JOURNAL OF MATHEMATICAL PHYSICS **50**, 123521 (2009)

Volume of the set of unistochastic matrices of order 3 and the mean Jarlskog invariant

Charles Dunkl[1,a)] and Karol Życzkowski[2,b)]
[1]*Department of Mathematics, University of Virginia, Charlottesville, Virginia 22904-4137, USA*
[2]*Institute of Physics, Jagiellonian University, 30-059 Cracow, Poland and Center for Theoretical Physics, Polish Academy of Sciences, 00-668 Warsaw, Poland*

(Received 4 September 2009; accepted 13 November 2009; published online 22 December 2009)

A bistochastic matrix B of size N is called *unistochastic* if there exists a unitary U such that $B_{ij}=|U_{ij}|^2$ for $i,j=1,\dots,N$. The set $\mathcal{U}_3$ of all unistochastic matrices of order $N=3$ forms a proper subset of the Birkhoff polytope, which contains all bistochastic (doubly stochastic) matrices. We compute the volume of the set $\mathcal{U}_3$ with respect to the flat (Lebesgue) measure and analytically evaluate the mean entropy of an unistochastic matrix of this order. We also analyze the Jarlskog invariant J, defined for any unitary matrix of order three, and derive its probability distribution for the ensemble of matrices distributed with respect to the Haar measure on $U(3)$ and for the ensemble which generates the flat measure on the set of unistochastic matrices. For both measures the probability of finding $|J|$ smaller than the value observed for the Cabbibo–Kobayashi–Maskawa matrix, which describes the violation of the CP parity, is shown to be small. Similar statistical reasoning may also be applied to the Maki–Nakagawa–Sakata matrix, which plays role in describing the neutrino oscillations. Some conjectures are made concerning analogous probability measures in the space of unitary matrices in higher dimensions. © *2009 American Institute of Physics.* [doi:10.1063/1.3272543]

I. INTRODUCTION

Bistochastic matrices appear in variety of problems in different branches of science. A bistochastic matrix (also called doubly stochastic) contains real non-negative entries, the sum of which in each column and in each row is equal to unity. Thus each column and each row of such a matrix can be interpreted as a probability vector. The structure of the set $\mathcal{B}_N$ of all bistochastic matrices of order N is well understood.[1] It is formed by the convex polytope of all permutation matrices of size N and it is often called the *Birkhoff polytope*.[2] Analytical expressions for its volume for small dimensionality[3,4] and for the leading terms of the volume in the asymptotic limit[5,6] are available in the literature.

In various physical problems it is assumed that the probabilities, which form the entries of a bistochastic matrix, arise as squared modulus of an element of a (*a priori* unknown) unitary matrix. A bistochastic matrix B which can be generated from a unitary matrix U by the relation $B_{ij}=|U_{ij}|^2$ is called *unistochastic* (or orthostochastic).

For instance, such matrices are used in high energy physics to characterize interactions between elementary particles, which can be divided into N generations. Since the Hamiltonians describing two kinds of physical interactions (usually called "strong" and "weak") do not commute, these two Hermitian operators determine a single unitary matrix of order N, which relates both eigenbases. This is the famous unitary matrix of Cabbibo–Kobayashi–Maskawa (CKM).[7]

[a)]Electronic mail: cfd5z@virginia.edu.
[b)]Electronic mail: karol@tatry.if.uj.edu.pl.

Note that the squared moduli of the CKM matrix V_{CKM} form the corresponding unistochastic matrix B, the entries of which represent probabilities which are accessible experimentally.[8–12]

According to the standard model of elementary particles there exist three generations of quarks, thus the case of a direct physical importance are unistochastic matrices of order $N=3$. On the other hand, the case $N=4$ could also become relevant in case a fourth generation of quarks should be discovered.[13–15] A similar problem in the neutrino physics is characterized by the Maki–Nakagawa–Sakata matrix (MNS matrix),[16] some parameters of which are still quite uncertain. A relation between the MNS and CKM matrices is studied in Ref. 17.

In practice, given a bistochastic matrix $B \in \mathcal{B}_N$ it is important to know, whether it belongs to the set $\mathcal{U}_N$ of unistochastic matrices. If this is the case one would like to describe the set of all unitary matrices such that $B_{ij}=|U_{ij}|^2$ for $i,j=1,\ldots,N$. These questions are of a particular interest for research in foundations of quantum mechanics and investigation of properties of transition probabilities,[18–20] scattering theory,[21] quantum counterparts of Markov processes and dynamics on graphs,[22,23] and the theory of quantum information processing.[24]

Any bistochastic matrix of order of 2 is unistochastic, so both sets coincide, $\mathcal{U}_2=\mathcal{B}_2$. The situation differs already for $N=3$. To show this fact Schur considered the symmetric combination of cycle-three permutation matrices, P and $P^{-1}=P^2$. This matrix

$$B_S = \frac{1}{2}(P+P^2) = \frac{1}{2}\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \tag{1}$$

is clearly bistochastic but it is easy to see that there is no corresponding unitary matrix. Hence, $\mathcal{U}_3 \nsubseteq \mathcal{B}_3$ and a similar relation holds for an arbitrary $N \geq 3$. Vaguely speaking, the moduli of a unitary matrix need to fulfill certain constraints, more stringent than the obvious fact that the sum of squared moduli in each row (each column) is equal to unity. As recently analyzed by Diţă,[25] this simple observation has important consequences for reconstruction of unitary matrices from experimental data.

A general question, if a given bistochastic matrix B is unistochastic, remains open, and only partial results are available.[26,27] Necessary and sufficient conditions for unistochasticity are known for $N=3$,[9,28,29] while the constraints for unistochasticity recently obtained by Diţă[25] for the general case of $N \geq 4$ are formulated implicitly and do not provide a constructive solution of the problem.

Geometrical properties of the set of $\mathcal{U}_3$ of unistochastic matrices of order of 3 were studied in Refs. 27 and 28. This set contains a four-dimensional (4D) unistochastic ball centered at the flat (van der Waerden) matrix W, for which all entries are equal to 1/3. The set $\mathcal{U}_3$ is not convex but it is star shaped. This means that if $B \in \mathcal{U}_3$ then the entire interval BW belongs to the set. The volume of $\mathcal{U}_3$ with respect to the Euclidean (Lebesgue) measure was numerically estimated by a Monte Carlo-type procedure.[27]

The main aim of this work is to derive an analytical formula for the volume of the set $\mathcal{U}_3$ of unistochastic matrices of order of 3 with respect to the flat, Euclidean measure. We obtain also a compact expression allowing one to average any function of elements of B over the set $\mathcal{U}_3$. In particular, we derive an explicit result for the average generalized entropies S_q which in the special case $q=1$ gives the mean Shannon entropy of the columns of unistochastic matrices. We compute also the average Jarlskog invariant J proportional to the area A of the unitarity triangle, which characterizes any unistochastic matrix.[8,9] Furthermore, we derive higher moments $\langle J^k \rangle$ and analyze the probability distribution $P(J)$, to get more insight into properties of the CKM matrix, which describes violation of the CP symmetry. Such an approach was recently suggested by Gibbons *et al.*,[30] who used other probability measures for this purpose.

All averages are computed with respect to the natural Euclidean measure in the set of unistochastic matrices and are compared with the averages with respect to the measure induced by the Haar measure on $U(3)$. This measure leading to the *unistochastic ensemble*,[26] called flag-manifold measure in Ref. 30, is not uniform in the set $\mathcal{U}_3$. Although approach presented, based on properties of intertwining operators associated with the group S_3,[31,32] is directly applicable to the case

$N=3$, we conjecture also some properties of the measures in the sets of unistochastic matrices in higher dimensions.

The paper is organized as follows. In Secs. II and III we present the necessary definitions and review key properties of the Birkhoff polytope $\mathcal{B}_3$ and its subset $\mathcal{U}_3$ containing unistochastic matrices. In Sec. IV we define a family of measures in the set $\mathcal{U}_3$ and compute its volume with respect to them. Similar study of average entropies is presented is Sec. IV while the average Jarlskog invariant and its distribution are investigated in Sec. V. The paper is concluded with Sec. VI, while some conjectures concerning the measures in the set of unitary matrices for $N\geq 4$ are relegated to Appendix.

II. THE BIRKHOFF POLYTOPE $\mathcal{B}_3$

A real square matrix B of order N is called *bistochastic* (or doubly stochastic) if it satisfies the following conditions:

$$\text{(i)}\ \ B_{ij}\geq 0,$$

$$\text{(ii)}\ \ \sum_i B_{ij}=1,$$

$$\text{(iii)}\ \ \sum_j B_{ij}=1. \tag{2}$$

Such a matrix B is often used to describe discrete dynamics, $p'=Bp$, in the space of probability vectors. Condition (i) implies that all elements of the transformed vector p' are nonnegative. Due to condition (ii) its 1-norm $\Sigma_i p_i$ is preserved. A matrix satisfying two first conditions is called *stochastic* and it sends the simplex of N-point probability vectors into itself. Condition (iii) implies that additionally the transposed matrix B^T is stochastic, which explains the name.

The uniform probability vector p_* with all components equal, $p_i=1/N$, stays clearly invariant with respect to any bistochastic matrix, $Bp_*=p_*$. Thus a bistochastic matrix describes a (weak) contraction of the probability simplex toward the uniform distribution p_*.

Let $\mathcal{B}_N$ denote the set of all bistochastic matrices of order N, called *Birkhoff polytope*. This convex polytope is well known in linear programming. Since it arises in the problem of assigning N workers to N tasks, given their efficiency ratings for each task, it is sometimes called the *assignment polytope*.

The Birkhoff polytope is equivalent to the convex hull of all $N!$ permutation matrices of size N. Hence a permutation matrix P forms an extremal point of $\mathcal{B}_N$. All corners of $\mathcal{B}_N$ are equivalent in the sense that a given corner can be obtained from another one by an orthogonal transformation. A bistochastic matrix belongs to the boundary of the Birkhoff polytope if and only if at least one of its entries is equal to zero.

There exists a unique bistochastic matrix W, with all entries equal, $W_{ij}=1/N$. It is also called a matrix of van der Waerden, since it saturates the van der Waerden inequality[1] concerning the permanent of bistochastic matrices, $\mathrm{per}(B)\geq N!N^{-N}$. It is easy to see that W is located symmetrically at the center of the Birkhoff polytope.

A bistochastic matrix B can be determined by its minor of size $(N-1)$. Hence the dimensionality of the Birkhoff polytope $\mathcal{B}_N$ equals $(N-1)^2$. For instance, the dimension of the set $\mathcal{B}_2$ is equal to 1, and this set forms indeed an interval between the identity matrix $\mathbb{1}_2$ and the 2-element permutation matrix. In other words, any bistochastic matrix of order of 2 can be written as

$$B_2(a)=\begin{bmatrix} a & 1-a \\ 1-a & a \end{bmatrix}, \quad \text{where}\ \ a\in[0,1]. \tag{3}$$

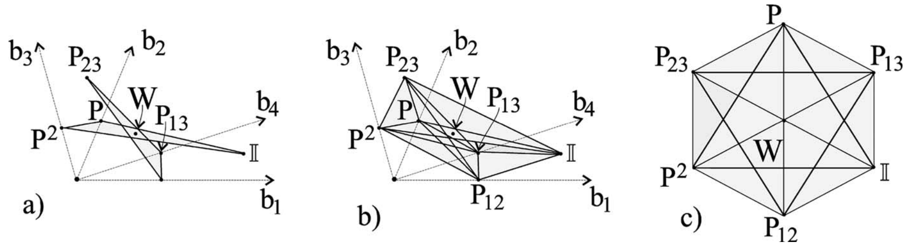

FIG. 1. Birkhoff polytope for $N=3$ plotted in parametrization (7), (a) two orthogonal equilateral triangles centered at the flat van der Waerden matrix W, (b) all 15 edges determining the polytope, and (c) the same polytope as seen "from above."

The length of this interval, equivalent to volume of $\mathcal{B}_2$, is equal to unity, if we consider it as a subset of $\mathbb{R}^1$. However, we are going to consider this set as an element of $\mathbb{R}^{N^2}$ then the distance between the points (1,0,1,0) and (0,1,0,1) is equal to 2, so in these units one has $\mathrm{vol}_1(\mathcal{B}_2)=2$.

In this paper we are going to work with the case $N=3$, so the Birkhoff polytope is defined as a convex hull of $3!=6$ permutation matrices,

$$P=\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix},\quad P^2=P^{-1}=\begin{bmatrix}0&0&1\\1&0&0\\0&1&0\end{bmatrix},\quad \mathbb{1}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}, \tag{4}$$

$$P_{12}=\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix},\quad P_{13}=\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix},\quad P_{23}=\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}. \tag{5}$$

We divided the permutation matrices into two triples, which belong to two totally orthogonal 2-planes. A uniform mixture in any triple produces the flat matrix W,

$$P+P^2+\mathbb{1}=W=P_{12}+P_{13}+P_{23}. \tag{6}$$

Working with the standard Hilbert–Schmidt distance, defined by $D^2(A,B):=\mathrm{Tr}(A-B)(A-B)^*$, we see that both triples form equilateral triangles. To produce a sketch of them in four dimensions we will use the following parametrization:

$$B(\vec{b})=B(b_1,b_2,b_3,b_4):=\begin{bmatrix}b_1&b_2&1-b_1-b_2\\b_3&b_4&1-b_3-b_4\\1-b_1-b_3&1-b_2-b_4&\sum_{i=1}^{4}b_i-1\end{bmatrix}. \tag{7}$$

Both triangles shown in Fig. 1(a), cross at their center W. Six permutation matrices form 15 edges, out of which all belong to the boundary of $\mathcal{B}_3$ and all are extremal. There are six long edges, of length $\sqrt{6}$, which form two equilateral triangles, and nine short edges of length of 2. If one plots all of them, as in Fig. 1(b), the sketch of $\mathcal{B}_3$ becomes complete, but not very illuminating. Another natural possibility is to look at the polytope "from above," the direction distinguished by the vector $\vec{b}=(1/3,1/3,1/3,1/3)$. The Birkhoff polytope then appears symmetrically as a regular hexagon with two inscribed equilateral triangles forming the Star of David—see Fig. 1(c). Note that all the diagonals of the hexagon belong to the boundary of $\mathcal{B}_3$ and that the distance PP_{12} of the diagonal is shorter than the side $P\mathbb{1}$ of the equilateral triangle. Any two-dimensional face of the polytope is formed by an isosceles triangle with two short edges and one long.

The polytope $\mathcal{B}_3$ is defined as the convex hull of six corners, so one could think, it may be decomposed into two 4D simplices, each determined by five points. This would be possible if we could select four corners, which span the base of a simplex and then allow two other corners to play the role of an apex for two simplices, with the same base. However, this would require that the edge connecting both apexes is not extremal or it includes one of the corner from the base of the simplex. Neither of these holds for the Birkhoff polytope, so its decomposition into two simplices is not possible.

To find a decomposition of $\mathcal{B}_3$ into three simplices take three corners of one equilateral triangle, e.g., $\Delta(P,P^2,\mathbb{1})$. Out of the orthogonal triangle select a side, say the one formed by the corners P_{12} and P_{13}. These five corners define a 4D simplex. The same construction performed for two other sides of the $\Delta(P_{12},P_{13},P_{23})$ produces two other simplices. It is easy to show that any point of $\mathcal{B}_3$ belongs to one of these simplices and that the 4D volume of any of their intersections is equal to zero. Such a triangulation of the Birkhoff polytope allows to find that its volume in $\mathbb{R}^9$ according to the Lebesgue measure is equal to 9/8. A detailed investigation of the geometry of the Birkhoff polytope is provided in Ref. 33.

III. THE SET $\mathcal{U}_3$ OF UNISTOCHASTIC MATRICES

A certain class of bistochastic matrices can be generated from unitary matrices. Let U denote a unitary matrix. Unitarity condition, $UU^*=\mathbb{1}$, implies that the matrix B defined by

$$B=f(U), \quad \text{so that } B_{ij}=|U_{ij}|^2, \tag{8}$$

is bistochastic. Any bistochastic matrix $B\in\mathcal{B}_N$ for which there exists unitary $U\in U(N)$ such that $B=f(U)$ is called *unistochastic*. The set of all unistochastic matrices of size N will be denoted as $\mathcal{U}_N$.

Note that the multiplication of U by any diagonal unitary matrices D_1 and D_2 changes the phases of entries of U, but does not modify the corresponding bistochastic matrix. Hence we define an equivalence relation,

$$U \approx U' = D_1 U D_2, \tag{9}$$

and observe that $B=f(U)=f(U')$.

If the unitary matrix U, appearing in (8), is orthogonal, the corresponding bistochastic matrix B is called *orthostochastic* (in some papers this name is used for unistochastic matrices as well). This is the case for any bistochastic matrix of size 2, since writing an orthogonal matrix $O=\begin{pmatrix}\cos\vartheta & \sin\vartheta\\ -\sin\vartheta & \cos\vartheta\end{pmatrix}$ and taking $\vartheta=\arccos a$, we see that $B=f(O)$ for any bistochastic matrix of order of 2 represented in the form (3). Therefore, any unistochastic matrix of order of 2 is also orthostochastic. This is no longer the case for $N\geq 3$, as it is explicitly shown later in this section.

Interestingly this simple mathematical observation has far reaching consequences for physics. In the theory of elementary particles one defines a discrete space-time symmetry called CP, which stands for *charge* conjugation and *parity*. Such a symmetry requires that a physical process in which all particles are replaced by their antiparticles is equivalent to the mirror image of the original process.

If such a symmetry were obeyed the CKM matrix V_{CKM} would be orthogonal, (and thus invariant with respect to the complex conjugation) or it would be equivalent to an orthogonal matrix with respect to (9). For $N=2$ this is the case for any unitary matrix from $U(2)$. As the CP symmetry was discovered in 1964 to be violated in experiments on decay of neutral mesons K, one could predict that the number N of generations of quarks in the theory, equal to the size of the CKM matrix, has to be greater than 2.

In her first paper on the CKM matrix Jarlskog[8] observed that for any unitary matrix U of size 3, the number

$$J := \mathrm{Im}(U_{11}U_{22}U_{12}^*U_{21}^*) \tag{10}$$

is invariant with respect to multiplication of the matrix U by diagonal unitary matrices and permutations. This quantity, now called the *Jarlskog invariant*, computed for the CKM matrix V_{CKM} can be considered as a measure of the violation of the CP symmetry.

Consider now an arbitrary bistochastic matrix B of size N. To check if this matrix is unistochastic, we need to know whether there exists a unitary matrix U such that $f(U)=B$ according to Eq. (8). The moduli of the unitary matrix are determined by the square roots of the entries of the bistochastic matrix, $|U_{ij}|=\sqrt{B_{ij}}$, and one needs to find a set of phases ϕ_{ij} which guarantees unitarity, where $U_{ij}=|U_{ij}|\exp(\mathrm{i}\phi_{ij})$.

Choose the two first columns of U, which we denote as $|u_1\rangle$ and $|u_2\rangle$. Their orthogonality relation, $\langle u_1|u_2\rangle=0$, implies that

$$\sum_{j=1}^{N} U_{j1}U_{j2}^{*}=0.$$

Introducing the notation $L_j=|U_{j1}|\cdot|U_{j2}|$ and $\theta_j=\phi_{j1}-\phi_{j2}$, we may rewrite this relation as

$$\sum_{j=1}^{N} L_j \exp(\mathrm{i}\theta_j)=0. \tag{11}$$

This form has a nice geometric interpretation: given a set of N line segments of lengths $L_1,\ldots,L_N$ we need to find the phases θ_j in such a way that the entire chain is closed. Obviously it cannot be done unless the longest link is shorter or is equal to the sum of all other links. We are free to change the order of summation in (11) and hence to relabel the links in such a way that they are ordered nonincreasingly, $L_1\geq L_2\geq\cdots\geq L_N$. Then the chain-link condition reads

$$L_1 \leq L_2+\cdots+L_N. \tag{12}$$

If it is satisfied the chain can be closed and forms a unitarity polygon.

This relation was imposed by the assumed orthogonality of the first two columns of U, but analogous conditions should be fulfilled by all links corresponding to any pair of columns of U. Similar conditions are due to the orthogonality between any two rows of U. This implies the total number of $N(N-1)$ constraints of the form (12), some of which can be dependent.[26,34] However, there is an example of a bistochastic matrix B of order of 4, which satisfies all chain-link condition for all pair of rows, but not for all pair of columns,[26,35] so in practice one has to check rows and columns separately. Furthermore, for $N\geq 4$ these conditions for B are only necessary but not sufficient to imply unistochasticity.

It is comforting to realize that the situation gets simpler for $N=3$. In this case the chain-link relation for the first two columns reduces to the triangle inequality,

$$|L_2-L_3| \leq L_1 \leq L_2+L_3, \tag{13}$$

and the first constraint is required if we relax the assumption that the links are ordered decreasingly. Although, in general, one should check similar relations stemming to other pairs of columns and rows of U, in the case $N=3$ the last column by construction has the right moduli and does not impose any further restrictions.[29] Thus in this case the relation allowing a chain to close is sufficient for unistochasticity,[9,28] and explicit formulas for the phases ϕ_{ij} are provided below. If there is a bistochastic matrix B such that in all relations (13) equality takes place, the phases θ_j are equal to zero or to π. Thus the corresponding matrix U is orthogonal, which means that B is orthostochastic. It is easy to show that a matrix B belongs to the boundary of the set $\mathcal{U}_3$ if and only if B is orthostochastic. The set $\mathcal{U}_3$ forms a 4D subset of $\mathcal{B}_3$ of a positive measure, while the set of orthostochastic matrices, at the boundary of $\mathcal{U}_3$, is three dimensional.[27,28]

For any given B of order of 3 it is straightforward to check whether link conditions (13) are fulfilled, so that B is unistochastic. For instance, nine short edges, (of length of 2) of the Birkhoff polytope belong to $\mathcal{U}_3$, while the long edges (of length of $\sqrt{6}$) do not belong to this set.

Let us take three such edges spanned by P, P^2, and $\mathbb{1}$, which form the equilateral triangle. At this plane, the set of orthostochastic matrices, for which $L_1=L_2+L_3$, forms a *deltoid*—see Fig. 3(a). This figure also called 3-*hypocycloid* may be obtained by sliding a circle of radius 1/3 inside

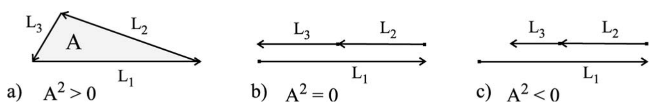

FIG. 2. Chain-link condition for unistochasticity: (a) a unistochastic matrix with a positive area of the unitarity triangle, $A^2>0$, (b) limiting case: an orthostochastic matrix with $A^2=0$, (c) a bistochastic matrix B not included into $\mathcal{U}_3$ for which $A^2<0$.

the unit circle. Thus the set of unistochastic matrices corresponds to the interior of the deltoid and is not convex. This set contains the maximal unistochastic ball of radius $r=\sqrt{2}/3$ centered at W, which touches the boundary at the deltoid.

Incidentally, the very same figure is related in a different way with the set of unistochastic matrices of order of 3. The spectra of these matrices are real or belong to the deltoid inscribed into the unit disk and stemming from the real eigenvalue equal to unity.[26]

Consider a unistochastic matrix B parametrized by (7). The length of the links reads

$$L_1=\sqrt{b_1b_2},\quad L_2=\sqrt{b_3b_4},\quad L_3=\sqrt{(1-b_1-b_2)(1-b_3-b_4)},\tag{14}$$

and the triangle inequality (13) provides the direct condition for unistochasticity. Let us write down the area A of this *unitarity triangle* with sides L_1, L_2, and L_3, and semiperimeter $p=(L_1+L_2+L_3)/2$. Making use of Heron's formula,

$$A=\sqrt{p(p-L_1)(p-L_2)(p-L_3)},\tag{15}$$

and substituting (14), we arrive with a compact expression for the squared area A^2.

It will be convenient to work with this quantity multiplied by 16,

$$Q(b) := 4b_1b_2b_3b_4-(b_1+b_2+b_3+b_4-1-b_1b_4-b_2b_3)^2=16A^2.\tag{16}$$

Here $b=\{b_1,b_2,b_3,b_4\}$ represents a vector in $\mathbb{R}^4$ which determines a bistochastic matrix in parametrization (7). In fact, we can form six unitarity triangles in this way, depending on what pair of columns or rows we wish to choose. Although their shapes differ due to unitarity their area A is the same[9] so the quantity A does not change under permutation of the unitary matrix U.

It is easy to see that all chain-link conditions are equivalent to the single condition for unistochasticity,

$$A^2(B)\geq 0.\tag{17}$$

In other words $B(b)\in\mathcal{U}_3$ if and only if $b\in\Omega$, where

$$\Omega := \{b\in\mathbb{R}^4 : b_1\geq 0, b_2\geq 0, b_1+b_2\leq 1, Q(b)\geq 0\}.\tag{18}$$

Also Ω is the closure of the connected component of $\{b:Q(b)>0\}$ which contains $(\frac{1}{3},\frac{1}{3},\frac{1}{3},\frac{1}{3})$ (see Ref. 31, Sec. 2).

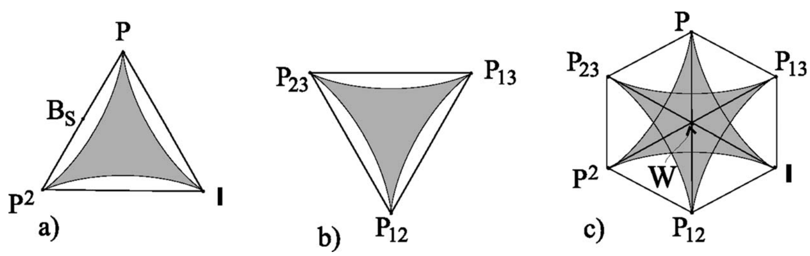

FIG. 3. Nonconvex set $\mathcal{U}_3$ of unistochastic matrices forms a proper subset $\mathcal{B}_3$. (a) Deltoid obtained by the cross section of $\mathcal{U}_3$ along the plane spanned by the equilateral triangle $\Delta(P,P^2,\mathbb{1})$, (b) a similar cross section along totally orthogonal plane, and (c) a view "from above" as in Fig. 2(c).

We relate these expressions to one of the standard parametrizations of a unitary matrix (see Ref. 25, p. 11). For any $U' \in U(3)$ there are diagonal matrices $D_1, D_2 \in U(3)$, such that $D_1 U' D_2$ =

$$U = \begin{bmatrix} c_{12} & s_{12}c_{13} & s_{12}s_{13} \\ s_{12}c_{23} & -c_{12}c_{13}c_{23} - e^{\mathrm{i}\delta}s_{13}s_{23} & e^{\mathrm{i}\delta}c_{13}s_{23} - c_{12}c_{23}s_{13} \\ s_{12}s_{23} & e^{\mathrm{i}\delta}c_{23}s_{13} - c_{12}c_{13}s_{23} & -c_{12}s_{13}s_{23} - e^{\mathrm{i}\delta}c_{13}c_{23} \end{bmatrix}. \tag{19}$$

The parameters are the angles $\theta_{12}, \theta_{13}, \theta_{23}, \delta$ and $c_{jk} := \cos\theta_{jk}, s_{jk} := \sin\theta_{jk}, 0 \le \theta_{jk} \le \pi/2$ for $1 \le j < k \le 3$. The function (8) determines thus a unistochastic matrix $B(b) = f(U)$, and its entries read

$$b_1 = c_{12}^2, \quad b_2 = s_{12}^2 c_{13}^2, \quad b_3 = s_{12}^2 c_{23}^2,$$

$$b_4 = c_{12}^2 c_{13}^2 c_{23}^2 + s_{13}^2 s_{23}^2 + 2c_{12}c_{13}c_{23}s_{13}s_{23}\cos\delta. \tag{20}$$

We explain how these parameters are used to determine the phases ϕ_{ij} for $2 \le i, j \le 3$. We consider only the nondegenerate case in which all entries of U are nonzero. This implies $c_{jk} > 0$ and $s_{jk} > 0$ for $1 \le j < k \le 3$. Equations (20) determine δ up to the sign (in $-\pi < \delta < \pi$). The fact that U and $\bar{U} := [\overline{U_{ij}}]_{i,j=1}^3$ produce the same values for $b_1, \ldots, b_4$ causes this ambiguity. We will adopt the normalization Im $U_{22} > 0$. This forces $\sin\delta < 0$, Im $U_{32} < 0$, Im $U_{23} < 0$ and Im $U_{33} > 0$. The cosines of the phases are computed using the entries in (19). The phases are related to the (interior) angles of the unitarity triangles derived from columns 1 and 2 and from columns 1 and 3. For the former case, using the lengths from Eq. (14) and denoting the angles $\theta_1, \theta_2, \theta_3$ by the label on the opposite side, we find

$$\cos\theta_3 = \frac{L_1^2 + L_2^2 - L_3^2}{2L_1L_2} = \frac{b_1b_2 + b_3b_4 - (1 - b_1 - b_3)(1 - b_2 - b_4)}{2\sqrt{b_1b_2b_3b_4}} = -\cos\phi_{22},$$

$$\cos\theta_2 = \frac{L_1^2 + L_3^2 - L_2^2}{2L_1L_3} = \frac{b_1b_2 + (1 - b_1 - b_3)(1 - b_2 - b_4) - b_3b_4}{2\sqrt{b_1b_2(1 - b_1 - b_3)(1 - b_2 - b_4)}} = -\cos\phi_{32}.$$

From the conditions $\sin\phi_{22} > 0$ and $\sin\phi_{32} < 0$, we obtain $\phi_{22} = \pi - \theta_3$ and $\phi_{32} = \theta_2 - \pi$ (note the interior angles satisfy $0 < \theta_1, \theta_2, \theta_3 < \pi$ thus $0 < \phi_{22} < \pi$ and $-\pi < \phi_{32} < 0$). By interchanging columns 2 and 3 we find the remaining nonzero phases [we use matrix entry notation from Eq. (7) for more concise statements],

$$\cos\phi_{23} = -\frac{b_1B_{13} + b_3B_{23} - B_{31}B_{33}}{2\sqrt{b_1b_3B_{13}B_{23}}}, \quad \sin\phi_{23} < 0,$$

$$\cos\phi_{33} = -\frac{b_1B_{13} + B_{31}B_{33} - b_3B_{23}}{2\sqrt{b_1B_{13}B_{31}B_{33}}}, \quad \sin\phi_{33} > 0.$$

For the example where each $b_i = \frac{1}{3}$ the unitarity triangle is equilateral, each $\theta_i = \pi/3$, and the above equations give $\phi_{22} = 2\pi/3 = \phi_{33}$ and $\phi_{32} = -2\pi/3 = \phi_{23}$.

We return to the consideration of the Jarlskog invariant. A straightforward computation yields

$$2\,\mathrm{Re}(U_{11}U_{22}U_{12}^*U_{21}^*) = 1 - b_1 - b_2 - b_3 - b_4 + b_1b_4 + b_2b_3, \tag{21}$$

and thus the square of the Jarlskog invariant (10) reads

$$[\mathrm{Im}(U_{11}U_{22}U_{12}^*U_{21}^*)]^2 = b_1b_2b_3b_4 - [\mathrm{Re}(U_{11}U_{22}U_{12}^*U_{21}^*)]^2. \tag{22}$$

Substituting expression (21) into above equation and comparing the outcome with (16), we see that

$$J^2 = \tfrac{1}{4}Q(b) = 4A^2. \tag{23}$$

Thus the squared Jarlskog invariant, proportional to the squared area of the unitarity triangle, may also be defined as in (16) for an arbitrary bistochastic matrix $B \in \mathcal{B}_3$. For simplicity we shall write according to the context $J=J(U)$ or $J=J(B)=J(B(f(U)))$, as it should not lead to misunderstanding.

The squared Jarlskog invariant J^2 is equal to zero if and only if B is orthostochastic, so there exists an orthogonal matrix O, such that $B_{ij}=O_{ij}^2$.

Following Haagerup[36] we shall call two unitary matrices U_1 and U_2 *equivalent*, written $U_1 \sim U_2$, if there exist two diagonal unitary matrices D_1 and D_2 and two permutation matrices P_1 and P_2 such that

$$U_2 = D_1 P_1 \cdot U_1 \cdot P_2 D_2. \tag{24}$$

Observe that due to permutation matrices this relation is more general than the relation (9).

Since multiplication by phases or permutations do not vary the area of the unitarity triangle we see that the squared Jarlskog invariants of two equivalent unitaries are equal, if $U_1 \approx U_2$ then $J^2(U_1)=J^2(U_2)$. Going back to the set of unitary matrices $U(3)$ we see that $J(U)=0$ if U is orthogonal, or more generally, if U is *equivalent* to an orthogonal matrix, $U=D_1P_1\cdot O\cdot P_2D_2$.

Thus $J^2(U)$ measures to what extend the matrix U can be transformed into an orthogonal matrix by means of enphasing and permutations. It is easy to see that J is maximal if the unitarity triangle is equilateral, $L_1=L_2=L_3=1/3$, so that $J^2_{\max}=1/108$—see Appendix A. This is the case for the flat matrix W of van der Waerden, which corresponds to the unitary Fourier matrix,

$$F_3 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 & 1 & 1 \\ 1 & \exp(\mathrm{i}\cdot 2\pi/3) & \exp(\mathrm{i}\cdot 4\pi/3) \\ 1 & \exp(\mathrm{i}\cdot 4\pi/3) & \exp(\mathrm{i}\cdot 2\pi/3) \end{bmatrix}, \tag{25}$$

which is an example of a complex Hadamard matrix of order of 3. Such a unitary matrix H of size N is distinguished by an extra condition that all its complex entries have the same modulus, $|H_{ij}|^2=1/N$ for $i,j=1,\ldots,N$.[37]

Any complex Hadamard matrix H of order of 3 is known to be equivalent to the Fourier matrix F_3,[36] which implies that $J^2(H)=J^2_{\max}$. From this perspective the set of Hadamard matrices is maximally distant from the set of orthogonal matrices. While complex Hadamard matrices correspond to the flat bistochastic matrix W, located at the center of the set of unistochastic matrices, the boundary of which is formed by the set of orthostochastic matrices.

On the other hand for any bistochastic matrix which is not unistochastic, the quantity $J^2 = Q/4$ defined by (16) is negative. Since $Q=16A^2$ one might say that in this case the area of the unitarity triangle is imaginary, since the three segments L_i cannot be closed to form a triangle. Among all bistochastic matrices of size of 3 the quantity Q is the smallest for the matrix B_S of Schur (1) for which $Q=-\frac{1}{16}$, see Appendix A. Indeed, looking at Fig. 3(a) we see that B_S is such a point of the Birkhoff polytope $\mathcal{B}_3$, for which the distance to the set $\mathcal{U}_3$ of unistochastic matrices, represented by the gray deltoid, is maximal.

IV. THE VOLUME OF THE SET OF UNISTOCHASTIC MATRICES OF ORDER OF 3

Since unistochastic matrices of order of 3 are used in various branches of theoretical physics, several geometric properties of the set $\mathcal{U}_3$ of these matrices were studied in Ref. 27. In particular, in that work the volume of this set was estimated numerically. In this section we shall improve these findings by deriving an analytical formula for this volume. To this end we need to introduce probability measures into the set of unistochastic matrices. A first natural choice will be as follows.

(a) The flat (Lebesgue) measure $\mu_{3/2}$ used before in Ref. 27. The above notation is due to the

fact that this measure belongs to a one-parameter class of measures μ_k, defined below. In general, any probability measure in the set $U(N)$ of unitary matrices induces by function (8) a measure into the set $\mathcal{U}_N$. Thus we will distinguish the following case.

(b) The measure μ_1 induced by the Haar measure on $U(3)$. This measure leads to the *unistochastic ensemble* or random unistochastic matrices.[26] Since it is related to the unitarily invariant measure on the flag manifold $U(3)/[U(1)]^3$ it was called *flag-manifold measure* in Ref. 30.

Suppose now that dm denotes the normalized Haar measure on $U(3)$, db is Lebesgue measure on $\mathbb{R}^4$, and g is a continuous function on $[0,1]^4 \subset \mathbb{R}^4$. Due to Ref. 31, Theorem 2.1, one may relate the integrals over the space of unitary matrices $U(3)$ and over the set $\Omega \in \mathbb{R}^4$, see Eq. (18), which determines the set of unistochastic matrices,

$$\int_{U(3)} g(|U_{11}|^2,|U_{12}|^2,|U_{21}|^2,|U_{22}|^2)dm = \frac{2}{\pi}\int_{\Omega} g(b_1,b_2,b_3,b_4)Q(b)^{-1/2}db. \tag{26}$$

The function $g(b_1,b_2,b_3,b_4)=g(b)$ determines the function $g(B)$ defined on the entire set $\mathcal{U}_3$ of unistochastic matrices.

We introduce new coordinates (b_1,s,t,r) with

$$b_2 = s(1-b_1), \quad b_3 = t(1-b_1),$$

$$b_4 = (1-s)(1-t) + b_1 st + 2r\sqrt{b_1 st(1-s)(1-t)}. \tag{27}$$

Note that $s=c_{23}^2, t=c_{13}^2, r=\cos\delta$ in terms of the matrix U.

Then $Q=4b_1(1-b_1)^2 s(1-s)t(1-t)(1-r^2)$, and Ω corresponds to

$$\{(b_1,s,t,r):(b_1,s,t) \in [0,1]^3, -1 \le r \le 1\}.$$

The Jacobian for the change of variables is

$$\left|\frac{\partial(b_1,b_2,b_3,b_4)}{\partial(b_1,s,t,r)}\right| = 2(1-b_1)^2\sqrt{b_1 s(1-s)t(1-t)}.$$

Making use of the expressions derived in Ref. 31, we may write an explicit form for an integral of a continuous function g with respect to the measure $Q(b)^{k-3/2}$ for arbitrary $k>1/2$. Hence it is natural to introduce a one-parameter family of measures μ_k on $\mathcal{U}_3$ which satisfy

$$\int_{\mathcal{U}_3} g(B)d\mu_k = \int_{\Omega} g(b)Q(b)^{k-3/2}db. \tag{28}$$

Thus the case $k=3/2$ corresponds to the flat (Lebesgue) measure on $\mathcal{U}_3$, while for $k=1$ this expression reduces to (26), so μ_1 represents the measure induced by the Haar measure on $U(3)$.

Integral at the right hand side of (28) can be rewritten as

$$\int_{\Omega} g(b)Q(b)^{k-3/2}db = \int_0^1\int_0^1\int_0^1\int_{-1}^1 g(b_1,s(1-b_1),t(1-b_1),b_4) \times b_1^{k-1}(1-b_1)^{2k-1}[4s(1-s)t(1-t)]^{k-1}(1-r^2)^{k-3/2}\; dr\; ds\; dt\; db_1, \tag{29}$$

where b_4 is given by (27). Setting $g=1$ and using the standard beta integrals we determine the normalization constant (Ref. 31, Proposition 3.2),

$$h_k := \int_{\Omega} Q(b)^{k-3/2}db = \frac{\pi\Gamma(k)^3}{(2k-1)\Gamma(3k)}, \quad k > \frac{1}{2}. \tag{30}$$

Suppose $n=1,2,3,\ldots$, then

$$h_n = \frac{\pi((n-1)!)^3}{(2n-1)(3n-1)!},$$

$$h_{n+1/2} = \frac{\pi^2}{2n}\frac{\left[\left(\frac{1}{2}\right)_n\right]^2}{\left(n+\frac{1}{2}\right)_{2n+1}},$$

where $(x)_n := \Pi_{i=1}^n (x+i-1)$ stands for the Pochhammer symbol. In particular,

$$h_{3/2} = \frac{\pi^2}{3\cdot 5\cdot 7} = \frac{\pi^2}{105} \tag{31}$$

gives the volume of $\mathcal{U}_3$ considered as a subset of $\mathbb{R}^4$. This is multiplied by 9 to produce the volume relative to $\mathbb{R}^9$ (Ref. 5)—see Appendix A.

Thus the ratio of the 4D volume of $\mathcal{U}_3$ in $\mathbb{R}^9$ to the volume of all bistochastic matrices is

$$\frac{\mathrm{vol}(\mathcal{U}_3)}{\mathrm{vol}(\mathcal{B}_3)} = 9\times\frac{\pi^2}{105}\Big/\left(\frac{9}{8}\right) = \frac{8\pi^2}{105} = 0.751\,969\ldots . \tag{32}$$

This is in agreement with the outcome of earlier numerical calculations [Ref. 27, Eq. (24)] which were based on roughly 10^7 sample points and yielded 0.7520 ± 0.0005.

For completeness let us add that the volume of the set $\mathcal{U}_3$ with respect to the flag-manifold measure μ_1 reads $h_1 = \pi/2$.

V. MEAN ENTROPY OF A N=3 UNISTOCHASTIC MATRIX

The Shannon entropy of an N-point probability vector $p=(p_1,\ldots,p_N)$ is defined by

$$S(p) = -\sum_{i=1}^{N} p_i \ln p_i. \tag{33}$$

This quantity measures to what extent the vector is mixed and varies from 0 for any pure vector $(1,0,\ldots,0)$ to $\ln N$ for the maximally mixed vector $p_* = (1/N,\ldots,1/N)$. In an analogous way one defines the entropy of a bistochastic matrix,

$$S(B) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N} B_{ij}\ln B_{ij}, \tag{34}$$

equal to the average entropy of its rows (or columns). For any permutation matrix P this entropy is equal to zero while its maximum value $\ln N$ is attained at the flat matrix W. The mean entropy of a bistochastic matrix was considered by Słomczyński[38] and later analyzed in Refs. 26 and 39.

To derive an expression for the average entropy of a unistochastic matrix with respect to any probability measure on $\mathcal{U}_3$ bi-invariant under the symmetric group $\mathcal{S}_3$, we observe that is equal to the expected value of $-3B_{11}\ln B_{11}$. This can be computed with respect to the probability measure $h_k^{-1}Q(b)^{k-3/2}db$ on Ω.

Let us denote the mean entropy by $\langle S\rangle_{\Omega,k}$, where parameter k labels the measure defined in (28). From (29) specialized to functions of b_1 we obtain

$$\langle S\rangle_{\Omega,k} = \frac{-3\Gamma(3k)}{\Gamma(k)\Gamma(2k)}\int_0^1 b_1^k \ln b_1 (1-b_1)^{2k-1} db_1 = \psi(3k+1)-\psi(k+1), \tag{35}$$

where the digamma function reads $\psi(x) := \Gamma'(x)/\Gamma(x), x>0$.

The recurrence relation $\psi(x+1)=\psi(x)+1/x$ is used in the following: Suppose $n=1,2,3,\ldots$ then

$$\langle S\rangle_{\Omega,n}=\sum_{j=n+1}^{3n}\frac{1}{j},\quad \langle S\rangle_{\Omega,n+1/2}=\sum_{j=n+1}^{3n+1}\frac{2}{2j+1}. \tag{36}$$

In particular, the average entropy for the flag-manifold measure μ_1 gives $\langle S\rangle_{\Omega,1}=\frac{5}{6}=0.833\ldots$, This result coincides with the average entropy of random complex vectors[40,41] which form a unitary matrix.

The mean entropy with respect to the flat measure $\mu_{3/2}$ on $\mathcal{U}_3$ reads $\langle S\rangle_{\Omega,3/2}=2\left(\frac{1}{5}+\frac{1}{7}+\frac{1}{9}\right)=\frac{286}{315}=0.907\,936\,51\ldots$. This quantity was approximated as 0.908 in Ref. 27, Eq. (25).

These data can be compared with the maximal entropy $S_{\max}=\ln 3\approx 1.099$, characteristic of the flat matrix W of van der Waerden.

For comparison let us now compute the mean entropy with respect to the Lebesgue measure averaged over the entire set $\mathcal{B}_3$ of bistochastic matrices.

Straightforward calculation allows us to integrate functions of b_1,b_2 over $\mathcal{B}_3$ with respect to the flat measure $db:=\prod_{i=1}^{4}db_i$. In general, we consider an arbitrary function f, integrable on the triangle with vertices {(0, 0), (1, 0) (0, 1)}. The result is

$$\int_{\mathcal{B}_3}f(b_1,b_2)db=\int_0^1 db_1\int_0^{1-b_1}f(b_1,b_2)(b_1b_2+(b_1+b_2)(1-b_1-b_2))db_2. \tag{37}$$

The corollary stated in the appendices specializes this to $\int_0^1 f(b_1)db$ and allows us to find the volume of the Birkhoff polytope. Thus relative to $\mathbb{R}^9$ one has $\mathrm{vol}_4(\mathcal{B}_3)=9/8$.

Formula (37) can also be used to compute the average entropy, equal to the expected value of $(-3b_1\ln b_1)$. The analytic result,

$$8\int_{\mathcal{B}_3}(-3b_1\ln b_1)db=\frac{53}{60},$$

agrees with the numerical estimation of 0.883 obtained earlier in Ref. 27. Note that this number is smaller than the mean entropy $\langle S\rangle_{\Omega,3/2}$ averaged over the set of unistochastic matrices, since these bistochastic matrices which do not belong to $\mathcal{U}_3$ are located close to the boundary of the Birkhoff polytope and are characterized by small entropy.

Let us now turn to the generalized entropy defined for any probability vector $\{p_i:1\le i\le N\}$,

$$S_q:=\frac{1}{q-1}\sum_{i=1}^{N}(p_i-p_i^q). \tag{38}$$

The parameter $q\neq 1$ is assumed to be non-negative. In the limiting case the generalized entropy converges to the standard (Shannon) entropy, $\lim_{q\to 1}S_q=-\sum_{i=1}^{N}p_i\ln p_i$.

Applying definition (38) to a unistochastic matrix $B\in\mathcal{U}_3$ similarly to the ordinary case, we let

$$S_q=\frac{1}{3(q-1)}\sum_{i=1}^{3}\sum_{j=1}^{3}(B_{ij}-B_{ij}^q).$$

We compute the expected value of this expression with respect to $h_k^{-1}Q(b)^{k-3/2}db$ on Ω for $k>\frac{1}{2}$.

It is the same as the expected value of $[3/(q-1)](B_{11}-B_{11}^q)$, indeed

$$\langle S_q\rangle_{\Omega,k} = \frac{3\Gamma(3k)}{(q-1)\Gamma(k)\Gamma(2k)}\int_0^1 (b_1 - b_1^q)b_1^{k-1}(1-b_1)^{2k-1}db_1 = \frac{1}{q-1}\left(1 - \frac{3\Gamma(k+q)\Gamma(3k)}{\Gamma(k)\Gamma(3k+q)}\right).$$

When k is an integer number n or a half-integer $k=n+\frac{1}{2}$ this expression is a rational function of q and can be expressed with help of the Pochhammer symbol $(x)_n$ defined above,

$$\langle S_q\rangle_{\Omega,n} = \frac{1}{q-1}\left(1 - \frac{(3n)!}{n!\,(q+n)_{2n}}\right), \tag{39}$$

$$\langle S_q\rangle_{\Omega,n+1/2} = \frac{1}{q-1}\left(1 - \frac{3\left(\frac{1}{2}+n\right)_{2n+1}}{\left(q+\frac{1}{2}+n\right)_{2n+1}}\right). \tag{40}$$

In particular, taking $n=1$ we arrive at handy expressions for the mean generalized entropies averaged over the Haar measure and flat measure, respectively, which allow for an explicit partial fraction expansion,

$$\langle S_q\rangle_{\Omega,1} = \frac{q+4}{(q+1)(q+2)} = \frac{3}{q+1} - \frac{2}{q+2},$$

$$\langle S_q\rangle_{\Omega,3/2} = \frac{2(4q^2+34q+105)}{(2q+3)(2q+5)(2q+7)} = \frac{63}{4(2q+3)} - \frac{45}{2(2q+5)} + \frac{35}{4(2q+7)}. \tag{41}$$

For completeness we provide also an expression for the generalized entropy averaged over the set $\mathcal{B}_3$ with respect to the flat measure obtained with the help of Corollary 2,

$$\langle S_q\rangle_{\mathcal{B}_3} = \frac{2}{q+1} + \frac{4}{q+2} - \frac{9}{q+3} + \frac{4}{q+4}. \tag{42}$$

These entropies characterize well the distribution of matrices generated by these measures. In particular, a comparison of both expressions in (42) shows that the Haar measure on $U(3)$ populates the region close to the boundary of $\mathcal{U}_3$ more densely than the vicinity of the flat matrix W around its center. Since the squared Jarlskog invariant J^2 is by construction equal to zero at the boundary of $\mathcal{U}_3$, we may expect that its mean value over the flat measure μ_1 is smaller than the average with respect to the Haar measure $\mu_{3/2}$. As shown in Sec. VI this is indeed the case.

VI. DISTRIBUTION OF THE JARLSKOG INVARIANT

The value of the Jarlskog invariant and its square at a cross section of the set $\mathcal{U}_3$ of unistochastic matrices is shown in Fig. 4. Recent papers of Gibbons *et al.*[12,30] analyzed squared Jarlskog invariant[8,9] averaged over several probability measures on the set of unitary matrices. In particular, these authors computed the expectation value $\langle J^2\rangle$ averaged over the "flag-manifold" measure induced by the Haar measure on $U(3)$ and analyzed numerically the probability distribution $P(|J|)$ with respect to this measure. In this section we proceed one step further and derive an analytical formula for this probability distribution.

We shall start computing the moments of the distribution of the variable $Q=4J^2$ defined in (16) as a function of a random unistochastic matrix B. This task is rather simple, since we can express the moments of Q with respect to any measure μ_k by the coefficients h_k defined in (30),

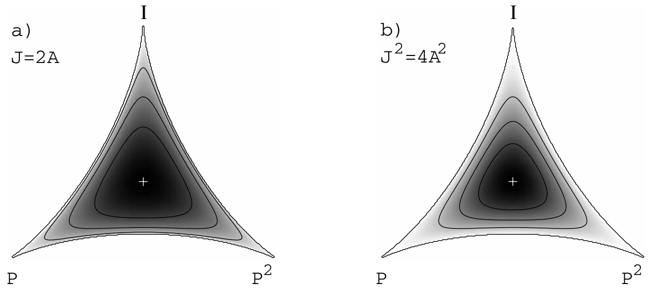

FIG. 4. The absolute value of the Jarlskog invariant $|J|$ (a), and its square J^2 (b), at the cross section of $\mathcal{U}_3$ along the plane formed by two permutation matrices and the identity. Dark color denotes high values of $|J|$ and J^2. The maximum is achieved at the van der Waerden matrix W located at the center of the deltoid. Note that outside the deltoid $J^2<0$ and the bistochastic matrix is not unistochastic.

$$\langle Q^n\rangle_k = \frac{h_{k+n}}{h_k} = 3^{-3n}\frac{\left(k-\frac{1}{2}\right)\Gamma\left(k+\frac{1}{3}\right)\Gamma\left(k+\frac{2}{3}\right)\Gamma(k+n)^2}{\left(k+n-\frac{1}{2}\right)\Gamma\left(k+n+\frac{1}{3}\right)\Gamma\left(k+n+\frac{2}{3}\right)\Gamma(k)^2}$$

$$= 3^{-3n}\frac{\left(k-\frac{1}{2}\right)(k)_n^2}{\left(k+n-\frac{1}{2}\right)\left(k+\frac{1}{3}\right)_n\left(k+\frac{2}{3}\right)_n}, \qquad (43)$$

where k determines the measure (28) while $n=0,1,2,\ldots$.

Setting $k=1$ and $n=1$ we find that the mean squared Jarlskog invariant, averaged over the Haar measure, reads $\langle J^2\rangle_1=\langle Q/4\rangle_1=1/720=1.389\times 10^{-3}$ in consistence with Ref. 30, Eq. (75). For comparison note that the average over the flat measure yields a larger value, $\langle J^2\rangle_{3/2}=3/1144=2.622\times 10^{-3}$. In general, the flat measure favors larger values of $|J|$ as it is shown in Fig. 5.

Having at our disposal the complete set of the moments of Q, we will determine the exact distribution function $P(Q)$ in terms of hypergeometric and related functions. To avoid nuisance factors in the calculations we will consider the random variable $X:=27Q=108J^2$ so that X takes values in $[0,1]$.

We know that the following relations hold for any $k>1/2$ and $n=0,1,2,\ldots$:

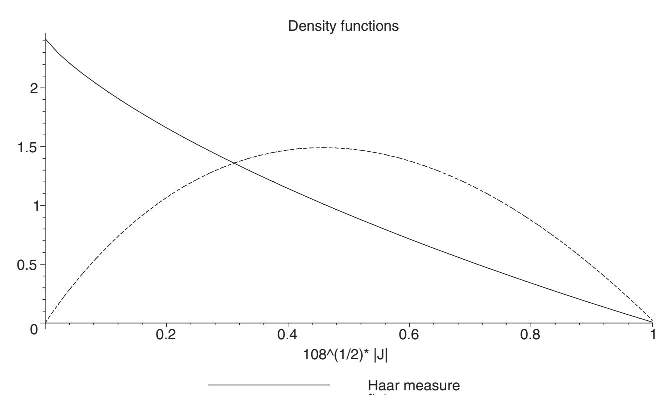

FIG. 5. Probability distribution $P_k(|J|)$ of the Jarlskog invariant for random unistochastic matrices generated according to the Haar measure μ_1 and the flat measure $\mu_{3/2}$.

$$\left(k-\frac{1}{2}\right)\int_0^1 x_0^n x_0^{k-3/2}dx_0 = \frac{\left(k-\frac{1}{2}\right)}{\left(k+n-\frac{1}{2}\right)},$$

$$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 x^{n+\alpha-1}(1-x)^{\beta-1}dx = \frac{(\alpha)_n}{(\alpha+\beta)_n},(\alpha,\beta>0).$$

In view of the expression (43) for the moments the above relations allow us to find an alternative representation of the desired probability distribution $P(X)$.

Let X_0, X_1, X_2 be independent random variables with the densities

$$f_0(x) = \left(k-\frac{1}{2}\right)x^{k-3/2}, \tag{44}$$

$$f_1(x) = \frac{\Gamma\left(k+\frac{1}{3}\right)}{\Gamma(k)\Gamma\left(\frac{1}{3}\right)}x^{k-1}(1-x)^{-2/3}, \tag{45}$$

$$f_2(x) = \frac{\Gamma\left(k+\frac{2}{3}\right)}{\Gamma(k)\Gamma\left(\frac{2}{3}\right)}x^{k-1}(1-x)^{-1/3}, \tag{46}$$

respectively, each being defined on $0 \le x \le 1$.

Then X has the same moments and the same probability distribution as the product $X_0X_1X_2$. This step, justified in Appendix B, enables us to arrive at the key result of this section: an explicit expression for the probability distribution for $X=108J^2$, where J denotes the Jarlskog invariant of a random unistochastic matrix generated according to the measure μ_k,

$$P_k(X) = c_k\left(k-\frac{1}{2}\right)X^{k-3/2}\int_X^1 {}_2F_1\left(\frac{1}{3},\frac{2}{3};1;1-t\right)t^{-1/2}dt. \tag{47}$$

It is assumed here that $0 < X \le 1$ and $k > 1/2$, the symbol ${}_2F_1$ stands for the hypergeometric function, while the normalization constant reads

$$c_k := \frac{\Gamma\left(k+\frac{1}{3}\right)\Gamma\left(k+\frac{2}{3}\right)}{\Gamma(k)^2}. \tag{48}$$

In Appendix B we made use of this expression to determine an explicit expansion for the density function $P_k(X)$. These results allowed us to show the distributions $P_1(X)$ and $P_{3/2}(X)$ in Fig. 5. Observe that in the case $k=1$ we obtain an expression for the distribution of $|J|$, considered as a function of a random unitary matrix U distributed with respect to the Haar measure on $U(3)$. For small values of $|J|$ this distribution behaves as $P(|J|) \sim \alpha |J|^\lambda$ with $\lambda=0$ and $\alpha = 8\pi \approx 25.133$.

To make a direct connection with the results of Ref. 30, we reproduce here the formula for the integrated probability distribution. For any $0 < y \le 1/6\sqrt{3}$ the probability of finding a unitary matrix U distributed according to the Haar measure on $U(3)$ with $|J(B)|$ less or equal y reads

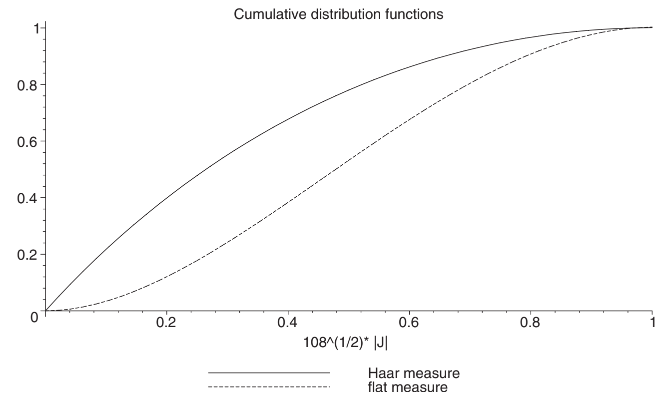

FIG. 6. Cumulative probability distribution for the functions plotted in Fig. 5.

$$P_1\{|J| \leq y\} = 8\pi y + \left\{ 24 \ln(4y^2)(y^2 + 4y^4 + 96y^6 + \cdots) - 72y^2 + 128y^4 + \frac{24\,384}{5}y^6 + \cdots \right\}. \tag{49}$$

In Appendix B we derive this formula obtained for the measure μ_1 on $\mathcal{U}_3$ as well as an analogous result for the flat measure $\mu_{3/2}$,

$$P_{3/2}\{|J| \leq y\} = 420y^2 - \frac{4480}{\pi}y^3 \cdots + \frac{1680}{\pi}y^3 \ln(4y^2) + \cdots. \tag{50}$$

Both cumulative distribution functions are compared in Fig. 6.

Recent experimental data show that the observed value of the Jarlskog invariant reads[42]

$$J_{\text{obs}} = 3.08[+0.16, -0.18] \times 10^{-5}. \tag{51}$$

This concrete number can be now compared with the probability distribution (49). A 95% confidence interval for $|J|$ is about $[0.00202, 1/6\sqrt{3} \approx 0.096]$, while the probability of getting a value of $|J|$ outside this interval is 5%.

Moreover, we get an explicit estimate for the probability of obtaining at random a unitary matrix, such that the absolute values of its Jarlskog invariant are smaller than the observed value,

$$P\{|J| \leq J_{\text{obs}}\} = 7.74[+0.41, -0.45] \times 10^{-4}. \tag{52}$$

The statistical hypothesis that the CKM matrix arises from the probability experiment of producing a random unitary matrix, with respect to Haar measure on $U(3)$, is rejected at the descriptive significance level of 0.08%.

Another benchmark introduced by Gibbons *et al.*,[30] $P_{\text{flag}}(|J| \leq 10^{-4}) \approx 2.5085 \times 10^{-3}$, is consistent with numerical data obtained in Eq. (92) of that paper. Note that Gibbons *et al.*[30] constructed several probability models for which values as small or smaller than J_{obs} are more likely.

For comparison we note that the flat measure $\mu_{3/2}$ in the set of unistochastic matrices yields a smaller probability. Using this measure $(k=\frac{3}{2})$ we obtain $P_{3/2}\{|J| \leq J_{\text{obs}}\} \approx 3.98 \times 10^{-7}$. Indeed this could be viewed as statistical evidence that the transition probabilities in the CKM matrix do arise from a unitary matrix. Specifically the so-called likelihood ratio test applied to the two probability densities for $|J|$ induced by $d\mu_1$ and $(8\pi^2/105)d\mu_{3/2}$ [the factor $(8\pi^2/105)$ comes from $P\{Q<0\} = 1-(8\pi^2/105)$ for the flat measure on $\mathcal{B}_3$] at $|J| = 3.08 \times 10^{-5}$ results in a factor of about 1200. This value is obtained from the formulas for $f_0(x)$ in Appendix B.

Let us express the Jarlskog invariant (10) by the standard parameters (19) of a unitary matrix of order of 3,

$$J(U) = -c_{12}c_{23}c_{13}s_{12}^2 s_{23}s_{13} \sin \delta. \tag{53}$$

Observed value of the Jarlskog invariant for the CKM matrix does not imply that the CP violating phase δ_{CKM} had to be small. In fact, $\delta_{\text{CKM}} \in [62°, 100°]$ so even the value $\pi/2$ is not ruled out—see, e.g., Ref. 43. Hence the small value (51) is due to the angles θ_{ij} in (19), which determine the bistochastic matrix.

Thus we are going to conclude this section with a simple statistical statement: The CKM matrix should not be considered as a generic unitary matrix drawn at random with respect to the Haar measure on $U(3)$. Furthermore, the matrix B_{CKM} of squared entries of V_{CKM} is rather unlikely to be an ordinary bistochastic matrix generated at random with respect to the flat measure in this set.

VII. CONCLUDING REMARKS

In this work we have analyzed the Birkhoff polytope $\mathcal{B}_3$ of $N=3$ bistochastic matrices and its subset $\mathcal{U}_3$ of unistochastic matrices. This set contains these bistochastic matrices which arise from squared moduli of entries of a unitary matrix. We have improved the result of Ref. 27 by computing the exact volume of $\mathcal{U}_3$ with respect to the flat (Lebesgue) measure and found that it takes more than three quarters of the volume of the Birkhoff polytope $\mathcal{B}_3$.

We have introduced a one-parameter family of probability measures μ_k into the set $\mathcal{U}_3$ of unistochastic matrices. Among the measures (28) the distinguished ones are the uniform (flat) measure $\mu_{3/2}$ and the measure μ_1, induced by the Haar measure on $U(3)$. Furthermore, the measure μ_k obtained in the limit $k \to \frac{1}{2}$ coincides with the measure induced by the Haar measure on the orthogonal group $O(3)$.

We derived explicit formulas which allow us to compute expectation values of a smooth function of an entry of B with respect to these measures. In this way we derived exact expressions for the mean entropy and the generalized entropy of a random unistochastic matrix with respect to the measures μ_k. These values can serve as a reference values in studying properties of concrete unitary and bistochastic matrices of order of 3, used in the theory of quantum information.

In high energy physics and the theory of CP symmetry breaking, one works with unitary matrices of order of 3 and characterizes them by the Jarlskog invariant (10). Computing all the moments of the squared Jarlskog invariant J^2 with respect to the measure μ_k, we could represent the probability distribution $P_k(J)$ as an integral (47) of the hypergeometric function ${}_2F_1$. Expanding this function in a series and integrating it term by term, we arrived at an explicit representation of the desired probability distributions. In particular, working with the Haar measure μ_1 we could derive analytical results on the distribution $P_1(|J|)$ consistent with the numerical results earlier obtained in Gibbons *et al.*[30] Our results support then the observation, that the unitary CKM matrix V_{CKM}, which describes the violation of the CP symmetry, should not be regarded as a generic unitary matrix of order of 3.

ACKNOWLEDGMENTS

It is a pleasure to thank I. Bengtsson and W. Tadej for numerous stimulating discussions and helpful correspondence. We acknowledge financial support by the special Grant No. DFG-SFB/38/2007 of Polish Ministry of Science and Higher Education and an European Research project COCOS (K.Ż.).

APPENDIX A: EXTREME VALUES OF THE PARAMETER Q

Consider a bistochastic matrix M of order three parametrized by (7) and the function $Q(M)$ defined in (16). The aim of this section is to show that for any $M \in \mathcal{B}_3$ this function takes values in $[-1/16, 1/27]$. Note that if the matrix is unistochastic, $M \in \mathcal{U}_3$ then Q is non-negative and is proportional to the squared area of the unitarity triangle.

It is straightforward to show that Q is invariant under transposition and permutation of rows or columns [for example, replacing (b_2,b_4) by $(1-b_1-b_2,1-b_3-b_4)$]. Let us first introduce parameters b_1,s,t,x with $0\leq b_1,s,t\leq 1$ and conditions on x to be determined. Motivated by the unistochastic situation let

$$b_2=(1-b_1)s,$$

$$b_3=(1-b_1)t,$$

$$b_4=(1-s)(1-t)+b_1st+x.$$

Four more conditions must be satisfied for $M(b)$ to be bistochastic [the inequalities $M(b)_{13}\geq 0$ and $M(b)_{31}\geq 0$ are already satisfied]. The simultaneous inequalities $M(b)_{23}\geq 0$ and $M(b)_{32}\geq 0$ are equivalent to

$$x\leq \min(u_1,u_2),$$

$$u_1 := s(1-t)+b_1t(1-s),$$

$$u_2 := t(1-s)+b_1s(1-t), \tag{A1}$$

and $\{M(b)_{22}\geq 0, M(b)_{33}\geq 0\}$ is equivalent to

$$-x\leq \min(\ell_1,\ell_2),$$

$$\ell_1 := (1-s)(1-t)+b_1st,$$

$$\ell_2 := st+b_1(1-s)(1-t). \tag{A2}$$

With these parameters,

$$Q(b)=Q'(b_1,s,t,x) := -(1-b_1)^2(x^2-4b_1st(1-s)(1-t)).$$

For fixed b_1,s,t this is decreasing in x^2; thus the maximum value occurs at $x=0$, and then maximizing over b_1,s,t, we obtain $Q(b)=\frac{1}{27}$ when $b_1=\frac{1}{3},s=\frac{1}{2}=t$ [that is, $M(b)_{ij}=\frac{1}{3}$ for $1\leq i,j\leq 3$].

Next we show that the minimum value of Q on $\mathcal{B}_3$ is $-\frac{1}{16}$, achieved at the Schur matrix (1): by permutations of rows or columns, and then transposition of M, if necessary, we may assume $0\leq t\leq s\leq\frac{1}{2}$. In this triangle the bounds $-\ell_2\leq x\leq u_2$ apply. Rather than considering $\min(Q'(b_1,s,t,-\ell_2),Q'(b_1,s,t,u_2))$ in this region, we will minimize $Q'(b_1,s,t,-\ell_2)$ in the triangle bounded by $t=0,t=s,t=1-s$ [vertices (0, 0), $(\frac{1}{2},\frac{1}{2})$, (1, 0)]; this works because $=Q'(b_1,s,t,u_2(b_1,s,t))=Q'(b_1,1-s,t,-\ell_2(b_1,1-s,t))$ [writing ℓ_2,u_2 as functions of (b_1,s,t)]. Then

$$Q'(b_1,1-s,t,-\ell_2(b_1,1-s,t))=-(1-b_1)^2(b_1(1-s)(1-t)-st)^2.$$

As a function of (s,t), Q' cannot have an interior minimum, so it suffices to check the edges of the triangle. On the edge $t=0$ we have $Q'=-(1-b_1)^2b_1^2(1-s)^2$ with minimum value of $-\frac{1}{16}$ at $b_1=\frac{1}{2},s=0$.

On the edge $t=1-s$ we obtain $Q'=-(1-b_1)^4s^2(1-s)^2$ with minimum value of $-\frac{1}{16}$ at $b_1=0,s=\frac{1}{2}=t$ (the Schur matrix for $x=-\frac{1}{4}$).

On the edge $s=t,Q'=-(1-b_1)^2(b_1(1-s)^2-s^2)^2$. This function has no interior minimum on the interval $0\leq s\leq\frac{1}{2}$. The end points $s=0$ and $s=\frac{1}{2}$ have already been considered.

There is a neat formula for the integral of functions of (b_1, b_2) over $\mathcal{B}_3$ with respect to the flat measure $db := \Pi_{i=1}^4 db_i$. The derivation of the formula involves adding over the four regions formed in the unit square by the lines $s=t, s+t=1$.

Proposition 1: Let f be integrable on $\{(x,y): x,y \geq 0, x+y \leq 1\}$, *then*

$$\int_{\mathcal{B}_3} f(b_1,b_2)db = \int_0^1 db_1 \int_0^{1-b_1} f(b_1,b_2)(b_1b_2 + (b_1+b_2)(1-b_1-b_2))db_2.$$

Observe that the integral kernel is an elementary symmetric function of $(b_1, b_2, 1-b_1-b_2)$.

Corollary 2: Let f be integrable on $[0, 1]$ *then*

$$\int_{\mathcal{B}_3} f(b_1)db = \frac{1}{6}\int_0^1 f(b_1)(1+5b_1)(1-b_1)^2 db_1.$$

Thus $\int_{\mathcal{B}_3} 1 db = \frac{1}{8}$.

When considering $\mathcal{B}_3$ as a subset of $\mathbb{R}^9$ the element of volume $db = \Pi_{i=1}^4 db_i$ is multiplied by 9. The map $B: \mathbb{R}^4 \to \mathbb{R}^9$, defined in Eq. (7), is affine onto a 4D linear manifold (translate of a subspace) and its Jacobian equals 9, calculated as the square root of the determinant of the Gram matrix of the images of the unit vectors relative to $B(\vec{0})$. For example,

$$b_1\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \to \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & -1 \end{bmatrix} + b_1\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix},$$

the Gram matrix is

$$\begin{bmatrix} 4 & 2 & 2 & 1 \\ 2 & 4 & 1 & 2 \\ 2 & 1 & 4 & 2 \\ 1 & 2 & 2 & 4 \end{bmatrix},$$

and its determinant equals 81.

In this way we may obtain the average entropy and directly derive expression (42) for the generalized entropy. Furthermore, $8\int_{\mathcal{B}_3} Q(b)db = \frac{1}{168}, 8\int_{\mathcal{B}_3} Q(b)^2 db = \frac{1}{5940}$, thus the standard deviation reads $\sigma_Q = 0.011\,53\ldots$.

Since the set $\mathcal{U}_3$ of unistochastic matrices is the subset of $\mathcal{B}_3$ for which $Q \geq 0$, equivalently

$$|x| \leq R := \{4b_1 st(1-s)(1-t)\}^{1/2},$$

we see that

$$\int_{\mathcal{U}_3} f(b_1)db = \int_0^1 f(b_1)(1-b_1)^2 db_1 \int_0^1 ds \int_0^1 dt \int_{-R}^{R} dx = \frac{\pi^2}{16}\int_0^1 f(b_1)b_1^{1/2}(1-b_1)^2 db_1,$$

and $\int_{\mathcal{U}_3} 1 db = \pi^2/105$ in agreement with Eq. (31). Observe that $R \leq \min(l_1, l_2, u_1, u_2)$ by the (well-known) inequality $2\sqrt{xy} \leq x+y$ for $x,y \geq 0$; this is the reason that the integral extends over $0 \leq s,t \leq 1$.

APPENDIX B: JARLSKOG INVARIANT AS A PRODUCT OF THREE RANDOM VARIABLES

In this appendix we show that the product of three random variables $X_0X_1X_2$ introduced in (44)–(46) has the same probability distribution as the rescaled squared Jarlskog invariant X

$:=27Q=108J^2$ of random unistochastic matrices generated with respect to the measure μ_k defined in (28). We shall start by quoting the lemma on probability distribution of a product of two independent random variables.

Lemma 3: Suppose Y_1, Y_2 are random variables on $[0, 1$ *with densities g_i and c.d.f.'s G_i, $i=1,2$ [that is, $G_i(x)=\int_0^x g_i(t)dt=P\{Y_i\le x\}$, $0\le x\le 1$]. Then the density for Y_1Y_2 is $\int_x^1 g_1(t)g_2(x/t)1/tdt$.*

Let us apply this lemma, which can be proven by direct integration, to a random variable $Y_2=X_0$ distributed as in (44).

Corollary 4: If $g_2(t)=(k-\frac{1}{2})t^{k-3/2}$ with $k>\frac{1}{2}$ then the density of Y_1Y_2 is

$$\left(k-\frac{1}{2}\right)x^{k-3/2}\int_x^1 g_1(t)t^{1/2-k}dt, \quad 0<x<1.$$

Making use of the normalization constant c_k introduced in (48), we can write down explicit form for the density of the product X_1X_2.

Proposition 5: The density f_{12} of X_1X_2 is $c_k x_2^{k-1}F_1(\frac{1}{3},\frac{2}{3};1;1-x), 0<x\le 1$.

Proof: By Lemma 3, the density is

$$f_{12}(x)=\frac{c_k}{\Gamma\left(\frac{1}{3}\right)\Gamma\left(\frac{2}{3}\right)}\int_x^1 t^{k-1}(1-t)^{-1/3}\left(\frac{x}{t}\right)^{k-1}\left(1-\frac{x}{t}\right)^{-2/3}\frac{dt}{t}$$

$$=\frac{c_k x^{k-1}}{\Gamma\left(\frac{1}{3}\right)\Gamma\left(\frac{2}{3}\right)}\int_x^1 (1-t)^{-1/3}(t-x)^{-2/3}t^{-1/3}dt$$

$$=\frac{c_k x^{k-1}}{\Gamma\left(\frac{1}{3}\right)\Gamma\left(\frac{2}{3}\right)}\int_0^1 (1-s)^{-1/3}s^{-2/3}(x+s(1-x))^{-1/3}ds,$$

using the substitution $t=x+s(1-x)$. Now restrict x to $\frac{1}{2}<x\le 1$ then $0\le(1-x)/x<1$ and we can expand

$$(x+s(1-x))^{-1/3}=x^{-1/3}\left(1-s\frac{x-1}{x}\right)^{-1/3}=x^{-1/3}\sum_{n=0}^{\infty}\frac{\left(\frac{1}{3}\right)_n}{n!}s^n\left(\frac{x-1}{x}\right)^n.$$

Integrating term by term we obtain

$$f_{12}(x)=\frac{c_k x^{k-1}}{\Gamma\left(\frac{1}{3}\right)\Gamma\left(\frac{2}{3}\right)}\sum_{n=0}^{\infty}\frac{\left(\frac{1}{3}\right)_n\Gamma\left(\frac{1}{3}+n\right)\Gamma\left(\frac{2}{3}\right)}{n!\,\Gamma(n+1)}(x-1)^n x^{-n-1/3}$$

$$=c_k x^{k-1}\sum_{n=0}^{\infty}\frac{\left(\frac{1}{3}\right)_n\left(\frac{1}{3}\right)_n}{n!\,n!}(-1)^n\sum_{j=0}^{\infty}\frac{\left(\frac{1}{3}+n\right)_j}{j!}(1-x)^{n+j}$$

$$=c_k x^{k-1}\sum_{m=0}^{\infty}\frac{\left(\frac{1}{3}\right)_m}{m!}(1-x)^m\sum_{n=0}^{m}\frac{\left(\frac{1}{3}\right)_n m!\,(-1)^n}{n!\,(1)_n(m-n)!},$$

where the summation variables are changed to n and $m=n+j$. The inner sum is evaluated with the Chu–Vandermonde sum,

$$\sum_{n=0}^{m}\frac{\left(\frac{1}{3}\right)_n m\,!\,(-1)^n}{n\,!\,(1)_n(m-n)!}=\sum_{n=0}^{m}\frac{\left(\frac{1}{3}\right)_n(-m)_n}{n\,!\,(1)_n}=\frac{\left(1-\frac{1}{3}\right)_m}{(1)_m}=\frac{\left(\frac{2}{3}\right)_m}{(1)_m}.$$

This shows that $f_{12}(x)=c_k x_2^{k-1}F_1\left(\frac{1}{3},\frac{2}{3};1;1-x\right)$ at least for $\frac{1}{2}<x\leq 1$, but both sides are analytic on $0<x<2$ so the equality holds for $0<x\leq 1$. ■

To derive an expression for the density of the triple product $X=X_0X_1X_2$, we need to combine Lemma 3 with Corollary 4. Hence we can write $f(x)=\left(k-\frac{1}{2}\right)x^{k-3/2}\int_x^1 f_{12}(t)t^{1/2-k}dt$ and $f_{12}(t)=c_k t_2^{k-1}F_1\left(\frac{1}{3},\frac{2}{3};1;1-t\right)$ since $t^{k-1}t^{1/2-k}=t^{-1/2}$. This completes the proof of formula (47) for the distribution of the rescaled squared Jarlskog invariant $X:=27Q=108J^2$.

Following Gibbons *et al.*[30] we shall now concentrate on the probability distribution for the absolute value of the Jarlskog invariant $|J|$ equal to $(X/108)^{1/2}$. Let $f_0(x)$ denote the density function of $X^{1/2}$, thus $f_0(x)=2xf(x^2)$. It is not hard to compute a series for $f_0(x)$ when x is near 1. We change the variable of integration $t=(1-s)^2$ and obtain

$$f_0(x)=4c_k\left(k-\frac{1}{2}\right)x^{2k-2}\int_0^{1-x}{}_2F_1\left(\frac{1}{3},\frac{2}{3};1;s(2-s)\right)ds.$$

We expand $(s(2-s))^j$ (for $j=1,2,\ldots$), collect the coefficients of s^m, and integrate term by term to get

$$\int_0^{1-x}{}_2F_1\left(\frac{1}{3},\frac{2}{3};1;s(2-s)\right)ds=\sum_{m=0}^{\infty}\frac{(1-x)^{m+1}}{m+1}\sum_{j=0}^{\lfloor m/2\rfloor}\frac{(-1)^j\left(\frac{1}{3}\right)_{m-j}\left(\frac{2}{3}\right)_{m-j}2^{m-2j}}{(m-j)\,!\,(m-2j)\,!\,j!},$$

there is no nice formula for the inner (j-) sum. Thus

$$f_0(x)=4c_k\left(k-\tfrac{1}{2}\right)x^{2k-2}(1-x)\left\{1+\tfrac{2}{9}(1-x)+\tfrac{22}{243}(1-x)^2+\tfrac{310}{6561}(1-x)^3+\cdots\right\}$$

for x near 1 (that is, not too close to zero). When $k=1$ we have $c_1=\Gamma(4/3)\Gamma(5/3)=4\pi\sqrt{3}/27$ and

$$f_0(x)=\frac{8\pi\sqrt{3}}{27}(1-x)\left(1+\frac{2}{9}(1-x)+\frac{22}{243}(1-x)^2+\frac{310}{6561}(1-x)^3+\cdots\right).$$

Lemma 6: $\int_0^1{}_2F_1\left(\frac{1}{3},\frac{2}{3};1;1-t\right)t^{-1/2}dt=3.$

Proof: Indeed

$$\int_0^1{}_2F_1\left(\frac{1}{3},\frac{2}{3};1;1-t\right)t^{-1/2}dt=\sum_{n=0}^{\infty}\frac{\left(\frac{1}{3}\right)_n\left(\frac{2}{3}\right)_n}{n\,!\,n!}\int_0^1(1-t)^n t^{-1/2}dt=\sum_{n=0}^{\infty}\frac{\left(\frac{1}{3}\right)_n\left(\frac{2}{3}\right)_n}{n\,!\,n!}\frac{\Gamma(n+1)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(n+\frac{3}{2}\right)}$$

$$=\frac{\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{3}{2}\right)}\sum_{n=0}^{\infty}\frac{\left(\frac{1}{3}\right)_n\left(\frac{2}{3}\right)_n}{n\,!\left(\frac{3}{2}\right)_n}=\frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{7}{6}\right)\Gamma\left(\frac{5}{6}\right)}$$

$$=\frac{\pi}{\frac{1}{6}\Gamma\left(\frac{1}{6}\right)\Gamma\left(\frac{5}{6}\right)}=6\sin\frac{\pi}{6}=3.$$

We used the Gauss sum ${}_2F_1(a,b;c;1)=\Gamma(c)\Gamma(c-a-b)/\Gamma(c-a)\Gamma(c-b)$ for $c>a+b$ and the equation $\Gamma(t)\Gamma(1-t)=\pi/\sin \pi t$. ■

Thus $f(x)=c_k\left(k-\frac{1}{2}\right)x^{k-3/2}\left(3-\int_0^x {}_2F_1\left(\frac{1}{3},\frac{2}{3};1;1-t\right)t^{-1/2}dt\right)$. To analyze the behavior for x near zero, we use the classical formulas for the hypergeometric series ${}_2F_1(a,b;c;t)$ at the singular point $t=1$. The special case $c=a+b$ is more complicated (see Ref. 44, p. 257),

$$ {}_2F_1\left(\frac{1}{3},\frac{2}{3};1;1-t\right)=\frac{\Gamma(1)}{\Gamma\left(\frac{1}{3}\right)\Gamma\left(\frac{2}{3}\right)}\sum_{n=0}^{\infty}\frac{\left(\frac{1}{3}\right)_n\left(\frac{2}{3}\right)_n}{n\,!\,n!}(A_n-\ln t)t^n, $$

$$ A_n := 2\psi(n+1)-\psi\left(n+\frac{1}{3}\right)-\psi\left(n+\frac{2}{3}\right). $$

By the triplication formula for the ψ-function [recall $\psi(t)=(d/dt)\Gamma(t)/\Gamma(t)$],

$$ \psi\left(n+\tfrac{1}{3}\right)+\psi\left(n+\tfrac{2}{3}\right)=3\psi(3n)-\psi(n)-3\ln 3=3\psi(3n+1)-\psi(n+1)-3\ln 3, $$

because $\psi(t+1)=\psi(t)=1/t$ for $t>0$; the latter formula is valid for $n\geq 0$. Thus,

$$ A_n=3(\psi(n+1)-\psi(3n+1))+3\ln 3=-\sum_{j=n+1}^{3n}\frac{3}{j}+3\ln 3. $$

Also

$$ \int_0^x (A_n-\ln t)t^{n-1/2}dt=\frac{2x^{n+1/2}}{2n+1}\left(-\ln x+A_n+\frac{2}{2n+1}\right). $$

Thus

$$ f(x)=c_k\left(k-\frac{1}{2}\right)x^{k-3/2}\left\{3-\frac{\sqrt{3}}{2\pi}\sum_{n=0}^{\infty}\frac{\left(\frac{1}{3}\right)_n\left(\frac{2}{3}\right)_n}{n\,!\,n!}\frac{2x^{n+1/2}}{2n+1}\left(-\ln x+A_n+\frac{2}{2n+1}\right)\right\}, $$

$$ f_0(x)=2c_k\left(k-\frac{1}{2}\right)x^{2k-2}\left\{3-\frac{\sqrt{3}}{2\pi}\sum_{n=0}^{\infty}\frac{\left(\frac{1}{3}\right)_n\left(\frac{2}{3}\right)_n}{n\,!\,n!}\frac{2x^{2n+1}}{2n+1}\left(-2\ln x+A_n+\frac{2}{2n+1}\right)\right\}, $$

and we have found the density function of $x=\sqrt{X}=6\sqrt{3}|J|$ exhibiting the behavior for x near zero.

In the expression for f_0 the first few terms inside the braces { } are

$$ 3-\frac{\sqrt{3}}{2\pi}\left[-\ln\left(\frac{x^2}{27}\right)\left(2x+\frac{4}{27}x^3+\frac{4}{81}x^5\right)+4x-\frac{22}{81}x^3-\frac{49}{405}x^5\right]. $$

The cumulative distribution function $F_0(x)=P\{\sqrt{X}<x\}=P\{|J|<x/6\sqrt{3}\}$ is

$$ F_0(x)=2c_k\left(k-\frac{1}{2}\right)x^{2k-1} $$
$$ \times\left\{\frac{3}{2k-1}-\frac{\sqrt{3}}{2\pi}\sum_{n=0}^{\infty}\frac{\left(\frac{1}{3}\right)_n\left(\frac{2}{3}\right)_n x^{2n+1}}{n\,!\,n\,!\,(2n+1)(n+k)}\left(-\ln\frac{x^2}{27}-\sum_{j=n+1}^{3n}\frac{3}{j}+\frac{2}{2n+1}+\frac{1}{n+k}\right)\right\}. $$

The important cases are

(1) $k=1$, Haar measure on $U(3)$,

$$F_0(x) = \frac{4\pi\sqrt{3}}{9}x - \frac{2}{9}\left(-\ln\left(\frac{x^2}{27}\right)\left(x^2 + \frac{1}{27}x^4 + \frac{2}{243}x^6 + \cdots\right) + 3x^2 - \frac{4}{81}x^4 - \frac{127}{7290}x^6 + \cdots\right),$$

(2) $k=\frac{3}{2}$, flat measure,

$$F_0(x) = \frac{70}{27}x^2\left\{\frac{3}{2} - \frac{\sqrt{3}}{2\pi}\left(-\ln\left(\frac{x^2}{27}\right)\left(\frac{2}{3}x + \frac{4}{135}x^3 + \cdots\right) + \frac{16}{9}x - \frac{86}{2025}x^3 + \cdots\right)\right\}.$$

Introducing a new variable, $y=x/6\sqrt{3}$, we arrive at expressions (49) and (50) presented in Sec. VI.

APPENDIX C: CONJECTURES OF MEASURES IN HIGHER DIMENSIONS

The method used in Sec. IV relies on one of the authors'[31] construction of a linear operator commuting with the action of $\mathcal{S}_3$, mapping homogeneous polynomials in three variables to homogeneous polynomials of the same degree, and depending on a parameter k (a particular case of the "Dunkl intertwining operator"). This operator is realized as an integral over $U(3)$. The case $k=1$ is based on a formula of Harish–Chandra [see Helgason (Ref. 45, p. 328),

$$\int_{U(N)} \exp(\mathrm{Tr}(D(x)UD(y)U^*))dm(U) = \frac{c_N}{a(x)a(y)}\sum_{w\in\mathcal{S}_N}\det(w)\exp(\langle xw,y\rangle),$$

where the symmetric group on N objects is identified with the set of permutation matrices in $O(N)$, for $x,y\in\mathbb{R}^N$ the inner product is $\langle x,y\rangle := \sum_{j=1}^N x_j y_j$, $D(x)$ is the diagonal matrix with $D(x)_{jj}=x_j$, $a(x) := \prod_{1\le i<j\le N}(x_i - x_j)$, dm is Haar measure, and c_N is a constant. The relation to unistochastic matrices follows from the equation

$$\mathrm{Tr}(D(x)UD(y)U^*) = \sum_{i,j=1}^N x_i|U_{ij}|^2 y_j.$$

The aforementioned linear operator is known algebraically, that is, with some computational effort one can determine the action on any (low-degree!) polynomial. For $N=3$ we were able to find a parametrized family of measures to implement the operator, roughly

$$p(x) \mapsto \int_{U(3)} p(xf(U))d\mu_k(U),$$

where $f(U)_{ij} := |U_{ij}|^2$ as in (8) for any polynomial p. It is known that

$$\int_{U(N)} g(|U_{ij}|^2)dm(U) = (N-1)\int_0^1 g(t)(1-t)^{N-2}dt$$

for any continuous function g and any matrix entry U_{ij}. As in Sec. V we can compute the average (generalized) entropy for the entries $|U_{ij}|^2$ with respect to Haar measure ($q\neq 1$),

$$\langle S_q\rangle_{\mathrm{Haar}} = \frac{1}{N(q-1)}\int_{U(N)}\sum_{i,j=1}^{N}(|U_{ij}|^2-|U_{ij}|^{2q})dm(U)$$
$$= \frac{N(N-1)}{q-1}\int_0^1 (t-t^q)(1-t)^{N-2}dt = \frac{1}{q-1}\left(1-\frac{N!}{(q+1)_{N-1}}\right)$$
$$= N!\sum_{i=0}^{N-2}\frac{(-1)^i}{i!(N-2-i)!(i+2)(q+i+1)},$$

the last equation is the partial fraction decomposition. Also $\lim_{q\to 1}\langle S_q\rangle = \psi(N+1)-\psi(2) = \Sigma_{j=2}^{N} 1/j$, in agreement with the known results for the mean entropy of random complex vectors distributed according to the unitarily invariant measure.[40,41]

We are thus tempted to speculate that there exists a measure μ_k on $U(N)$, such that

$$\int_{U(N)} g(|U_{ij}|^2)d\mu_k(U) = \frac{\Gamma(Nk)}{\Gamma(k)\Gamma((N-1)k)}\int_0^1 g(t)t^{k-1}(1-t)^{(N-1)k-1}dt,$$

with $g, |U_{ij}|^2$ as above. However, we must emphasize that there is an important difference between $\mathcal{U}_3$ and $\mathcal{U}_N$ with $N \geq 4$. There is a single inequality characterizing $\mathcal{U}_3$ inside $\mathcal{B}_3$ [the condition is $Q(b)\geq 0$] while $(N-2)^2$ inequalities occur, in general.[25] Another difference is that the elements of U cannot necessarily be determined from the values $\{|U_{ij}|^2 : 1\leq i,j\leq N\}$ (that is, up to left and right multiplications by diagonal unitary matrices and permutation of rows or columns). For instance, for the flat matrix W_4 of van der Waerden, with all entries equal to 1/4 there exists a one-parameter family of unitary matrices $U(\alpha)$ (rescaled complex Hadamard matrices[37]), such that $|[U(\alpha)]_{ij}|^2 = [W_4]_{ij} = 1/4$.

It would be interesting to be able to fit the "flat" measure on $\mathcal{U}_N$ (inherited from $\mathcal{B}_N$) into the μ_k framework suggested above, but this appears to be a sizable research problem in itself.

[1] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and is Applications* (Academic, New York, 1979).
[2] G. Birkhoff, Rev.-Univ. Nac. Tucuman, Ser. A **A5**, 147 (1946).
[3] C. S. Chan and D. P. Robbins, Exp. Math. **8**, 291 (1999).
[4] M. Beck and D. Pixton, Discrete Comput. Geom. **30**, 623 (2003).
[5] E. R. Canfield and B. D. McKay, e-print arXiv:0705.2422; see also http://www.ojac.org/message.html.
[6] V. Cappellini, H.-J. Sommers, W. Bruzda, and K. Życzkowski, J. Phys. A **42**, 365209 (2009).
[7] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).
[8] C. Jarlskog, Phys. Rev. Lett. **55**, 1039 (1985).
[9] C. Jarlskog and R. Stora, Phys. Lett. B **208**, 268 (1988).
[10] P. Diţă, Mod. Phys. Lett. A **20**, 1709 (2005).
[11] P. Diţă, e-print arXiv:0804.3282.
[12] G. W. Gibbons, S. Gielen, C. N. Pope, and N. Turok, Phys. Rev. Lett. **102**, 121802 (2009).
[13] J. D. Bjorken and I. Dunietz, Phys. Rev. D **36**, 2109 (1987).
[14] L. Lavoura, Phys. Rev. D **40**, 2440 (1989).
[15] G. Auberson, A. Martin, and G. Mennessier, Commun. Math. Phys. **140**, 523 (1991).
[16] Z. Maki, M. Nakagawa, and S. Sakata, Prog. Theor. Phys. **28**, 870 (1962).
[17] M.-C. Chen and K. T. Mahanthappa, Phys. Rev. D **62**, 113007 (2000).
[18] A. Landé, *From Dualism to Unity in Quantum Physics* (Cambridge University Press, Cambridge, 1960).
[19] C. Rovelli, Int. J. Theor. Phys. **35**, 1637 (1996).
[20] J. D. Louck, Found. Phys. **27**, 1085 (1997).
[21] G. Mennessier and J. Nyuts, J. Math. Phys. **15**, 1525 (1974).
[22] G. Tanner, J. Phys. A **34**, 8485 (2001).
[23] P. Pakoński, G. Tanner, and K. Życzkowski, J. Stat. Phys. **111**, 1331 (2003).
[24] R. F. Werner, J. Phys. A **34**, 7081 (2001).
[25] P. Diţă, J. Math. Phys. **47**, 083510 (2006).
[26] K. Życzkowski, M. Kuś, W. Słomczynski, and H.-J. Sommers, J. Phys. A **36**, 3425 (2003).
[27] I. Bengtsson, A. Ericsson, M. Kuś, W. Tadej, and K. Życzkowski, Commun. Math. Phys. **259**, 307 (2005).
[28] Y.-H. Au-Yeung and Y.-T. Poon, Linear Algebr. Appl. **27**, 69 (1979).
[29] H. Nakazato, Nihonkai Mathematical Journal **7**, 83 (1996).
[30] G. W. Gibbons, S. Gielen, C. N. Pope, and N. Turok, Phys. Rev. D **79**, 013009 (2009).
[31] C. F. Dunkl, Trans. Am. Math. Soc. **347**, 3347 (1995).
[32] C. F. Dunkl, Monatsh. Math. **126**, 181 (1998).

[33] R. A. Brualdi and P. M. Gibson, J. Comb. Theory, Ser. A **22**, 194 (1977).
[34] P. Pakoński, K. Życzkowski, and M. Kuś, J. Phys. A **34**, 9303 (2001).
[35] P. Pakoński, Ph.D. thesis, Jagiellonian University, 2002.
[36] U. Haagerup, *Operator Algebras and Quantum Field Theory (Rome)* (International Press, Cambridge, MA, 1996), pp. 296–322.
[37] W. Tadej and K. Życzkowski, Open Syst. Inf. Dyn. **13**, 133 (2006) and an online updated version at http://chaos.if.uj.edu.pl/~karol/hadamard.
[38] W. Słomczyński, Open Syst. Inf. Dyn. **9**, 201 (2002).
[39] H. G. Gadiyar, K. M. S. Maini, R. Padma, and H. S. Sharatchandra, J. Phys. A **36**, L109 (2003).
[40] K. R. W. Jones, J. Phys. A **23**, L1247 (1990).
[41] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States* (Cambridge University Press, Cambridge, 2006).
[42] J. Charles, A. Höcker, H. Lacker, S. Laplace, F. R. Le Diberder, J. Malclés, J. Ocariz, M. Pivk, and L. Roos, Eur. Phys. J. C **41**, 1 (2005); updated results and plots available at http://ckmfitter.in2p3.fr.
[43] Y. Koide and H. Nishiura, Phys. Rev. D **79**, 093005 (2009).
[44] N. N. Lebedev, *Special Functions and Applications* (Dover, New York, 1972).
[45] S. Helgason, *Groups and Geometric Analysis* (Academic, New York, 1984).