Commun. Math. Phys. 259, 307–324 (2005)
Digital Object Identifier (DOI) 10.1007/s00220-005-1392-8

Communications in
**Mathematical
Physics**

# Birkhoff's Polytope and Unistochastic Matrices, N = 3 and N = 4

**Ingemar Bengtsson**[1], **Åsa Ericsson**[1], **Marek Kuś**[2], **Wojciech Tadej**[3], **Karol Życzkowski**[2,4]

[1] Stockholm University, AlbaNova, Fysikum, 106 91 Stockholm, Sweden.
E-mail: ingemar@physto.se; asae@physto.se
[2] Centrum Fizyki Teoretycznej, Polska Akademia Nauk, Al. Lotników 32/44, 02–668 Warszawa, Poland.
E-mail: marek@cft.edu.pl
[3] Cardinal Stefan Wyszynski University, Warszawa, Poland. E-mail: wtadej@wp.pl
[4] Instytut Fizyki im. Smoluchowskiego, Uniwersytet Jagielloński, ul. Reymonta 4, 30–059 Kraków, Poland. E-mail: karol@tatry.if.uj.edu.pl

Received: 8 March 2004 / Accepted: 24 March 2005
Published online: 15 July 2005 – 

**Abstract:** The set of bistochastic or doubly stochastic $N \times N$ matrices is a convex set called Birkhoff's polytope, which we describe in some detail. Our problem is to characterize the set of unistochastic matrices as a subset of Birkhoff's polytope. For $N = 3$ we present fairly complete results. For $N = 4$ partial results are obtained. An interesting difference between the two cases is that there is a ball of unistochastic matrices around the van der Waerden matrix for $N = 3$, while this is not the case for $N = 4$.

## 1. Introduction

Unistochastic matrices arise in many different contexts including error correcting codes, quantum information theory and particle physics. To define them, we first recall that an $N \times N$ matrix $B$ is said to be *bistochastic* if its matrix elements satisfy

$$\text{i: } B_{ij} \geq 0, \qquad \text{ii: } \sum_i B_{ij} = 1, \qquad \text{iii: } \sum_j B_{ij} = 1 \, . \tag{1}$$

The set of bistochastic matrices is a convex polytope known as *Birkhoff's polytope*.

One way of constructing a bistochastic matrix is to begin with a unitary matrix $U$ and let

$$B_{ij} = |U_{ij}|^2 \, . \tag{2}$$

However, it is well-known [1] that not all bistochastic matrices arise in this way. If there is such a $U$, then we will call $B$ *unistochastic*. If $U$ is also real, that is orthogonal, then we call $B$ *orthostochastic*. (Much of the mathematics literature uses the term *orthostochastic* to mean any matrix satisfying (2) and does not distinguish the subclass for which $U$ is real. We will see later that the distinction is important.) In this paper, we consider the problem of characterizing the unistochastic subset of Birkhoff's polytope.

Before summarizing our results, we mention some physical applications. In quantum mechanics, the transition probabilities associated with a finite basis form bistochastic matrices. In studies of the foundations of quantum theory, the attempt to build some group structure into these transition probabilities leads to the requirement that they form unistochastic matrices. A sample of the literature includes Landé [2], Rovelli [3] and Khrennikov [4].

In the attempt to formulate quantum mechanics on graphs (in the laboratory on thin strips of, say, gold film) the question of what Markov processes have quantum counterparts in the given setting again leads to unistochastic matrices [5–7]. In this connection studies of the spectra and entropies of unistochastic matrices chosen at random have been made [8].

In particle physics, a related question arises. In the theory of weak interactions one encounters the unitary Kobayashi-Maskawa matrices (one for quarks and one for neutrinos), and Jarlskog raised the question to what extent such a matrix can be parametrized by the easily measured moduli of its matrix elements. The physically interesting case here is $N = 3$ [9], and possibly also $N = 4$, should a fourth generation of quarks be discovered [10]. The question of determining $U$ from $B$ also arises in scattering theory, with no restriction on $N$ [11].

Our main result involves the *van der Waerden matrix* $J_N$, whose matrix elements satisfy $(J_N)_{ij} = \frac{1}{N}$. This matrix is unistochastic, and any corresponding unitary matrix is known as a *complex Hadamard matrix*. An example is the *Fourier matrix*, whose matrix elements are

$$U_{jk} = \frac{1}{\sqrt{N}} q^{jk} \ , \qquad 0 \leq j, k \leq N - 1 \ . \tag{3}$$

Here $q = e^{2\pi i/N}$ is a root of unity. Complex Hadamard matrices have a long history in mathematics [12–14], and have recently arisen in quantum information theory [15–17].

In this paper we study the set of unistochastic matrices, and the precise way in which it forms a subset of Birkhoff's polytope. Our main result is that for $N = 4$ every neighborhood of the van der Waerden matrix contains matrices that are not unistochastic. This is in striking contrast with the $N = 3$ case for which $J_3$ is at the center of a ball of unistochastic matrices inside a star-shaped region bounded by the set of orthostochastic matrices.

This paper is organized as follows. In Sect. 2 we consider the set of all bistochastic matrices, and describe the cases $N = 3$ and $N = 4$ in detail ($N = 2$ is trivial). In Sect. 3 we discuss some generalities concerning unistochastic matrices, and then characterize the unistochastic subset in the case $N = 3$. Most of our results can be found elsewhere but, we believe, not in this coherent form. In Sect. 4 we consider $N = 4$, prove our main result, and relate some already known facts [10] to our explicit description of Birkhoff's polytope. Section 5 summarises our conclusions. Some technical matters are found in three appendices.

## 2. Birkhoff's Polytope

The set $\mathcal{B}_N$ of bistochastic $N \times N$ matrices has $(N - 1)^2$ dimensions. To see this, note that the last row and the last column are fixed by the conditions that the row and column sums should equal one. The remaining $(N - 1)^2$ entries can be chosen freely, within limits. Birkhoff proved that $\mathcal{B}_N$ is a convex polytope whose extreme points, or corners, are the $N!$ permutation matrices [18]. It is called Birkhoff's polytope. All its corners are equivalent in the sense that they can be transformed into each other by means of orthogonal transformations. A bistochastic matrix belongs to the boundary of $\mathcal{B}_N$ if and only if at least one of its entries is zero. The boundary consists of corners, edges, faces, 3-faces and so on; the highest dimensional faces are called facets and consist of matrices with only one zero entry. For a detailed account of $\mathcal{B}_N$, especially its face structure, see Brualdi et al. [19]. We will be even more detailed concerning $\mathcal{B}_3$ and $\mathcal{B}_4$. We will use a quite explicit notation for the 24 permutation matrices in $\mathcal{B}_4$; see Appendix A for the details.

It is convenient to regard the convex polytope $\mathcal{B}_N$ as a subset of a vector space, with the van der Waerden matrix $J_N$ as its origin. The distance squared between two matrices is chosen to be

$$D^2(A, B) = \text{Tr}(A - B)(A^\dagger - B^\dagger) \ , \tag{4}$$

where the dagger denotes Hermitian conjugation. The distance squared between an arbitrary bistochastic matrix $B$ and the van der Waerden matrix $J_N$ is then given by

$$D^2(B, J_N) = \sum_{i,j} B_{ij}^2 - 1 \ . \tag{5}$$

In particular, the distance between $J_N$ and a corner of the polytope becomes $D = \sqrt{N-1}$. Permutations of rows or columns are orthogonal transformations of the polytope, since they preserve distance and leave the van der Waerden matrix invariant. They also take permutation matrices (corners) into permutation matrices, hence they are symmetry operations of Birkhoff's polytope as well.

The (Shannon) *entropy* of a bistochastic matrix is defined as the entropy of the rows averaged over the columns,

$$S = -\frac{1}{N}\sum_i \sum_j B_{ij} \ln B_{ij} \ . \tag{6}$$

Its maximum value $\ln N$ is attained at $J_N$. For some of its properties consult Słomczyński [20] et al. [8].

When $N = 2$ there are just two permutation matrices and $\mathcal{B}_2$ is a line segment between these two points. A general bistochastic matrix can be parametrized as

$$B = \begin{bmatrix} c^2 & s^2 \\ s^2 & c^2 \end{bmatrix} , \qquad c \equiv \cos\theta \ , \qquad s \equiv \sin\theta \ , \qquad 0 \le \theta \le \frac{\pi}{2} \ . \tag{7}$$

When $N = 3$ we have six permutation matrices forming the vertices of a four dimensional polytope. It admits a simple description:

**Theorem 1.** *The 6 corners of $\mathcal{B}_3$ are the corners of two equilateral triangles placed in two totally orthogonal 2-planes and centered at $J_3$.*

To prove this we form two triangles as convex combinations of permutation matrices. Using a notation that is consistent with Appendix A they are

$$\Delta_1 = p_0 P_0 + p_3 P_3 + p_4 P_4 = \begin{bmatrix} p_0 & p_3 & p_4 \\ p_4 & p_0 & p_3 \\ p_3 & p_4 & p_0 \end{bmatrix} , \qquad p_0 + p_3 + p_4 = 1 \tag{8}$$

and

$$\Delta_2 = p_1 P_1 + p_2 P_2 + p_5 P_5 = \begin{bmatrix} p_1 & p_2 & p_5 \\ p_2 & p_5 & p_1 \\ p_5 & p_1 & p_2 \end{bmatrix} , \qquad p_1 + p_2 + p_5 = 1 \ . \tag{9}$$

The calculation we have to do is to check that $D^2(P_0, P_3) = D^2(P_0, P_4) = D^2(P_3, P_4) = 6$ and similarly for the other triangle, and also that

$$\mathrm{Tr}(\Delta_1 - J_3)(\Delta_2^\dagger - J_3) = 0 \tag{10}$$

for all values of $p_i$. This is so.

There are thus 6 corners and $6 \cdot 5/2 = 15$ edges, all of which are extremal. The last is a rather exceptional property; in 3 dimensions only the simplex has it. There are 9 short edges of length squared $D^2 = 4$ and 6 long edges of length squared $D^2 = 6$, namely the sides of the two equilateral triangles. A useful overview of $\mathcal{B}_3$ is given by its graph, where we exhibit all corners and all edges (see Fig. 1). All the 2-faces are triangles with one long and two short edges. The 3-faces in a 4 dimensional polytope are facets and here they are made of matrices with a single zero. They are irregular tetrahedra with two long edges, one from each equilateral triangle (see Fig. 4).

The volume of $\mathcal{B}_3$ is readily computed because it can be triangulated using only three simplices. The total volume is 9/8. As $N$ grows the total volume of $\mathcal{B}_N$ becomes increasingly hard to compute; mathematicians know it for $N \leq 10$ [21].

The next case is the 9 dimensional polytope $\mathcal{B}_4$. It has 24 corners and 276 edges. The latter come in four types and we give the classification including the angle they subtend at $J_4$ and whether they consist of unistochastic matrices or not (see Sects. 3 and 4):

| | Length squared | Unistochastic | Angle at origin | Number of edges |
|---|---|---|---|---|
| $4U$ | 4 | Yes | Acute | 72 |
| 6 | 6 | No | 90 degrees | 96 |
| 8 | 8 | No | Obtuse | 72 |
| $8U$ | 8 | Yes | Obtuse | 36 |

All edges except the $8U$ ones are extremal. The 2-faces consist of triangles and squares. (Interestingly, for all $N$ it is true that the 2-faces of Birkhoff's polytope $\mathcal{B}_N$ are either triangles or rectangles [19].) There are 18 squares bounded by edges of type $4U$ and their diagonals are of type $8U$. Three squares meet at each corner. If we pick four permutation matrices we obtain a 3-face, with six exceptions. The exceptions form 6 regular tetrahedra centered at $J_4$, whose edges are non-extremal $8U$ edges. They are denoted $T_i$

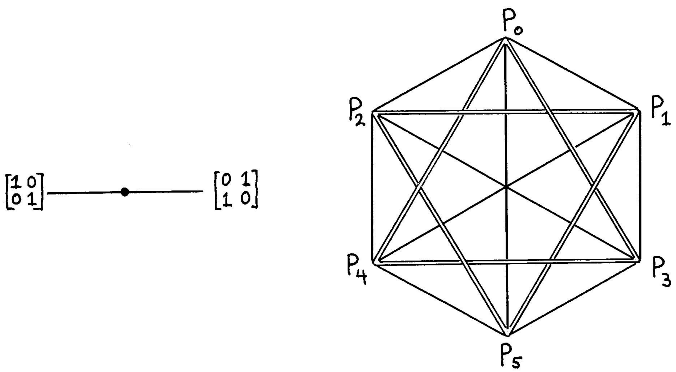


**Fig. 1.** Left: Birkhoff's polytope for $N = 2$ (centered at $J_2$). Right: The graph of Birkhoff's polytope for $N = 3$; single lines have $D^2 = 4$ and double $D^2 = 6$. The double edges form the triangles mentioned in Theorem 1

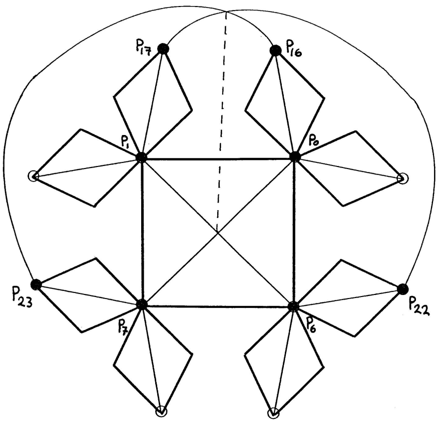


**Fig. 2.** How to begin to draw the surface of $\mathcal{B}_4$. Two tetrahedra whose edges are the non-extremal diagonals of squares are shown. The dashed line goes through the polytope; it connects the midpoints of two opposing $8U$ edges of two tetrahedra that are otherwise disjoint

and explicitly listed in Appendix A. When regular tetrahedra are mentioned below it is understood that we refer to one of these six. In a sense the structure can now be drawn; see Fig. 2. The facets consist of matrices with one zero, so there are 16 facets.

A subset of $\mathcal{B}_4$ that has no counterpart for $\mathcal{B}_3$ is the set of matrices that are tensor products of two by two bistochastic matrices. This subset splits naturally into several two dimensional components, and it turns out that they sit in $\mathcal{B}_4$ as doubly ruled surfaces inside the regular tetrahedra. Thus the following matrix, parametrised with two angles, is a tensor product of two matrices of the form (7):

$$\begin{bmatrix} c_1^2c_2^2 & c_1^2s_2^2 & s_1^2c_2^2 & s_1^2s_2^2 \\ c_1^2s_2^2 & c_1^2c_2^2 & s_1^2s_2^2 & s_1^2c_2^2 \\ s_1^2c_2^2 & s_1^2s_2^2 & c_1^2c_2^2 & c_1^2s_2^2 \\ s_1^2s_2^2 & s_1^2c_2^2 & c_1^2s_2^2 & c_1^2c_2^2 \end{bmatrix}, \qquad c_1 \equiv \cos\theta_1, \text{ etc.} \tag{11}$$

These matrices form a doubly ruled surface inside the regular tetrahedron $T_1$, analogous to that depicted in Fig. 4.

An interesting way to view $\mathcal{B}_4$, and one that will recur in Sect. 4, stems from the following observation:

**Theorem 2.** *The 24 corners of $\mathcal{B}_4$ belong to a set of nine orthogonal hyperplanes through $J_4$. Each regular tetrahedron belongs to six hyperplanes and contains the normal vectors of the remaining three hyperplanes. Each hyperplane contains four regular tetrahedra and its normal vector is the intersection of the remaining two regular tetrahedra.*

Again the proof is a simple calculation, once the explicit form of the hyperplanes is known. They are denoted $\Pi_i$ and listed in Appendix A. From now on, *hyperplane* always refers to one of these nine. Figure 3 in a sense illustrates the theorem.

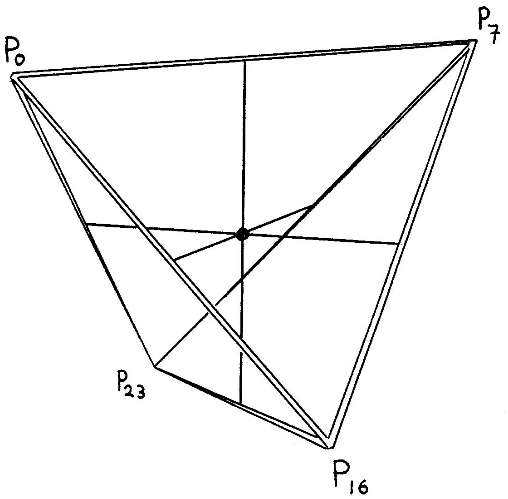


**Fig. 3.** A regular tetrahedron centered at $J_4$. It contains the normal vectors of three orthogonal hyperplanes and belongs entirely to another six. There are six such regular tetrahedra and pairs of them intersect along the normal vectors they contain. (Note that the dashed line in Fig. 2 represents such a normal vector.)

It is quite helpful to have an incidence table for tetrahedra and hyperplanes available. It is

$$\begin{array}{cccccccccc} & \Pi_1 & \Pi_2 & \Pi_3 & \Pi_4 & \Pi_5 & \Pi_6 & \Pi_7 & \Pi_8 & \Pi_9 \\ T_1 & & X & X & X & & X & X & X & \\ T_2 & & X & X & X & X & & X & & X \\ T_3 & X & & X & & X & X & X & X & \\ T_4 & X & X & & & X & X & X & & X \\ T_5 & X & & X & X & X & & & X & X \\ T_6 & X & X & & X & & X & & X & X \end{array} \tag{12}$$

where the tetrahedra $T_i$ and the hyperplanes $\Pi_i$ are listed in Appendix A.

For later purposes we will need some information about exactly how the hyperplanes divide the space into $2^9$ hyperoctants. For this reason we look at the rays

$$B_i(t) = J_4 + tV_i \ , \tag{13}$$

where $V_i$ is a vector constructed in terms of the normal vectors $n_1, \ldots, n_9$ of the hyperplanes (see Appendix A), namely

$$V_1 \equiv n_1 + n_2 + n_3 + n_4 + n_5 + n_6 + n_7 + n_8 + n_9 = \frac{1}{4} \begin{bmatrix} 9 & -3 & -3 & -3 \\ -3 & 1 & 1 & 1 \\ -3 & 1 & 1 & 1 \\ -3 & 1 & 1 & 1 \end{bmatrix} , \tag{14}$$

$$V_2 \equiv n_1 + n_2 + n_3 + n_4 + n_5 + n_6 + n_7 + n_8 - n_9 = \frac{1}{4} \begin{bmatrix} 7 & -1 & -1 & -5 \\ -1 & -1 & -1 & 3 \\ -1 & -1 & -1 & 3 \\ -5 & 3 & 3 & -1 \end{bmatrix} , \tag{15}$$

$$V_3 \equiv n_1 + n_2 + n_3 + n_4 - n_5 + n_6 + n_7 + n_8 - n_9 = \frac{1}{4}\begin{bmatrix} 5 & 1 & -3 & -3 \\ 1 & -3 & 1 & 1 \\ -3 & 1 & -3 & 5 \\ -3 & 1 & 5 & -3 \end{bmatrix}. \quad (16)$$

All other cases can be obtained from one of these three by permutations of rows and columns. The various hyperoctants are convex cones centered on these rays. This gives a classification of the hyperoctants into six different types (since the parameter $t$ can be positive or negative) called respectively type $\mathrm{I}_\pm$, $\mathrm{II}_\pm$ and $\mathrm{III}_\pm$. Type I has 16 representatives and is especially noteworthy. For type $\mathrm{I}_-$ the centered ray hits the boundary in the center of one of the 16 facets, at the matrix $B_1(-\frac{1}{9})$. In the other direction we also hit a quite distinguished point. There are 16 ways of setting one entry of a bistochastic matrix equal to one, and this gives rise to 16 copies of $\mathcal{B}_3$ sitting in the boundary of $\mathcal{B}_4$. For the octants $\mathrm{I}_+$ the centered ray hits the boundary precisely at the center of such a $\mathcal{B}_3$, at the matrix $B_1(\frac{1}{3})$.

In Sect. 4 we will see how the structure of the unistochastic subset is related to the structure of Birkhoff's polytope, and in particular to those of its features that we have stressed.

## 3. The Unistochastic Subset, $N$ = 3

Let us begin with some generalities concerning the unistochastic subset $\mathcal{U}_N$ of $\mathcal{B}_N$. The dimension of $\mathcal{B}_N$ is $(N-1)^2$, and the dimension of the group of unitary $N$ by $N$ matrices, $U(N)$, is $N^2$. Therefore the map $U(N) \to \mathcal{B}_N$ cannot be one-to-one. Now it is clear that multiplying a row or a column by a phase factor—an operation that we refer to as *rephasing*—will result in the same bistochastic matrix via Eq. (2). Therefore the map is naturally defined as a map from a double coset space to $\mathcal{B}_N$. The double coset space is

$$U(1) \times \cdots \times U(1) \setminus U(N) \,/\, U(1) \times \cdots \times U(1)\,, \quad (17)$$

with $N$ $U(1)$ factors acting from the right and $N-1$ factors from the left, say. The dimension of this set is $(N-1)^2$, so now the dimensions match. There is a complication because the double coset space is not a smooth manifold. The action from the left of the $U(1)$ factors on the right coset space (in itself a well behaved flag manifold) has fixed points. These fixed points are easy to locate however (and always map to the boundary of $\mathcal{B}_N$), so that for most practical purposes we can think of our map as a map between smooth manifolds.

In general we will see that the image of our map is a proper subset of $\mathcal{B}_N$, and the map is many-to-one. There is not much we can usefully say about the general case, except for two remarks: The unistochastic subset $\mathcal{U}_N$ has the full dimension $(N-1)^2$ while the unistochastic subset of the boundary of $\mathcal{B}_N$ has dimension $(N-1)^2-2$; why this is so will presently become clear.

For $N=2$ every bistochastic matrix is orthostochastic. A unitary matrix that maps to the matrix in Eq. (7) is

$$U = \begin{bmatrix} c & s \\ s & -c \end{bmatrix}, \qquad c \equiv \cos\theta\,, \qquad s \equiv \sin\theta\,, \qquad 0 \le \theta \le \frac{\pi}{2}\,. \quad (18)$$

The matrix is given in *dephased* form. This means that the first row and the first column are real and positive. This fixes the $U(1)$ factors mentioned above (unless there is a zero entry in one of these places) and from now on we shall present all unitary matrices in this form. For any $N$ it is straightforward to check whether a given edge of $\mathcal{B}_N$ is unistochastic. For $N = 3$ the edges of length squared equal to 4 are unistochastic, and for $N = 4$ we have the results given in table (11).

Given a $3 \times 3$ bistochastic matrix it is easy to check whether it is unistochastic or not [22, 9]. We form the moduli $r_{ij} = \sqrt{B_{ij}}$ and write down the matrix

$$U = \begin{bmatrix} r_{00} & r_{01} & \bullet \\ r_{10} & r_{11}e^{i\phi_{11}} & \bullet \\ r_{20} & r_{21}e^{i\phi_{21}} & \bullet \end{bmatrix} . \tag{19}$$

This matrix is given in dephased form. If it is unitary, the original matrix $B$ is unistochastic. The unitarity conditions simply say that the first two columns are orthogonal. The last column by construction has the right moduli and does not impose any further restrictions, hence it is not written explicitly. The problem is whether phases $\phi_{11}$ and $\phi_{21}$ can be found so that the matrix is unitary. This problem can be translated into the problem of forming a triangle from three line segments of given lengths

$$L_0 = r_{00}r_{01}, \qquad L_1 = r_{10}r_{11}, \qquad L_2 = r_{20}r_{21}. \tag{20}$$

This is possible if and only if the "chain–links" conditions are fulfilled, i.e.

$$|L_1 - L_2| \le L_0 \le L_1 + L_2 . \tag{21}$$

The bistochastic matrix $B$ corresponding to $U$ sits at the boundary of $\mathcal{U}_3$ if and only if one of these inequalities is saturated. When the inequalities (21) hold the solution for the phases is

$$\cos\phi_{11} = \frac{L_2^2 - L_0^2 - L_1^2}{2L_0L_1}, \qquad \cos\phi_{21} = \frac{L_1^2 - L_2^2 - L_0^2}{2L_0L_2}, \tag{22}$$

$$\cos(\phi_{11} - \phi_{21}) = \frac{L_0^2 - L_1^2 - L_2^2}{2L_1L_2} . \tag{23}$$

There is a two-fold ambiguity (corresponding to taking the complex conjugate of the matrix, $U \to U^*$). The area $A$ of the triangle is easily computed and the chain–links conditions are equivalent to the single inequality $A \ge 0$. As a matter of fact we can form six so called unitarity triangles in this way, depending on what pair of columns or rows that we choose. Although their shapes differ their area is the same, by unitarity [9].

Because we can easily decide if a given matrix is unistochastic, it is easy to characterize the unistochastic set $\mathcal{U}_3$. We single out the following facts (some of which are known [22, 23]) for attention:

**Theorem 3.** *The unistochastic subset $\mathcal{U}_3$ of $\mathcal{B}_3$ is a non-convex star shaped four dimensional set whose boundary consists of the set of orthostochastic matrices. It contains a unistochastic ball of maximal radius $\sqrt{2}/3$, centered at $J_3$. The set meets the boundary of $\mathcal{B}_3$ in a doubly ruled surface in each facet.*

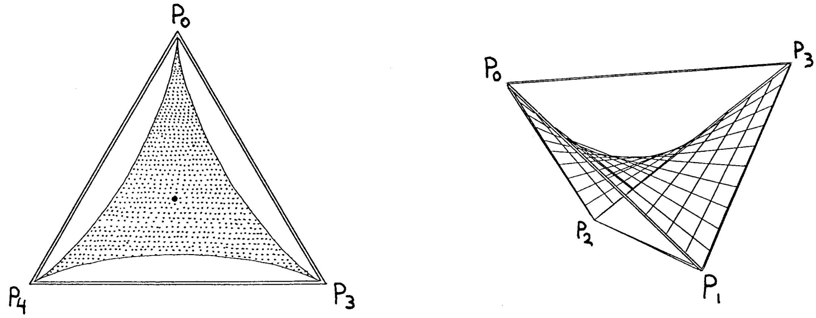


**Fig. 4.** Birkhoff's polytope for $N = 3$. Left: One of the two orthogonal equilateral triangles centered at $J_3$, with its unistochastic subset (the boundary is the famous hypocycloid). Right: A facet, an irregular tetrahedron, with its doubly ruled surface of unistochastic matrices

The relative volume of the unistochastic subset is, according to our numerics,

$$\frac{\mathrm{vol}(\mathcal{U}_3)}{\mathrm{vol}(\mathcal{B}_3)} \approx 0.7520 \pm 0.0005 \ . \tag{24}$$

We did not attempt an analytical calculation; details of our numerics are in Appendix B.

Theorem 3 is easy to prove. To see that $\mathcal{U}_3$ is non-convex we just draw its intersection with one of the equilateral triangles that went into the definition of the polytope, and look at it (see Fig. 4). An amusing side remark is that the boundary of the unistochastic set in this picture is a 3-hypocycloid [8]. It can be obtained by rolling a circle of radius $1/3$ inside the unit circle. The maximal unistochastic ball is centered at $J_3$ and touches the boundary at the hypocycloid, as one might guess from the picture; its radius was deduced from results presented in ref. [24]. To see that the boundary consists of orthostochastic matrices, observe that when the chain–links conditions are saturated the phases in $U$ will equal $\pm 1$. That the set is star shaped then follows from an explicit check that there is only one orthostochastic matrix on any ray from $J_3$. Finally Fig. 4 includes a picture of the unistochastic subset of a facet. The reason why it has codimension one is that a matrix on the boundary of $\mathcal{B}_N$ has a zero entry, which means that the number of phases available in the dephased unitary matrix drops with one, and then the dimension of the unistochastic set also drops with one; the argument goes through for any $N$.

Finally let us make some remarks on entropy. First we compare the Shannon entropy averaged over $\mathcal{B}_3$ to the Shannon entropy averaged over $\mathcal{U}_3$, using the flat measure in both cases. Numerically we find that

$$\langle S \rangle_{\mathcal{B}_3} \approx 0.883 \quad \text{and} \quad \langle S \rangle_{\mathcal{U}_3} \approx 0.908 \ , \tag{25}$$

with all digits significant. Observe that the latter average is larger since some matrices of small entropy close to the boundary of $\mathcal{B}_3$ are not unistochastic and do not contribute to the average over $\mathcal{U}_3$. The above data may be compared with the maximal possible entropy $S_{\mathrm{max}} = \ln 3 \approx 1.099$, attained at $J_3$, and also with

$$\langle S \rangle_{\mathrm{Haar}} = \frac{1}{2} + \frac{1}{3} \approx 0.833 \ , \tag{26}$$

which is the average taken over $\mathcal{U}_3$ with respect to the measure induced by the Haar measure on $U(3)$. This analytical result follows from the observation that $\langle S \rangle_{\mathrm{Haar}}$ coincides with the average entropy of squared components of complex random vectors, which was computed by Jones [25].

## 4. The Unistochastic Subset, N = 4

The case $N = 4$ is more difficult. It is also clear from the outset that it will be qualitatively different—thus the dimension of the orthogonal group is too small for the boundary of the unistochastic set $\mathcal{U}_4$ to be formed by orthostochastic matrices alone. There are other differences too, as we will see.

Given a bistochastic matrix we can again define $r_{ij} = \sqrt{B_{ij}}$ and consider

$$U = \begin{bmatrix} r_{00} & r_{01} & r_{02} & \bullet \\ r_{10} & r_{11}e^{i\phi_{11}} & r_{12}e^{i\phi_{12}} & \bullet \\ r_{20} & r_{21}e^{i\phi_{21}} & r_{22}e^{i\phi_{22}} & \bullet \\ r_{30} & r_{31}e^{i\phi_{31}} & r_{32}e^{i\phi_{32}} & \bullet \end{bmatrix}. \tag{27}$$

Phases must now be chosen so that this matrix is unitary, and more especially so that the three columns we focus on are orthogonal. Geometrically this is the problem of forming three quadrilaterals with their sides given and six free angles. This is not a simple problem, and in practice we have to resort to numerics to see whether a given bistochastic matrix is unistochastic (see Appendix B for details). There are some easy special cases though. One easy case is that of a matrix belonging to the boundary of $\mathcal{B}_N$. Then the matrix $U$ must contain one zero entry and when we check the orthogonality of our three columns two of the equations reduce to the problem of forming triangles. This fixes four of the angles, and the final orthogonality relation is easily dealt with. Another easy case concerns the regular tetrahedra. They turn out to consist of orthostochastic matrices; for the tetrahedron $T_1$ (see Appendix A) a corresponding orthonormal matrix is

$$O_1 = \begin{bmatrix} \sqrt{p_0} & \sqrt{p_7} & \sqrt{p_{16}} & \sqrt{p_{23}} \\ \sqrt{p_7} & -\sqrt{p_0} & -\sqrt{p_{23}} & \sqrt{p_{16}} \\ \sqrt{p_{16}} & \sqrt{p_{23}} & -\sqrt{p_0} & -\sqrt{p_7} \\ \sqrt{p_{23}} & -\sqrt{p_{16}} & \sqrt{p_7} & -\sqrt{p_0} \end{bmatrix}. \tag{28}$$

This saturates a bound saying that the maximum number of $N \times N$ permutation matrices whose convex hull is unistochastic is not larger than $2^{\left[\frac{N}{2}\right]}$, where $[N/2]$ denotes the integer part of $N/2$ [26].

Let us now turn our attention to $\mathcal{J}_4$. Hadamard [27] observed that up to permutations of rows and columns the most general form of the complex Hadamard matrix is

$$H(\phi) = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & e^{i\phi} & -1 & -e^{i\phi} \\ 1 & -1 & 1 & -1 \\ 1 & -e^{i\phi} & -1 & e^{i\phi} \end{bmatrix}. \tag{29}$$

One can show that this is a geodesic in $U(N)$. What is new, compared to $N = 3$, is that the van der Waerden matrix is orthostochastic because $H(0)$ is real. Moreover, there is a continuous set of dephased unitaries mapping to the same $B$. In a calculational *tour de force*, Auberson et al. [10] were able to determine all bistochastic matrices whose dephased unitary preimages contain a continuous ambiguity (and they found that the ambiguity is given by one parameter in all cases). There are three such families. Using the notation of ref. [10] they consist of matrices of the following form:

$$\text{Type A:}\quad \begin{bmatrix} a & b & c & d \\ b & a & d & c \\ e & f & g & h \\ f & e & h & g \end{bmatrix}, \qquad \text{Type C:}\quad \begin{bmatrix} a & a & \frac{1}{2}-a & \frac{1}{2}-a \\ b & b & \frac{1}{2}-b & \frac{1}{2}-b \\ c & c & \frac{1}{2}-c & \frac{1}{2}-c \\ d & d & \frac{1}{2}-d & \frac{1}{2}-d \end{bmatrix}, \tag{30}$$

$$\text{Type B:}\quad \begin{bmatrix} s_1^2 s_2^2 & c_1^2 s_2^2 & c_3^2 c_2^2 & s_3^2 c_2^2 \\ s_1^2 c_2^2 & c_1^2 c_2^2 & c_3^2 s_2^2 & s_3^2 s_2^2 \\ c_1^2 c_4^2 & s_1^2 c_4^2 & s_3^2 s_4^2 & c_3^2 s_4^2 \\ c_1^2 s_4^2 & s_1^2 s_4^2 & s_3^2 c_4^2 & c_3^2 c_4^2 \end{bmatrix}. \tag{31}$$

Here $c_1 = \cos\theta_1$, $s_1 \equiv \sin\theta_1$, and so on. Type A consists of nine five dimensional sets, type B of nine four dimensional sets, and type C of six three dimensional sets. In trying to understand their location in $\mathcal{B}_4$ the observation in Sect. 2 concerning the nine orthogonal hyperplanes begins to pay dividends. (In particular, consult the incidence table (12).) Type A consists of the linear subspaces obtained by taking all intersections of four hyperplanes that contain exactly two regular tetrahedra. Type C consists of the linear subspaces obtained by taking all intersections of six hyperplanes that contain no permutation matrices at all. Type B finally consists of curved manifolds confined to one hyperplane. Auberson's families are not exclusive. In particular tensor product matrices belong to families A and B, which means that there are two genuinely different ways of introducing a free phase in the corresponding unitary matrix. Outside the three sets A, B and C Auberson et al. find a 12-fold discrete ambiguity in the dephased unitaries, dropping to 4-fold for symmetric matrices [10].

Tensor product matrices $B_4 = B_2 \otimes B_2'$ appear because $4 = 2 \times 2$ is a composite number. That they are always unistochastic follows from a more general result:

**Lemma 1.** *Let $B_K$ and $B_M$ be unistochastic matrices of size $K$ and $M$, respectively. Then the matrix $B_N = B_K \otimes B_M$ of size $KM$ is unistochastic. The corresponding dephased unitary matrices contain at least $(K-1)(M-1)$ free phases.*

That $B_N$ is unistochastic follows from properties of the Hadamard and the tensor products. By definition, the Hadamard product $A \circ B$ of two matrices is the matrix whose matrix elements are the products of the corresponding matrix elements of $A$ and $B$. Then $B_K = U_K \circ U_K^*$ and $B_M = U_M \circ U_M^*$ implies that $B_N = (U_K \circ U_K^*) \otimes (U_M \circ U_M^*) = (U_K \otimes U_M) \circ (U_K^* \otimes U_M^*)$, so it is unistochastic. The existence of free phases is an easy generalization of Proposition 2.9 in Haagerup [28].

The hyperplane structure of $\mathcal{B}_4$ reverberates in the structure of the unistochastic set in several ways. Let us consider how the tangent space of $U(N)$ behaves under the map to $\mathcal{B}_N$. In equations, this means that we fix a unitary matrix $U_0$ and expand

$$U(t) = e^{iht} U_0 = (1 + iht - \frac{1}{2} h^2 t^2 + \ldots) U_0, \tag{32}$$

where $h$ is an Hermitian matrix. Then we study bistochastic matrices with elements $B_{ij}(t) = |U_{ij}(t)|^2$ to first order in $t$. The following features are true for all $N$:

- Generically the tangent space of $U(N)$ maps onto the tangent space of $\mathcal{B}_N$. We checked this statement by generating unitary matrices at random using the Haar measure on the group. It implies that the dimension of the unistochastic set is equal to that of $\mathcal{B}_N$.
- A matrix element in $B$ receives a first order contribution only if it is non-vanishing. Hence the map of the tangent space of $U(N)$ to the tangent space of $\mathcal{B}_N$ is degenerate at the boundary of the polytope. In general such behaviour is to be expected at the boundary of the unistochastic set $\mathcal{U}_N$.
- If $U_0$ is real the map is degenerate in the sense that the tangent space maps to an $N(N-1)/2$ dimensional subspace of the tangent space of $\mathcal{B}_N$.
- If $U_0$ maps to a corner of the polytope then the first order contributions vanish. To second order we pick up the tip of a convex cone whose extreme rays are the $N(N-1)/2$ edges of type $4U$, emanating from that corner.

For $N = 4$ the story becomes interesting when we choose $U_0$ equal to the Hadamard matrix $H(\phi)$. Then we find that the tangent space at $U_0$ maps into one of the nine hyperplanes; which particular one depends on how we permute rows and columns in Eq. (29). The question therefore arises whether the orthostochastic van der Waerden matrix belongs to the boundary of the unistochastic set—or not since *a priori* such degeneracies can occur also in the interior of the set.

We know that we can form curves of unistochastic matrices starting from $J_4$ and moving out into the nine hyperplanes. Can we form such curves that go directly out into one of the $2^9$ hyperoctants? Here the division of the $2^9$ hyperoctants into six different types becomes relevant. We have investigated whether their central rays given in Eqs. (13–16) consist of unistochastic matrices, or not. Let us begin with the 16 hyperoctants of type I, where the central ray $B_1(t) = J_4 + tV_1$ hits the boundary in the center of one of the 16 $\mathcal{B}_3$ sitting in the boundary (at $t = 1/3$), and in the center of one of the 16 facets (at $t = -1/9$). Of these two points, the first is unistochastic, the second is not. A one parameter family of candidate unitary matrices that maps to the central ray is

$$U(t) = \frac{1}{2}\begin{bmatrix} \sqrt{1-3t} & \sqrt{1-3t} & \sqrt{1-3t} & \bullet \\ \sqrt{1+t} & \sqrt{1+t}e^{\phi_{11}} & \sqrt{1+t}e^{i\phi_{12}} & \bullet \\ \sqrt{1+t} & \sqrt{1+t}e^{i\phi_{21}} & \sqrt{1+t}e^{\phi_{22}} & \bullet \\ \sqrt{1+t} & \sqrt{1+t}e^{\phi_{31}} & \sqrt{1+t}e^{i\phi_{32}} & \bullet \end{bmatrix}, \tag{33}$$

where $t > 0$ and we permuted the columns relative to Eq. (14) in order to get the unitarity equations in a pleasant form. (We do not need to give the phases for the last column.) The conditions that the first three columns be orthogonal read

$$e^{i\phi_{11}} + e^{i\phi_{21}} + e^{i\phi_{31}} + L = 0, \tag{34}$$

$$e^{i\phi_{12}} + e^{i\phi_{22}} + e^{i\phi_{32}} + L = 0, \tag{35}$$

$$e^{i(\phi_{11}-\phi_{12})} + e^{i(\phi_{21}-\phi_{22})} + e^{i(\phi_{31}-\phi_{32})} + L = 0, \tag{36}$$

where

$$L = \frac{1-3t}{1+t}. \tag{37}$$

In Appendix C we prove that the system of equations (34–36)

1. has no real solutions for $L > 1$,

2. for $0 < L < 1$ has the solution

$$\begin{array}{l} \phi_{11} = 0,\ \phi_{21} = \phi,\ \phi_{31} = -\phi \\ \phi_{12} = \phi,\ \phi_{22} = 0,\ \phi_{32} = -\phi \end{array}, \qquad \cos\phi = \frac{t-1}{t+1} = -\frac{L+1}{2}\,. \tag{38}$$

It follows that the central ray is unistochastic for the hyperoctants of type $\mathrm{I}_+$ (and the unitary matrices on the central ray tend to the real Hadamard matrix at $t = 0$). In the other direction the central ray is not unistochastic for type $\mathrm{I}_-$. Thus we have proved

**Theorem 4.** *For $N = 4$ there are non-unistochastic matrices in every neighbourhood of the van der Waerden matrix $J_4$. At $J_4$ the map $U(4) \to \mathcal{B}_4$ aligns the tangent space of $U(4)$ with one of the nine orthogonal hyperplanes.*

The structure of the unistochastic set is dramatically different depending on whether $N = 3$ or $N = 4$. It is only in the former case that there is a ball of unistochastic matrices surrounding the van der Waerden matrix. On the other hand, the hyperoctants are not empty—some of them do contain unistochastic matrices all the way down to $J_4$.

Concerning the other hyperoctants, for types $\mathrm{II}_-$, $\mathrm{III}_+$, and $\mathrm{III}_-$ the central rays hit the boundary of the polytope in points that are not unistochastic, but numerically we find that a part of the ray close to $J_4$ is unistochastic. For type $\mathrm{II}_+$ we hit the boundary in a unistochastic point and numerically we find the entire ray to be unistochastic. There is still much that we do not know. We do not know if the hyperoctants of type $\mathrm{I}_-$ are entirely free of unistochastic matrices, nor do we know if $\mathcal{U}_4$ is star shaped, or what its relative volume may be. What is clear from the results that we do have is that the global structure of Birkhoff's polytope reverberates in the structure of the unistochastic subset in an interesting way—it is a little bit like a nine dimensional snowflake, because the nine hyperplanes in $\mathcal{B}_4$ can be found through an analysis of the behaviour of $\mathcal{U}_4$ in the neighbourhood of $J_4$.

## 5. Conclusions

Our reasons for studying the unistochastic subset of Birkhoff's polytope have been summarized in the introduction. Because the problem is a difficult one we concentrated on the cases $N = 3$ and $N = 4$. Our descriptions of Birkhoff's polytope for these two cases are given in Theorems 1 and 2, respectively, and a characterization sufficient for our purposes of the unistochastic set for $N = 3$ is given in Theorem 3. For $N = 4$ the dimension of the unistochastic set is again equal to that of the polytope itself, but its structure differs dramatically from the $N = 3$ case. In particular Theorem 4 states that for $N = 4$ there are non-unistochastic matrices in every neighbourhood of the van der Waerden matrix. Hence there does not exist a unistochastic ball surrounding the van der Waerden matrix. We observed that the structure of the unistochastic set at the center of the polytope reflects the global structure of the latter in an interesting way.

It is natural to ask to what extent the difference between the two cases is due to the fact that 3 is prime while 4 is not. Although this is not the place to discuss the cases $N > 4$, let us mention that we have reasons to believe that the dimension of the unistochastic set is equal to that of $\mathcal{B}_N$ for all values of $N$ [29]. On the other hand it is only when $N$ is a prime number that we have been able to show that there is a unistochastic ball surrounding the van der Waerden matrix.

*Acknowledgements.* We thank Göran Björck, Prot Pakoński, Wojciech Słomczyński, and Gregor Tanner for discussions, Petre Dĭţa for email correspondence, Uffe Haagerup for supplying us with a copy of Petrescu's thesis, and an anonymous referee for comments. Financial support from the Swedish Research Council VR, and from the Polish Ministry of Scientific Research under grant No PBZ-MIN-008/P03/2003, is gratefully acknowledged.

## Appendix A: Notation

For $N = 4$ we have defined the 24 permutation $P_0, \ldots, P_{23}$ matrices in a lexicographical order. They can be regarded as the corners of 6 regular tetrahedra, that can be written in the form

$$T_1 = p_0 P_0 + p_7 P_7 + p_{16} P_{16} + p_{23} P_{23} = \begin{bmatrix} p_0 & p_7 & p_{16} & p_{23} \\ p_7 & p_0 & p_{23} & p_{16} \\ p_{16} & p_{23} & p_0 & p_7 \\ p_{23} & p_{16} & p_7 & p_0 \end{bmatrix},$$

$$T_2 = p_1 P_1 + p_6 P_6 + p_{17} P_{17} + p_{22} P_{22} = \begin{bmatrix} p_1 & p_6 & p_{17} & p_{22} \\ p_6 & p_1 & p_{22} & p_{17} \\ p_{22} & p_{17} & p_6 & p_1 \\ p_{17} & p_{22} & p_1 & p_6 \end{bmatrix},$$

$$T_3 = p_2 P_2 + p_{10} P_{10} + p_{13} P_{13} + p_{21} P_{21} = \begin{bmatrix} p_2 & p_{10} & p_{13} & p_{21} \\ p_{13} & p_{21} & p_2 & p_{10} \\ p_{10} & p_2 & p_{21} & p_{13} \\ p_{21} & p_{13} & p_{10} & p_2 \end{bmatrix},$$

$$T_4 = p_3 P_3 + p_{11} P_{11} + p_{12} P_{12} + p_{20} P_{20} = \begin{bmatrix} p_3 & p_{11} & p_{12} & p_{20} \\ p_{12} & p_{20} & p_3 & p_{11} \\ p_{20} & p_{12} & p_{11} & p_3 \\ p_{11} & p_3 & p_{20} & p_{12} \end{bmatrix},$$

$$T_5 = p_4 P_4 + p_8 P_8 + p_{15} P_{15} + p_{19} P_{19} = \begin{bmatrix} p_4 & p_8 & p_{15} & p_{19} \\ p_{19} & p_{15} & p_8 & p_4 \\ p_8 & p_4 & p_{19} & p_{15} \\ p_{15} & p_{19} & p_4 & p_8 \end{bmatrix},$$

$$T_6 = p_5 P_5 + p_9 P_9 + p_{14} P_{14} + p_{18} P_{18} = \begin{bmatrix} p_5 & p_9 & p_{14} & p_{18} \\ p_{18} & p_{14} & p_9 & p_5 \\ p_{14} & p_{18} & p_5 & p_9 \\ p_9 & p_5 & p_{18} & p_{14} \end{bmatrix}.$$

These expressions also implicitly define our numbering convention for the permutation matrices.

The nine hyperplanes mentioned in Theorem 2 consist of matrices of the form

$$\Pi_1 = \begin{bmatrix} B_{00} & B_{01} & \bullet & \bullet \\ B_{10} & B_{11} & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \end{bmatrix}, \qquad \Pi_2 = \begin{bmatrix} B_{00} & B_{01} & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ B_{20} & B_{21} & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \end{bmatrix},$$

$$\Pi_3 = \begin{bmatrix} B_{00} & B_{01} & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ B_{30} & B_{31} & \bullet & \bullet \end{bmatrix}, \qquad \Pi_4 = \begin{bmatrix} B_{00} & \bullet & B_{02} & \bullet \\ B_{10} & \bullet & B_{12} & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \end{bmatrix},$$

$$\Pi_5 = \begin{bmatrix} B_{00} & \bullet & B_{02} & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ B_{20} & \bullet & B_{22} & \bullet \\ \bullet & \bullet & \bullet & \bullet \end{bmatrix}, \qquad \Pi_6 = \begin{bmatrix} B_{00} & \bullet & B_{02} & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ B_{30} & \bullet & B_{32} & \bullet \end{bmatrix},$$

$$\Pi_7 = \begin{bmatrix} B_{00} & \bullet & \bullet & B_{03} \\ B_{10} & \bullet & \bullet & B_{13} \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \end{bmatrix}, \qquad \Pi_8 = \begin{bmatrix} B_{00} & \bullet & \bullet & B_{03} \\ \bullet & \bullet & \bullet & \bullet \\ B_{20} & \bullet & \bullet & B_{23} \\ \bullet & \bullet & \bullet & \bullet \end{bmatrix},$$

$$\Pi_9 = \begin{bmatrix} B_{00} & \bullet & \bullet & B_{03} \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ B_{30} & \bullet & \bullet & B_{33} \end{bmatrix},$$

where the matrix elements that are explicitly written are assumed to sum to one (hence this holds also for the remaining three blocks taken separately).

The normal vectors of these hyperplanes are the matrices

$$n_1 = \frac{1}{4} \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}$$

and so on.

## Appendix B: Numerics

**I**. Entropy averaged over $\mathcal{B}_3$. To generate a random bistochastic matrix from the flat measure on $\mathcal{B}_3 \subset \mathbb{R}^4$, we have drawn at random a point $(x, y, z, t)$ in the 4-dimensional hypercube. It determines a minor of a $N = 3$ matrix $B$, and the remaining five elements of $B$ may be determined by the unit sum conditions in Eq. (1). Condition i is fulfilled if the sums in both rows and both columns of the minor does not exceed unity, and the sum of all four elements is not smaller than one. If this was the case, the random matrix $B$ was accepted to the ensemble of random bistochastic matrices. If additionally, the chain links condition (21) were satisfied, the matrix was accepted to the ensemble of unistochastic matrices, generated with respect to the flat measure on $\mathcal{U}_3$. The mean entropies, (25), were computed by taking an average over both ensembles consisting of $10^7$ random matrices, respectively.

**II**. Numerical verification, whether a given bistochastic matrix $B$ is unistochastic. We have performed a random walk in the space of unitary matrices. Starting from an arbitrary random initial point $U_0$ we computed $B_0 = U_0 \circ U_0^*$ and its distance to the analyzed matrix, $D_0 = D(B_0, B)$, as defined in (4). We fixed a small parameter $\alpha \approx 0.1$, generated a random Hermitian matrix $H$ from the Gaussian unitary

ensemble [30], and found a unitary perturbation $V = \exp(-i\alpha H)$. The matrix $U_{n+1} = VU_n$ was accepted as a next point of the random trajectory if the distance $D_{n+1}$ was smaller than the previous one, $D_n$. If a certain number (say 100) of random matrices $V$ did not allow us to decrease the distance, we reduced the angle $\alpha$ by half, to start a finer search. A single run was stopped if the distance $D$ was smaller than $\epsilon = 10^{-6}$ (numerical solution found), or $\alpha$ got smaller than a fixed cut off value (say $\alpha_{\min} = 10^{-4}$). In the latter case, the entire procedure was repeated a hundred times, starting from various unitary random matrices $U_0$, generated from the Haar measure on $U(4)$ [31]. The smallest distance $D_{\min}$ and the closest unistochastic matrix $B_{min} = U_n \circ \bar{U}_n$ were recorded.
To check the accuracy of the algorithm we constructed several random unistochastic matrices, $B = U \circ U^*$, and verified that the random walk procedure gave their approximations with $D_{\min} < \epsilon$.

## Appendix C: A System of Equations

In order to curtail a plethora of indices in Eqs. (34–36) and ease the subsequent notation, let us introduce shorthand: $\varphi_j = \phi_{j1}$, $\psi_j = -\phi_{j2}$, $j = 1, 2, 3$. With that the system reads

$$e^{i\varphi_1} + e^{i\varphi_2} + e^{i\varphi_3} = -L, \tag{39}$$

$$e^{i\psi_1} + e^{i\psi_2} + e^{i\psi_3} = -L, \tag{40}$$

$$e^{i(\varphi_1+\psi_1)} + e^{i(\varphi_2+\psi_2)} + e^{i(\varphi_3+\psi_3)} = -L. \tag{41}$$

We shall prove the following:

**Lemma 2.** *The system of Eqs. (39–41)*

1. *has no real solutions for* $L > 1$,
2. *for* $0 < L < 1$ *has the solution*

$$\begin{matrix} \varphi_1 = 0, & \varphi_2 = \phi, & \varphi_3 = -\phi \\ \psi_1 = -\phi, & \psi_2 = 0, & \psi_3 = \phi \end{matrix}\,, \qquad \cos\phi = \frac{t-1}{t+1} = -\frac{L+1}{2}\,, \tag{42}$$

*unique up to obvious permutations,*

3. *has continuous families of solutions for* $L = 0, 1$.

Indeed, each of the unimodal numbers $e^{i\varphi_k}$, $k = 1, 2, 3$ is a root of:

$$\begin{aligned} P(\lambda) &= (\lambda - e^{i\varphi_1})(\lambda - e^{i\varphi_2})(\lambda - e^{i\varphi_3}) \\ &= \lambda^3 - (e^{i\varphi_1} + e^{i\varphi_2} + e^{i\varphi_3})\lambda^2 + (e^{i(\varphi_1+\varphi_2)} + e^{i(\varphi_1+\varphi_3)} + e^{i(\varphi_2+\varphi_3)})\lambda \\ &\quad - e^{(i\varphi_1+\varphi_2+\varphi_3)} \\ &= \lambda^3 - (e^{i\varphi_1} + e^{i\varphi_2} + e^{i\varphi_3})\lambda^2 + (e^{-i\varphi_3} + e^{-i\varphi_2} + e^{-i\varphi_1})e^{(i\varphi_1+\varphi_2+\varphi_3)}\lambda \\ &\quad - e^{(i\varphi_1+\varphi_2+\varphi_3)} \\ &= \lambda^3 + \lambda^2 L - \lambda L e^{i\Phi} - e^{i\Phi} = \lambda^2(\lambda + L) - (1 + \lambda L)e^{i\Phi}, \end{aligned} \tag{43}$$

where $\Phi = \varphi_1 + \varphi_2 + \varphi_3$, and we used (39) and the reality of $L$. Thus each $\lambda = e^{i\varphi_k}$, $(k = 1, 2, 3)$, fulfills:

$$\lambda^2(\lambda + L) = (1 + \lambda L)e^{i\Phi}. \tag{44}$$

Analogously, $\mu = e^{i\psi_k}$, $(k = 1, 2, 3)$, fulfills

$$\mu^2(\mu + L) = (1 + \mu L)e^{i\Psi}, \tag{45}$$

with $\Psi = \psi_1 + \psi_2 + \psi_3$.

Observe now, that if $\lambda = e^{i\varphi_k}$ and $\mu = e^{i\psi_k}$ are solutions of (39–41) with the same number $k$ $(k = 1, 2, 3)$ then, upon the same reasoning applied to (41), $\lambda\mu$ fulfills

$$\lambda^2\mu^2(\lambda\mu + L) = (1 + \lambda\mu L)e^{i(\Phi+\Psi)}\,. \tag{46}$$

Multiplying (44) by (45) and finally by (46) after exchanging its sides, we obtain, after division by $\lambda^2\mu^2 e^{i(\Phi+\Psi)} \neq 0$,

$$(L + \lambda)(L + \mu)(L\lambda\mu + 1) = (L\lambda + 1)(L\mu + 1)(L + \lambda\mu), \tag{47}$$

which, upon substitution $\lambda = e^{i\varphi_k}$, $\mu = e^{i\psi_k}$ and putting everything on one side factorizes to

$$L(L - 1)(e^{i\varphi_k} - 1)(e^{i\psi_k} - 1)(e^{i(\varphi_k+\psi_k)} - 1) = 0, \tag{48}$$

(any computer symbolic manipulation program can be helpful in revealing (48) from (47)).

Hence, if $L \neq 0, 1$, then for each pair $(\varphi_k, \psi_k)$, $k = 1, 2, 3$, either: a) one of the angles is zero or b) they are opposite. The latter case can not occur for all three pairs since then $e^{i(\varphi_1+\psi_1)} + e^{i(\varphi_2+\psi_2)} + e^{i(\varphi_3+\psi_3)} = 3 \neq -L$, hence at least one of $\varphi_k$ or $\psi_k$ equals zero. Up to unimportant permutations we can assume $\varphi_3 = 0$, but then, since $e^{i\varphi_1} + e^{i\varphi_2} + e^{i\varphi_3} = -L \in \mathbb{R}$, we immediately get $\varphi_1 = -\varphi_2$. This determines also all other angles (also up to some unimportant permutation), and we end up with the solution announced in point 2 above as the only possibility, but such a solution exists only if $L \leq 1$.

To prove point 3, observe that

1. for $L = 0$,

$$\varphi_1 = \varphi, \quad \varphi_2 = \varphi + 2\pi/3, \quad \varphi_3 = \varphi + 4\pi/3, \tag{49}$$
$$\psi_1 = \psi, \quad \psi_2 = \psi + 2\pi/3, \quad \psi_3 = \psi + 4\pi/3, \tag{50}$$

   is a legitimate solution of (39–41) for arbitrary $\varphi$ and $\psi$,
2. for $L = 1$,

$$\varphi_1 = \varphi, \quad \varphi_2 = \pi, \quad \varphi_3 = \varphi + \pi, \tag{51}$$
$$\psi_1 = -\varphi + \pi, \quad \psi_2 = \pi, \quad \psi_3 = -\varphi, \tag{52}$$

   is a solution for an arbitrary $\varphi$.

## References


1. Marshall, A.W., Olkin, I.: *Inequalities: Theory of Majorization and its Applications.* New York: Academic Press 1979
2. Landé, A.: *From Dualism to Unity in Quantum Physics*. Cambridge: Cambridge U. P. 1960
3. Rovelli, C.: Relational Quantum Mechanics. Int. J. of Theor. Phys. **35**, 1637 (1996)
4. Khrennikov, A.: Linear representations of probabilistic transformations induced by context transition. J. Phys. **A34**, 9965 (2001)
5. Tanner, G.: Unitary-stochastic matrix ensembles, and spectral statistics. J. Phys. **A34**, 8485 (2001)
6. Pakoński, P., Życzkowski, K., Kuś, M.: Classical 1D maps, quantum graphs and ensembles of unitary matrices. J. Phys. **A34**, 9303 (2001)
7. Pakoński, P., Tanner, G., Życzkowski, K.: Families of Line-Graphs and Their Quantization. J. Stat. Phys. **111**, 1331 (2003)
8. Życzkowski, K., Kuś, M., Słomczyński, W., Sommers, H.-J.: Random unistochastic matrices. J. Phys. **A36**, 3425 (2003)
9. Jarlskog, C., Stora, R.: Unitarity polygons and $CP$ violation areas and phases in the standard electroweak model. Phys. Lett. **B208**, 268 (1988)
10. Auberson, G., Martin, A., Mennessier, G.: On the Reconstruction of a Unitary Matrix from its Moduli. Commun. Math. Phys. **140**, 523 (1991)
11. Mennessier, G., Nyuts, J.: Some unitarity bounds for finite matrices. J. Math. Phys. **15**, 1525 (1974)
12. Sylvester, J.J.: Thoughts on inverse orthogonal matrices, simultaneous sign-successions, and tesselated pavements in two or more colours, with applications to Newton's rule, ornamental tile-work, and the theory of numbers. Phil. Mag. **34**, 461 (1867)
13. Petrescu, M.: *Existence of continuous families of complex Hadamard matrices of certain prime dimensions and related results*. UCLA thesis, Los Angeles 1997
14. Dĭţa, P.: Some results on the parametrization of complex Hadamard matrices. J. Phys. **A37**, 5355 (2004)
15. Zeilinger, A., Żukowski, M., Horne, M.A., Bernstein, H.J., Greenberger, D.M.: Einstein-Podolsky-Rosen correlations in higher dimensions. In: J. Anandan, J. L. Safko, eds. *Fundamental Aspects of Quantum Theory*, Singapore: World Scientific, 1993
16. Törmä, P., Stenholm, S., Jex, I.: Hamiltonian theory of symmetric optical network transforms. Phys. Rev. **A52**, 4853 (1995)
17. Werner, R.F.: All Teleportation and Dense Coding Schemes. J. Phys. **A34**, 7081 (2001)
18. Birkhoff, G.: Tres observaciones sobre el algebra lineal. Univ. Nac. Tucumán Rev. **A5**, 147 (1946)
19. Brualdi, R.A., Gibson, P.M.: Convex Polyhedra of Doubly Stochastic Matrices. I. Applications of the Permanent Function. J. Comb. Theory **A22**, 194 (1977)
20. Słomczyński, W.: Subadditivity of Entropy for Stochastic Matrices. Open Sys. Inf. Dyn. **9**, 201 (2002)
21. Beck, M., Pixton, D.: The volume of the $10^{th}$ Birkhoff polytope. arXiv: math.CO/0305322
22. Au-Yeung, Y.-H., Poon, Y.-T.: $3 \times 3$ Orthostochastic Matrices and the Convexity of Generalized Numerical Ranges. Lin. Alg. Appl. **27**, 69 (1979)
23. Nakazato, H.: Sets of $3 \times 3$ Orthostochastic Matrices. Nikonkai Math. J. **7**, 83 (1996)
24. Gadiyar, H.G., Maini, K.M.S., Padma, R., Sharatchandra, H.S.: Entropy and Hadamard matrices. J. Phys. **A36**, L109 (2003)
25. Jones, K.R.W.: Entropy of random quantum states. J. Phys. **A23**, L1247 (1990)
26. Au-Yeung, Y.-H., Cheng, C.-M.: Permutation Matrices Whose Convex Combinations Are Orthostochastic. Lin. Alg. Appl. **150**, 243 (1991)
27. Hadamard, M.J.: Résolutions d'une question relative aux déterminants. Bull. Sci. Math. **17**, 240 (1893)
28. Haagerup, U.: Orthogonal maximal Abelian $*$-subalgebras of the $n \times n$ matrices and cyclic $n$-roots. In: *Operator Algebras and Quantum Field Theory, Rome (1996)*, Cambridge, MA: Internat. Press, 1997
29. Tadej, W.: Unpublished
30. Mehta, M.L.: *Random Matrices*. II ed., New York: Academic, 1991
31. Poźniak, M., Życzkowski, K., Kuś, M.: Composed ensembles of random unitary matrices. J. Phys. **A31**, 1059 (1998)


Communicated by M.B. Ruskai