real varieties*, Proc. Amer. Math. Soc. **15** (1964), 275–280. MR 0161339 (28 #4547)
9. F. L. Nazarov, *On the theorems of Turán, Amrein and Berthier, and Zygmund*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) **201** (1992), no. Issled. po Linein. Oper. Teor. Funktsii. 20, 117–123, 191. MR 1172761 (93e:42002)
10. Roberto Imbuzeiro Oliveira, *On the convergence to equilibrium of Kac's random walk on matrices*, Ann. Appl. Probab. **19** (2009), no. 3, 1200–1231. MR 2537204 (2010j:60190)
11. Natesh S Pillai and Aaron Smith, *Kac's walk on n-sphere mixes in $n \log n$ steps*, arXiv preprint arXiv:1507.08554 (2015).
12. ______, *On the mixing time of Kac's walk and other high-dimensional Gibbs samplers with constraints*, arXiv preprint arXiv:1605.08122 (2016).
13. R. Potrie, *Communication via mathoverflow.*, `http://mathoverflow.net/questions/45199/`.
14. I. Rivin, *Communication via mathoverflow.*, `http://mathoverflow.net/questions/84471/`.
15. Sergiy Sidenko, *Kac's random walk and coupon collector's process on posets*, ProQuest LLC, Ann Arbor, MI, 2008, Thesis (Ph.D.)–Massachusetts Institute of Technology. MR 2717533
16. Aaron Smith et al., *A Gibbs sampler on the n-simplex*, The Annals of Applied Probability **24** (2014), no. 1, 114–130.
17. René Thom, *Sur l'homologie des variétés algébriques réelles*, Differential and Combinatorial Topology (A Symposium in Honor of Marston Morse), Princeton Univ. Press, Princeton, N.J., 1965, pp. 255–265. MR 0200942 (34 #828)

Stanford University, Department of Mathematics, 450 Serra Mall, Stanford, CA 94305

E-mail address: `jyj@math.stanford.edu`

Current address: Google Inc., 1600 Amphitheatre Parkway, Mountain View, CA 94043

E-mail address: `yunjiangster@gmail.com`