LINEAR ALGEBRA AND ITS APPLICATIONS

ELSEVIER Linear Algebra and its Applications 289 (1999) 243–259

The Marcus–de Oliveira conjecture, bilinear forms, and cones

Alexander Kovačec [1]

Departamento de Matemática, Universidade de Coimbra, Apartado 3008, P-3000 Coimbra, Portugal

Received 27 May 1996; accepted 2 May 1997

Submitted by G.P.H. Styan

Abstract

The well-known determinantal conjecture of de Oliveira and Marcus (OMC) confines the determinant $\det(X+Y)$ of the sum of normal $n \times n$ matrices X, Y to a certain region in the complex plane. Even the subconjecture obtained by specializing it to $n=4$, X Hermitian and Y normal is still open. We view the subconjecture as a special case of an assertion concerning a certain family of bilinear forms on $\mathbb{R}^{16} \times \mathbb{C}^{16}$ and give a method that may prove useful for establishing it for many of such matrix pairs, independent of their spectrum; in particular we apply it successfully in the case of a prominent unitary similarity of Drury's threatening OMC. Unfortunately we find the assertion, extended naturally to pairs of complex arguments to be false and the ideas outlined inapplicable for the general OMC($n=4$) case. We also report on some computer experiments, formulate OMC($n=4$) as a statement about cones, and find it would be implied by establishing the emptiness of certain semialgebraic sets defined by systems of quadratic and linear relations.

1. Introduction

a. We use the notation $\mathbb{Q}, \mathbb{R}, \mathbb{C}; M_{n,m}(\mathbb{R})$ etc.; $A^{-1}, A \oplus B, AB, Ax, x^*Ay$ etc. with their universally adopted meaning in Linear Algebra and assume that sizes and properties of the objects involved make sense. Notation, definitions, conventions, properties, and observations on many of the other objects that we will use are listed in telegram style, as follows:

[1] E-mail: kovacec@mat.uc.pt. Research sponsored by JNICT Projecto 1329/CEN.

PII: S0024-3795(97)10004-0

$\mathbb{R}^n, \mathbb{C}^n$	columnspaces of real and complex n-vectors.
$[n]$; I, J, I^c, etc.; $\lvert I\rvert$	set $\{1, 2, \dots, n\}$; subsets of $[n]$. I^c being the complement of I, etc.; cardinality of I.
$S_n; \sigma; \langle k_1, \dots, k_s\rangle$	symmetric group acting on $[n]$; element of S_n; cycle of length s in S_n.
a_{ij} or A_{ij} etc.	entry with address (i,j) of complex matrix A, etc.
$A^{\mathrm{T}}, \bar{A}, A^*, \lvert A\rvert, A^{\updownarrow}, A^{\leftrightarrow}, A^{\mathrm{c}}$	matrices derived from $A \in M_{m,n}$, whose (i,j)-entries are given respectively by a_{ji}, $\bar{a}_{ij}$, $\bar{a}_{ji}$, $\lvert a_{ij}\rvert$, $a_{m+1-i,j}$, $a_{i,n+1-j}, a_{m+1-i,n+1-j}$. Matrices $A^{\leftrightarrow}, A^{\updownarrow}, A^{\mathrm{c}}$ are obtained respectively by reversing the order of rows, columns, and by reflecting entries of A at center. A is called centrosymmetric if $A^{\mathrm{c}} = A$. One has $A^{\mathrm{T}\leftrightarrow} = A^{\updownarrow\mathrm{T}}$, $A^{\mathrm{c}} = A^{\updownarrow\leftrightarrow}$.
$A \circ B, A^{\circ 2}$	Hadamard Product of $A, B \in M_{m,n}(\mathbb{C})$: $(A \circ B)_{ij} = (A_{ij}B_{ij})$; $A^{\circ 2} = A \circ A$.
$A[I\|J], [\,], [a]$	the submatrix $(a_{ij})_{i \in I, j \in J}$ of A, empty matrix and the 1×1 matrix with a single entry $a \in \mathbb{C}$, respectively.
$0_{k\times l}/J_{k\times l}; I_k$	the $k \times l$ matrices containing only entries 0/1; the $k \times k$ identity matrix.
$P, P_\sigma; O, Q; U$	usually intended for permutations; real orthogonal matrices; unitary matrices respectively. $(P_\sigma)_{ij} = \delta_{\sigma(i),j}$. In geometric context, O can mean 'origin'.
$O(n), \mathrm{SU}(n)$	real orthogonal group, special unitary group.
$\mathrm{supp}(A)$	support of a matrix $A = (a_{ij})$; that is the set of (i,j) for which $a_{ij} \neq 0$. If J is $(0,1)$-matrix and $\mathrm{supp}(A) \subseteq \mathrm{supp}(J)$, then $A \circ J = A$.
$s(A)$	sum of the entries of matrix A.
$\langle A, B\rangle$	Frobenius inner product of $A, B \in M_n(\mathbb{C})$. $\langle A, B\rangle = \mathrm{tr}(AB^*) = s(A \circ \overline{B}) = \langle UAV, UBV\rangle$ for unitary U, V. For real A and orthogonal Q_1, Q_2, $A \mapsto Q_1AQ_2$ is Euclidean isometry; if $Q_1 = Q_2^{\mathrm{T}}$, a rotation. See [10], Problems 5.1.4, etc.
$C_k(A)$	for $k = 0, \dots, n$ the k-th compound of $n \times n$ A.: size is $\binom{n}{k} \times \binom{n}{k}$, $C_0(A) = [\det([\,])] = [1]$, $C_1(A) = A$, $C_n(A) = [\det(A)]$, $\lvert C_k(P_\sigma)\rvert$ is a permutation matrix, $C_k(U)$ is unitary, $\lvert C_k(U)\rvert = \lvert C_{n-k}(U)\rvert^{\mathrm{c}}$; for $n = 4$ the $\lvert C_2(U)\rvert^{\circ 2}$ are centrosymmetric doubly stochastic. More in [10], 0.8.1, 0.8.4.
$\mathrm{co}S, \mathrm{aff}S$	convex and affine hull (span) of a subset S of a real vector space.
$\overrightarrow{AB}$	vector from point A to point B in affine space.

Ω_n	the polytope of $n \times n$ doubly stochastic matrices.
$\bigwedge, \neg$	logical symbols for 'for all', and 'not'.
$R(\sigma, U)$	The $2^n \times 2^n$ real matrix $\oplus_{k=0}^n(\lvert C_k(P_\sigma)\rvert - \lvert C_k(U)\rvert^{\circ 2})$.
$x_I, y_J; x_{1234}$, etc.	The products $\prod_{i\in I} x_i, \prod_{j\in J} y_j$; abbreviation for $x_{\{1,2,3,4\}}$ etc.
$\mathscr{P}_{k\mathrm{perm}}$	Polytope $\mathrm{co}\{\lvert C_k(P_\sigma)\rvert:\ \sigma \in S_n\}$; $\mathscr{P}_{1\mathrm{perm}} = \Omega_n$. We mostly use $\mathscr{P}_{2\mathrm{perm}}$.

The choice of notation also complies for the most part with current standards in matrix theory, combinatorics, geometry, and logic, as given by [10–12,8]. Given $\mathscr{R} \subseteq M_{p,q}(\mathbb{R})$, we will write $x^*\mathscr{R}y$ as a shorthand for $\{x^*Ry:\ R \in \mathscr{R}\}$; similarly, if $\mathscr{R} \subseteq \mathbb{R}$, then $0 < \mathscr{R}$ might stand for 'for all $r \in \mathscr{R}$ there holds $0 < r$'; further devices of this sort are used frequently to lighten notation. For 'left/right hand side of ...' we may use the abbreviations lhs(.)/rhs(.). There is only little further specific symbology; it will be explained along the text. Phrases like 'for all k' etc. should be interpreted as 'for all k for which ... makes sense' or alike, etc. Occasional use of logic expressions will clarify relations between various statements; references to [8] are meant for those seeking rapid information and standard literature.

b. Contents. Consider the following two predicates whose arguments are n-tuples $x = [x_1, \ldots, x_n]$ and $y = [y_1, \ldots, y_n]$ of complex numbers and unitary matrices U.

$\mathrm{OM}_1(x, y)$: *If X and Y are normal matrices in $\mathbb{C}^{n\times n}$ with eigenvalues $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ then*

$$\det(X + Y) \in \mathrm{co}\left\{\prod_{i=1}^{n}(x_i + y_{\sigma(i)}):\ \sigma \in S_n\right\}. \tag{1}$$

$\mathrm{OM}_2(x, y, U)$: *There holds*

$$0 \in \mathrm{co}\left\{\sum_{I,J:|I|=|J|}(\lvert\det P_\sigma[I|J]\rvert - \lvert\det U[I|J]\rvert^2)\prod_{i\in I^c} x_i \prod_{i\in J} y_j:\ \sigma \in S_n\right\}. \tag{2}$$

Using that normal matrices are unitarily similar to diagonal ones, the nontrivial arguments in [20] prove these predicates are equivalent in the following sense:

$$\mathrm{OM}_1(x,y) \iff \bigwedge_{U\in \mathrm{SU}(n)} \mathrm{OM}_2(x, y, U).$$

The well-known conjecture explicitly put forth by de Oliveira [20] in 1982, later discovered by Miranda and Bebiano in a 1973 paper by Marcus [16], states that $\mathrm{OM}_1(x, y)$ holds for all $x, y \in \mathbb{C}^n$. This extends a theorem of Fiedler

[9] according to which this is true for real x, y; that is (1) holds for Hermitian X, Y.

Of course, if the $n!$ expressions occurring in rhs(2) can be convexly combined so as to make the coefficients of all $x_I y_J$ equal to zero, then (2) will hold. In terms of matrices $R(\sigma, U) = \oplus_{k=0}^{n}(|C_k(P_\sigma)| - |C_k(U)|^{\circ 2})$, this establishes the implication

$$0_{2^n \times 2^n} \in \text{co}\{R(\sigma, U)\colon\ \sigma \in S_n\} \Rightarrow \bigwedge_{x,y \in C^n} \text{OM}_2(x, y, U). \tag{3}$$

If $n = 3$, then $|U| = |C_1(U)| = |C_2(U)|^c$ and lhs(3) can be inferred from Birkhoff's representation theorem (Theorem B, Section 2) for doubly stochastic matrices. This argument yields the main result of [1]. But for $n > 3$ the $n!$ k-ply permutations of S_n form in general a very small proper subset of all the $\binom{n}{k}!$ permutation matrices of size $\binom{n}{k}$. Thus it is not clear whether $\oplus_{k=0}^{n}|C_k(U)|^{\circ 2}$ lies for every unitary U in the polytope $\text{co}\{\oplus_{k=0}^{n}|C_k(P_\sigma)|\colon\ \sigma \in S_n\}$. We shall call unitary U for which this holds *internal*; otherwise *external*. There where good reasons, established by work of Bebiano, Merikoski, Virtanen, da Providência [18,1,19] to believe that all unitary matrices are internal; the OM-conjecture would follow. It came as an unpleasant surprise when Drury [5] gave an example of a matrix $U \doteq U_{\text{Dru}}$ in SU(4) which is external; i.e. for which there holds ¬lhs(3); see Section 2.

In Drury's example and other known external U, the distance of $0_{16\times 16}$ from $\text{co}\{R(\sigma, U)\colon\ \sigma \in S_4\}$ is minute – see Section 5. Writing (2), in terms of x_I and y_J, this suggested to us that a generalization of OM_2 to a statement about a certain family of bilinear forms might hold; after all it is clear that (4) below holds for all $x, y \in \mathbb{C}^{2^n} \times \mathbb{C}^{2^n}$ if U is internal and much information is lost upon mapping the $R(\sigma, U)$ via x, y as below to the complex plane.

Main Theorem. (a) *For $n = 4$, all $(x, y) \in \mathbb{R}^{16} \times \mathbb{C}^{16}$, and $U = U_{\text{Dru}}$, there holds*

$$0 \in \text{co}\{x^* R(\sigma, U) y\colon\ \sigma \in S_n\}. \tag{4}$$

(b) *However, for any external U there exist $(x, y) \in \mathbb{C}^{16} \times \mathbb{C}^{16}$ such that* (4) *does not hold.*

More important than part (a) itself seems to be the method establishing it. External unitary matrices seem to be scarce. For reasons we give later we put forth the following somewhat vague conjecture.

Conjecture. The heuristics outlined in Section 3 for establishing part (a) of the theorem will succeed for any external $U \in \text{SU}(4)$ instead of $U = U_{\text{Dru}}$.

As follows from the discussions above, could part (a) of the Main Theorem be established without any restriction concerning the $U \in \text{SU}(4)$, this would imply (1) whenever X is Hermitian and Y is a normal 4×4 matrix. This would extend Fiedler's result for $n = 4$. We are unwilling to conjecture anything for the case $n \geqslant 5$. If we are to admit complex x, y unrestrictedly however, then part (b) shatters the hopes for the bilinear generalization completely.

The Main Theorem is established by means of some facts on the representation theory of the symmetric group S_4. These facts allow a reduction of the dimensionality of the problem from 16×16 matrices to one involving 2×2 matrices. They are collected in Section 2 which terminates with a proof of part (a) of the Main Theorem. As a byproduct we extend Fiedler's result to the case of pairs X, Y of matrices where X is Hermitian and Y normal with one nonreal eigenvalue. The heuristic considerations underlying the judicious choice of certain parameters that worked for the historically important example U_{Dru} and as conjectured may work for the other external $U \in \mathrm{SU}(4)$ too, are explained in Section 3. Many of the calculations of this paper were done with MATLAB [17]. In Section 4 we prove part (b). We also give for the case $U = U_{\mathrm{Dru}}$ explicit x, y showing (4) to fail. In Section 5 we report on numerical experiments with a polytope that contains the same geometric information as the convex hull of the $R(\sigma, U_{\mathrm{Dru}})$, $\sigma \in S_4$. We formulate OMC as a conjecture about cones and find that OMC($n = 4$) is a bit 'less true' than Fiedler's result for Hermitian matrices. We use this new formulation to show that OMC could be inferred from the inconsistency of certain families of homogeneous polynomial equations of degrees 1 and 2 in many real variables, a number of which are required to be nonnegative. Since many of these equations might be superfluous, the proof of the inconsistency might lie within reach of current computing facilities. For another attack to the case $n = 4$ see the beautiful recent paper [7] by Drury.

2. Proof of part (a) of main theorem

The relevance of the representation theory of the symmetric group and in particular of the work of Saxl [22] for the OM-conjecture was discovered by Drury [5]. We elaborate first on his observations. The reader might consult [8], Articles 362 H, 368 G, for a first information on idempotents; we learned the details originally from [14]. Other approaches via Specht Modules are in [13] or [21].

The general theory associates Young tableaus to irreducible representations. The group of interest for us is S_4. The following three tableaus called alike here as later their corresponding representations will be of importance for us:

$$\underset{\pi_0,}{\begin{array}{|c|c|c|c|}\hline 1&2&3&4\\\hline\end{array}} \qquad \underset{\pi_1,}{\begin{array}{|c|c|c|}\hline 1&2&3\\\hline 4& \multicolumn{2}{c}{} \\\cline{1-1}\end{array}} \qquad \underset{\pi_2.}{\begin{array}{|c|c|}\hline 1&2\\\hline 3&4\\\hline\end{array}}$$

In the group algebra $\mathbb{Q}[S_4]$ one calculates the primitive idempotents e_i, $i = 0, 1, 2$ defined by

$$e_i \doteq \sum\{(\mathrm{sgn}\ \sigma)\sigma\colon\ \sigma \text{ stabilizes columns of } \pi_i\} \times \sum\{\sigma\colon\ \sigma \text{ stabilizes rows of } \pi_i\},$$

where $\doteq$ means 'equal up to some nonzero multiple'; see [14].

Thus in our case,

$$e_0 \doteq \sum_{\sigma\in S_4} \sigma,$$
$$e_1 \doteq (\mathrm{id} - \langle 1,4\rangle)(\mathrm{id} + \langle 1,2\rangle + \langle 1,3\rangle + \langle 2,3\rangle + \langle 1,2,3\rangle + \langle 1,3,2\rangle),$$
$$e_2 \doteq (\mathrm{id} - \langle 1,3\rangle - \langle 2,4\rangle + \langle 1,3\rangle\langle 2,4\rangle)(\mathrm{id} + \langle 1,2\rangle + \langle 3,4\rangle + \langle 1,2\rangle + \langle 3,4\rangle).$$

By the theory of semisimple rings, $\mathbb{Q}[S_4]$ has among its direct summands the minimal left ideals $\mathbb{Q}[S_4]e_0$, $\mathbb{Q}[S_4]e_1$, and $\mathbb{Q}[S_4]e_2$.

Extending the trivial 0-ply, the 1-ply, and the 2-ply matrix representations of S_4, namely $\sigma \overset{\pi^{(0)}}{\mapsto} [1]$, $\sigma \overset{\pi^{(1)}}{\mapsto} |C_1(P_\sigma)|$, and $\sigma \overset{\pi^{(2)}}{\mapsto} |C_2(P_\sigma)|$ to $\mathbb{Q}[S_4]$, we obtain as images matrix subalgebras of $M_1(\mathbb{R})$, $M_4(\mathbb{R})$, and $M_6(\mathbb{R})$ respectively. By [22], Lemma 2.2, these representations break up as follows into irreducible representations π_i with representation spaces given by image($\mathbb{Q}[S_4]e_i$):

$$\pi^{(0)} = \pi_0, \quad \pi^{(1)} = \pi_0 \oplus \pi_1, \quad \pi^{(2)} = \pi_0 \oplus \pi_1 \oplus \pi_2.$$

From this we read off the dimensions $\dim \pi_0 = 1$, $\dim \pi_1 = 4 - \dim \pi^{(0)} = 3$, $\dim \pi_2 = 6 - \dim \pi^{(1)} = 2$. Thus Saxl's theorem predicts the existence of 4×4 and 6×6 regular matrices Q, $\tilde{Q}$, respectively, such that

$$Q^{-1}|C_1(P_\sigma)|Q = \begin{bmatrix} 1 & & & \\ & * & * & * \\ & * & O(\sigma) & * \\ & * & * & * \end{bmatrix} \quad \text{and}$$

$$\tilde{Q}^{-1}|C_2(P_\sigma)|\tilde{Q} = \begin{bmatrix} 1 & & & & \\ & * & * & * & \\ & * & O(\sigma) & * & \\ & * & * & * & \\ & & & & O'(\sigma) \end{bmatrix}. \tag{5}$$

The orthogonality of the permutations $|C_1(P_\sigma)|$ and $|C_2(P_\sigma)|$ implies the orthogonality of the 3×3 and 2×2 matrices $O(\sigma)$ and $O'(\sigma)$, for these describe the action of the permutations with respect to their invariant subspaces. Basis vectors in $\mathbb{R}^4$ and $\mathbb{R}^6$ for the 1-, 3-, and (for the 6×6 case) also 2- dimensional subspaces left invariant by the images mentioned, will serve as columns for Q and $\tilde{Q}$.

Evaluating the e_i, we find in $M_4(\mathbb{R})$ that $4e_0 = J_{4\times 4}$, the matrix with only unit entries, and $e_2 = 0_{4\times 4}$, as predicted by Saxl's theorem; similarly in $M_6(\mathbb{R})$, $6e_0 = J_{6\times 6}$. More interesting idempotents are

$$\frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & -3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & 3 \end{bmatrix}, \quad \frac{1}{4}\begin{bmatrix} 1 & 1 & -1 & 1 & -1 & -1 \\ 1 & 1 & -1 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & 1 & -1 & 1 & 1 \\ -1 & -1 & 1 & -1 & 1 & 1 \end{bmatrix},$$

$$e_1 \in M_4(\mathbb{R}), \qquad e_1 \in M_6(\mathbb{R}),$$

$$\frac{1}{5}\begin{bmatrix} 2 & -1 & -1 & -1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -2 & 1 & 1 & 1 & 1 & -2 \\ -2 & 1 & 1 & 1 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & -1 & -1 & -1 & -1 & 2 \end{bmatrix}$$

$$e_2 \in M_6(\mathbb{R}).$$

Calculating and ordering adequately orthogonal bases for the spaces generated by the sets $\{|C_1(P_\sigma)|e_ix\colon\ \sigma \in S_4\}$ and $\{|C_2(P_\sigma)|e_ix\colon\ \sigma \in S_4\}$ – almost any x is suitable here – one finds the desired transformation matrices as follows:

$$Q = \begin{bmatrix} -1/2 & -1/\sqrt{2} & 0 & 1/2 \\ -1/2 & 0 & 1/\sqrt{2} & -1/2 \\ -1/2 & 0 & -1/\sqrt{2} & -1/2 \\ -1/2 & 1/\sqrt{2} & 0 & 1/2 \end{bmatrix},$$

$$\tilde{Q} = \begin{bmatrix} 1/\sqrt{6} & 1/2 & -1/2 & 0 & 1/2 & 1/\sqrt{12} \\ 1/\sqrt{6} & 1/2 & 1/2 & 0 & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & 0 & 0 & -1/\sqrt{2} & -1/2 & 1/\sqrt{12} \\ 1/\sqrt{6} & 0 & 0 & 1/\sqrt{2} & -1/2 & 1/\sqrt{12} \\ 1/\sqrt{6} & -1/2 & -1/2 & 0 & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & -1/2 & 1/2 & 0 & 1/2 & 1/\sqrt{12} \end{bmatrix}.$$

Notably Q and $\tilde{Q}$ are also orthogonal, so $Q^{\mathrm{T}} = Q^{-1}, \tilde{Q}^{\mathrm{T}} = \tilde{Q}^{-1}$. This can be explained from going through the proof of Maschke's Theorem in [21], 1.5.3. One finds that the representation of any finite, with respect to the ordinary scalar product orthogonal group, will be orthogonally equivalent to the direct sum of its irreducible representations: for in this case one can take in [21] for $\langle , \rangle$ the ordinary scalar product and spare the averaging process. Furthermore, since in our case $\langle e \rangle$ is an invariant space of dimension 1, $Q, \tilde{Q}$, have one column$\doteq e$ and thus their other columns have column sums 0. The following was an important step in work on OMC.

Theorem MV. *For any unitary U, the direct sum $\oplus_{k=0}^{n}|C_k(U)|^{\circ 2}$ lies in the affine hull of the $\oplus_{k=0}^{n}|C_k(P_\sigma)|$. Equivalently*

$$0_{2^n \times 2^n} \in \operatorname{aff}\{R(\sigma, U):\ \sigma \in S_n\}. \tag{6}$$

Proof. This is a restatement of a theorem of Merikoski and Virtanen [18], Theorem 4. □

It is worth noting that it was later established in [5,19], that any sequence $(t_\sigma)_{\sigma \in S_n}$ yielding $|C_{\lfloor n/2 \rfloor}(U)|^{\circ 2} = \sum_{\sigma \in S_n} t_\sigma |C_{\lfloor n/2 \rfloor}(P_\sigma)|^{\circ 2}$, serves also as coefficient sequence for an analogous linear combination for the direct sums. Thus the search for suitable sequences – they exist by Theorem MV – can be 'localized' to the natural k nearest to $n/2$.

Lemma 1. *With the matrices $Q, \tilde{Q}$ above, there holds for any unitary 4×4 matrix U, that:*

(a) *$Q^{\mathrm{T}}|C_1(U)|^{\circ 2}Q$ is equal to the left upper 4×4 submatrix of $\tilde{Q}^{\mathrm{T}}|C_2(U)|^{\circ 2}\tilde{Q}$.*
(b) *The latter matrix is of the direct sum form* $[1] \oplus (3 \times 3) \oplus (2 \times 2)$.

Proof. By Theorem MV there are reals s_σ summing to 1 such that $|C_1(U)|^{\circ 2} = \sum_\sigma s_\sigma |C_1(P_\sigma)|$, and $|C_2(U)|^{\circ 2} = \sum_\sigma s_\sigma |C_2(P_\sigma)|$. Now since by (5) statements (a), (b) hold whenever U is a permutation, the lemma follows. □

In the present context, the main message of [5] is that the s_σ cannot, in general, be chosen all nonnegative.

Define next the polytope $\mathscr{B}_n = \{\mathrm{t} = (t_\sigma)_{\sigma \in S_n}:\ t_\sigma \geqslant 0, \sum t_\sigma = 1\}$ of affine dimension $n! - 1$ along with the $(n^2 - 2n + 1)$-dimensional polytope Ω_n of $n \times n$ doubly stochastic matrices. Birkhoff's representation theorem [10], p. 527, can be stated as follows:

Theorem B. *There is a surjective map given by*

$$(t_\sigma) \overset{\beta}{\mapsto} \sum_\sigma t_\sigma |C_1(P_\sigma)|,$$

$$\mathscr{B}_n \to \Omega_n.$$

We would possibly profit from partial answers to the problems posed by Brualdi [3] concerning $(\mathscr{B}_n, \beta, \Omega_n)$; conversely, perhaps representation theory of S_n can help to advance Brualdi's. We return to the case $n = 4$.

Corollary 2. (a) *There exists a real orthogonal* 16×16 *matrix* $\mathbf{Q}$ *such that for all unitary* 4×4 *matrices* U *and* $\sigma \in S_n$ *there exist matrices* $3(\sigma, U)$ *and* $2(\sigma, U)$ *of sizes* 3×3 *and* 2×2 *respectively, such that*

$$\mathbf{Q}^{\mathrm{T}} R(\sigma, U)\mathbf{Q} = 0_{5\times 5} \oplus \left(\bigoplus_{j=1}^{3} 3(\sigma, U)\right) \oplus 2(\sigma, U). \tag{7}$$

(b) *If* $(t_\sigma) \in \beta^{-1}(|C_1(U)|^2)$ *then for some* 2×2 *matrix* $T = T((t_\sigma), U)$ *there holds*

$$\mathbf{Q}^{\mathrm{T}} \sum_{\sigma} t_\sigma R(\sigma, U)\mathbf{Q} = 0_{14\times 14} \oplus T. \tag{8}$$

Proof. (a) We write Q_0, Q_1, Q_2, Q_3, Q_4, respectively, for the matrices, $[1], Q, \tilde{Q}, Q^{\updownarrow}, [1]$. Just as the blocks in the definition of $R(\sigma, U)$, these matrices have the respective sizes $1, 4, 6, 4, 1$. Define $\dot{\mathbf{Q}} = \oplus_{j=0}^{4} Q_i$. We then obtain

$$\begin{aligned}\dot{\mathbf{Q}}^{\mathrm{T}} R(\sigma, U)\dot{\mathbf{Q}} &= \bigoplus_{i=0}^{4} Q_i^{\mathrm{T}}(|C_i(P_\sigma)| - |C_i(U)|^{\circ 2})Q_i \\ &= \bigoplus_{i=0}^{4}(Q_i^{\mathrm{T}}|C_i(P_\sigma)|Q_i - Q_i^{\mathrm{T}}|C_i(U)|^{\circ 2}Q_i).\end{aligned}$$

For $i = 0$ and $i = 4$, the 1×1 matrices $|C_i(U)|^{\circ 2}$ are for any unitary matrix U, in particular for the permutations equal to $[1]$. Therefore the direct summands with these indices are $[0]$. As a consequence of Lemma 1(b) the summand of index 2 is of the form $[0] \oplus 3(\sigma, U) \oplus 2(\sigma, U)$ where the matrices entering have the sizes claimed. By Lemma 1(a) the summand of index $i = 1$ is then $[0] \oplus 3(\sigma, U)$. This latter summand is also that of index $i = 3$: this is a consequence of the fact $|C_3(U)|^{\mathrm{c}} = |C_1(U)|$, and the computational rules for $\leftrightarrow, \updownarrow$ given in Section 1a. It follows from a count that there is a permutation P such that the matrix $\mathbf{Q} = \dot{\mathbf{Q}}P$ is as desired. (b) By hypothesis we have the left and therefore the right hand side of the implication

$$\begin{aligned}|C_1(U)|^{\circ 2} &= \sum t_\sigma |C_1(P_\sigma)| \\ \Rightarrow [0] \oplus 0_{3\times 3} &= \sum t_\sigma Q_1^{\mathrm{T}}(|C_1(P_\sigma)| - |C_1(U)|^{\circ 2})Q_1 \\ &= \sum t_\sigma [0] \oplus 3(\sigma, U).\end{aligned}$$

Canceling here the $[0]$ on the extremes, a look at (7) shows with $T = \sum_\sigma t_\sigma 2(\sigma, U)$ validity of (8). $\square$

Lemma 3. (a) *Assume* $\mathscr{A} \subseteq \mathbb{R}^q$ *to be finite. If there holds* $0 \in \mathrm{co}\{\mathscr{A}^{\mathrm{T}} y\}$ *for all* $y \in \mathbb{R}^q$, *then this holds also for all* $y \in \mathbb{C}^q$.

(b) *Assume* $\mathscr{R} \subset M_{p,q}(\mathbb{R})$ *to be finite. If for an* $x \in \mathbb{R}^p$ *and all* $y \in \mathbb{R}^q$ *there holds* $0 \in \mathrm{co}\{x^{\mathrm{T}} \mathscr{R} y\}$, *then this holds also for all* $y \in \mathbb{C}^q$.

Proof. (a) Let $\mathscr{A} = \{a_1, \ldots, a_m\}$. Consider the set of $1 \times q$ row vectors $\{a_1^{\mathrm{T}}, \ldots, a_m^{\mathrm{T}}\}$. The hypothesis of the lemma implies that this set must contain $0_{1\times q}$ in its convex hull – otherwise there existed by the hyperplane separation theorem, see [8], Article 89 or [10], p. 534, a real y such that all $a_i^{\mathrm{T}} y$ are > 0. Thus there exist $t_i \geqslant 0$ summing to 1 such that $\sum_i t_i a_i^{\mathrm{T}} = 0_{1\times q}$. But then for any $y \in \mathbb{C}^q$, $\sum_i t_i a_i^{\mathrm{T}} y = 0_{1\times q} y = 0$. (b) It is a direct consequence of (a) obtained from putting $\mathscr{A}^{\mathrm{T}} = x^{\mathrm{T}} \mathscr{R}$. □

As one consequence which we shall not use later we mention the following slight extension of Fiedler's result [9].

Corollary 4. *The Oliveira Marcus conjecture is true for all pairs of normal matrices $X, Y \in M_{n,n}(\mathbb{C})$ for which at most one matrix has at most one nonreal eigenvalue.*

Proof. This claim is equivalent to that $\mathrm{OM}_2(x, y, U)$ is valid whenever the n-tuple x is real while y (say) has at most one nonreal entry. Single out this one entry as a variable – call it $\tilde{y}$. Fix all the other real quantities in the expressions in the rhs(2). With suitable real a_σ, b_σ, $\sigma \in S_n$, these expressions can be seen as affine forms of the type

$$a_\sigma + b_\sigma \tilde{y} = [a_\sigma, b_\sigma] \begin{bmatrix} 1 \\ \tilde{y} \end{bmatrix}.$$

By Fiedler's result these forms contain for all real $\tilde{y}$ the real 0 in their convex hull. Passing to homogeneous linear forms, we can infer the claim from Lemma 3(a). □

Lemma 5. *Let A, B, C be real invertible 2×2 matrices such that $A^{-1}B$, $B^{-1}C$, $C^{-1}A$ have no real eigenvalues. Then there holds either for all $x \in \mathbb{R}^2 \setminus \{0_{2\times 1}\}$ or for no such x that $0_{2\times 1} \in \mathrm{co}\{Ax, Bx, Cx\}$.*

Proof. By hypothesis for any $x \in R^2 \setminus \{0_{2\times 1}\}$ and all $r \in \mathbb{R}$, $A^{-1}Bx \neq rx$. Equivalently, for all $x \neq 0$, $0_{2\times 1}, Ax, Bx$ are three distinct noncollinear points. Thus the straight line $g_{Ax,Bx}$ through Ax, Bx does not contain 0. By similar reasoning the straight lines $g_{Bx,Cx}$ and $g_{Cx,Ax}$ never hit the origin of the Euclidean plane. Consequently either the origin is contained for all $x \neq 0$ in the interior of the triangle with vertices Ax, Bx, Cx, or for no such x. □

Proof of Part (a) of the Main Theorem. Define the numbers $a = 0.16823\ldots$, $b = 0.56912\ldots$. With these numbers the matrix U_{Dru} is given by

$$U_{\text{Dru}} = \begin{bmatrix} a & b & b & b \\ -b & a & b & -b \\ -b & -b & a & b \\ -b & b & -b & a \end{bmatrix}.$$

Let us enumerate the permutations of S_4 lexicographically. Thus, assuming the permutations given in word representation, set up the correspondence $1 \leftrightarrow 1234, 2 \leftrightarrow 1243, \ldots$; see Section 3 for a dictionary. Write accordingly t_1 for t_{1234} etc. With $a_1 = 0.0283\ldots, a_2 = 0.2956\ldots, a_3 = 0.3239\ldots$, one can show the following $\mathbf{t} = (t_\sigma)$ to lie in $\beta^{-1}(|C_1(U_{\text{Dru}})|^{\circ 2})$:

$$\mathbf{t}' = \begin{pmatrix} 2 & 7 & 18 & 23 & 8 \\ a_1 & a_1 & a_3 & a_3 & a_2 \end{pmatrix}, \qquad \mathbf{t}'' = \begin{pmatrix} 3 & 11 & 14 & 22 & 24 \\ a_1 & a_3 & a_3 & a_1 & a_2 \end{pmatrix},$$
$$\mathbf{t}''' = \begin{pmatrix} 6 & 10 & 15 & 19 & 17 \\ a_1 & a_3 & a_1 & a_3 & a_2 \end{pmatrix}.$$

The first table should be read e.g. as $t_2 = a_1, t_7 = a_1, t_{18} = a_3, \ldots$, etc, and $t_i = 0$ for all $i \notin \{2, 7, 18, 23, 8\}$; similar for the other tables. The matrices $T' = T(\mathbf{t}', U_{\text{Dru}})$ etc., see Corollary 2(b), that arise from these choices are given approximately by

$$T' = \begin{pmatrix} 0.3323 & 0.8249 \\ 0.3952 & -0.3721 \end{pmatrix}, \qquad T'' = \begin{pmatrix} 0.3323 & -0.3952 \\ -0.8249 & -0.3721 \end{pmatrix},$$
$$T''' = \begin{pmatrix} -0.7243 & 0.2149 \\ -0.2149 & 0.6845 \end{pmatrix}.$$

One checks that the eigenvalues of $T'^{-1}T''$, $T''^{-1}T'''$, and $T'''^{-1}T'$ are nonreal – curiously there occurs the same pair in the three cases, namely $-0.6553 \pm 0.7553i$.

Since $(\mathbb{R}^{16})^{\mathrm{T}}\mathbf{Q}^{\mathrm{T}} = (\mathbb{R}^{16})^{\mathrm{T}}$, and $\mathbb{R}^{16} = \mathbf{Q}\mathbb{R}^{16}$, as columnspaces, we find the equivalence

$$\bigwedge_{x,y\in\mathbb{R}^{16}} 0 \in \operatorname{co}\{x^{\mathrm{T}}R(\sigma, U)y \colon \ \sigma \in S_4\}$$
$$\iff \bigwedge_{x,y\in\mathbb{R}^{16}} 0 \in \operatorname{co}\{x^{\mathrm{T}}\mathbf{Q}^{\mathrm{T}}R(\sigma, U)\mathbf{Q}y \colon \ \sigma \in S_4\} \tag{9}$$

and from Corollary 2(b),

$$\mathscr{T} = 0_{14\times 14} \oplus \{T', T'', T'''\} \subseteq \operatorname{co}\{\mathbf{Q}^{\mathrm{T}}R(\sigma, U_{\text{Dru}})\mathbf{Q} \colon \ \sigma \in S_4\}. \tag{10}$$

For any $y \in \mathbb{R}^{16}$, $(0_{14\times 14} \oplus T')y$ has $T'\binom{y_{15}}{y_{16}}$ as nontrivial part; similarly for T'', T'''. Now, since the first columns of the latter matrices, obtained by multiplication with $\binom{1}{0}$, contain $0_{2\times 1}$ in their convex hull, we see by Lemma 5 that for all $x, y \in \mathbb{R}^{16}$, $0 \in \mathrm{co}\{x^{\mathrm{T}}\mathscr{T}y\}$. Thus from relation (10), we find the validity of rhs(9) and thus of lhs(9). The latter fact implies part (a) of the Main Theorem by means of Lemma 3. □

3. On finding adequate elements in $\beta^{-1}(|C_1(U)|^{\circ 2})$

The proof in Section 2 uses the possibility of writing matrices in Ω_n in various ways as convex combinations of permutations. We give the ideas that guided our choices of $\mathbf{t}', \mathbf{t}'', \mathbf{t}''' \in \beta^{-1}(|C_1(U_{\mathrm{Dru}})|^{\circ 2})$. Probably part (a) of the Main Theorem can be established for other explicitly given external U by similar considerations.

Real multiples of 2×2 rotations $\neq I_2$ have no real eigenvalues. Inverses of reflections/rotations are reflections/rotations. Products of two reflections or two rotations are rotations. So if we could choose $\mathbf{t}', \mathbf{t}'', \mathbf{t}'''$ such that (i) T', T'', T''' are all small perturbations either of multiples of rotations or of reflections, and (ii) their first columns would include the origin of the Euclidean plane, then Lemma 5 guarantees $0 \in \mathrm{co}\{T'x, T''x, T'''x\}$ for all x.

To achieve this, we note from (5) and Corollary 2(b), that since $\sum_\sigma t_\sigma = 1$,

$$\begin{aligned} T((t_\sigma), U) &= \mathrm{lhs}(8)[15,16|15,16] \\ &= \sum_\sigma t_\sigma(\tilde{Q}^T|C_2(P_\sigma)|\tilde{Q})[5,6|5,6] - \underbrace{(\tilde{Q}^{\mathrm{T}}|C_2(U)|^{\circ 2})\tilde{Q})[5,6|5,6]}_{=:T(U)} \\ &= \left(\sum_\sigma t_\sigma \mathcal{O}(\sigma)\right) - T(U). \end{aligned}$$

It turns out that there occur only certain six orthogonal matrices $\mathcal{O}(\sigma)$. Putting $c = \sqrt{3} = 2\cos(\pi/6) = 1.7321\ldots$, the following table gives in the uppermost part these matrices, called in the line below $\mathcal{O}_1, \ldots, \mathcal{O}_6$, followed by their type. The column of $\mathcal{O}_i$ contains in its lowermost part the indications of the associated σ in word representation. The permutations are preceded by their lexicographic ordinal numbers in the sense mentioned in Section 2.

$$\frac{1}{2}\left\{ \underset{\mathcal{O}_1}{\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}}, \underset{\mathcal{O}_2}{\begin{bmatrix} 1 & c \\ c & -1 \end{bmatrix}}, \underset{\mathcal{O}_3}{\begin{bmatrix} 1 & -c \\ -c & -1 \end{bmatrix}}, \underset{\mathcal{O}_4}{\begin{bmatrix} -1 & -c \\ c & -1 \end{bmatrix}}, \underset{\mathcal{O}_5}{\begin{bmatrix} -1 & c \\ -c & -1 \end{bmatrix}}, \underset{\mathcal{O}_6}{\begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix}} \right\}$$

$\mathcal{O}_1$	$\mathcal{O}_2$	$\mathcal{O}_3$	$\mathcal{O}_4$	$\mathcal{O}_5$	$\mathcal{O}_6$
rotation	reflection	reflection	rotation	rotation	reflection
1 1234	2 1243	3 1324	4 1324	5 1423	6 1432
8 2143	7 2134	11 2423	12 2431	9 2314	10 2341
17 3412	18 3421	14 3142	13 3124	16 3241	15 3214
24 4321	23 4312	22 4231	21 4213	20 4132	19 4123

The standard method to find elements in $\beta^{-1}(D)$, $D \in \Omega_n$, is the Birkhoff algorithm:

Input: $D \in \Omega_n$
Output: A set of $t_\sigma \geqslant 0$ such that $\sum_\sigma t_\sigma P_\sigma = D$.
$\dot{D} := D$
while $\dot{D} \neq 0_{n\times n}$
choose a $\sigma \in S_n$ such that $t_\sigma := \min \dot{D}_{\sigma i,i} > 0$.
$\dot{D} := \dot{D} - t_\sigma P_\sigma$
end

The most elegant exposition of the properties of this algorithm is probably by Brualdi [3]. He bases his analysis on geometry rather than on the combinatorics used earlier.

Noting that the first columns of the reflections $\mathcal{O}_2', \mathcal{O}_3', \mathcal{O}_6'$ contain the origin of the plane, to find a $\mathbf{t}'$ hopefully appropriate for a given external U we simply would try to use in the Birkhoff-algorithm, applied to $D = |C_1(U)|^{\circ 2}$, as often as possible multiples of P_σ with σ's pertaining to the $\mathcal{O}_2'$-column. This might result in that the 'reflectory part' stemming from $\mathcal{O}_2'$ would outweigh the term $T(U)$ plus eventual further perturbations stemming from subtracting multiples of P_σ with σ's different from the $\mathcal{O}_2'$ group and yet necessary to complete the while-loop. To find $\mathbf{t}'', \mathbf{t}'''$ we would proceed similarly with $\mathcal{O}_3'$ and $\mathcal{O}_6'$, respectively, taking the rôles of $\mathcal{O}_2'$. This strategy worked well for $U = U_{\mathrm{Dru}}$. In each of $\mathbf{t}', \mathbf{t}'', \mathbf{t}'''$, we had only to use one permutation not belonging to the group desired. Its coefficient, always $a_3 = 0.2956\ldots$, is small enough so as not to harm our intentions.

If the enumeration of the facets of $\mathscr{P}_{2\mathrm{perm}} = \mathrm{co}\{|C_2(P_\sigma)|:\ \sigma \in S_4\}$ in [5], pp. 255–256 is complete, the discussion there shows that implementability of the above heuristics for all matrices $U \in \mathrm{SU}(4)$, for which $\tilde{D} = |C_2(U)|^{\circ 2}$ satisfies $\tilde{D}_{12} - \tilde{D}_{33} + \tilde{D}_{61} + \tilde{D}_{62} + \tilde{D}_{52} > 1$ would imply part (a) of the Main Theorem without the restriction $U = U_{\mathrm{Dru}}$.

4. Proof of part (b) of the Main Theorem

Let U be an external unitary 4×4 matrix; that is assume it satisfies $\neg$lhs(3). Using (7) we see by the hyperplane separation theorem that this is equivalent to finding a real 3×3 matrix $\mathbf{3}$ and a 2×2 matrix $\mathbf{2}$ such that for all σ, $s(3(\sigma, U) \oplus 2(\sigma, U) \circ (\mathbf{3} \oplus \mathbf{2})) > 0$. We can write $\mathbf{3} = \mathbf{3}_1 + \mathbf{3}_2 + \mathbf{3}_3$ where the $\mathbf{3}_i$ have rank 1. Define matrices

$$K = 0_{5\times 5} \oplus \bigoplus_{j=1}^{3} \mathbf{3}_j \oplus \mathbf{2}, \qquad J = 0_{5\times 5} \oplus \bigoplus_{j=1}^{3} J_{3\times 3} \oplus J_{2\times 2}.$$

K is a direct sum of matrices of rank $\leqslant 2$. Thus by [10], 0.4.6 and p. 457, we can write $K = \oplus_{i=0}^{4}(a_i b_i^{\mathrm{T}} + \tilde{a}_i \tilde{b}_i^{\mathrm{T}})$ with columns $a_i, b_i, \tilde{a}_i, \tilde{b}_i$ of lengths $5, 3, 3, 3, 2,$

respectively, for $i = 0, 1, 2, 3, 4$. Let $a = [a_0^T, \ldots, a_4^T]$. Define similarly 16-rows $b, \tilde{a}, \tilde{b}$. Then $K = J \circ (a^T b + \tilde{a}^T \tilde{b})$; i.o.w. K can be embedded in a rank 2 matrix. Define complex 16-columns $x = a^T + \sqrt{-1}\tilde{a}^T$, $y = b^T + \sqrt{-1}\tilde{b}^T$ and note that $\mathrm{supp}(\mathbf{Q}^T R(\sigma, U)\mathbf{Q}) \subseteq \mathrm{Supp}\,(J)$. Thus for every σ,

$$\begin{aligned}
0 &< s(3(\sigma, U) \oplus 2(\sigma, U) \circ (\mathbf{3} \oplus \mathbf{2})) = s(\mathbf{Q}^T R(\sigma, U)\mathbf{Q} \circ K) \\
&= s((\mathbf{Q}^T R(\sigma, U)\mathbf{Q} \circ J) \circ (a^T b + \tilde{a}^T \tilde{b})) \\
&= \sum_{i,j} (\mathbf{Q}^T R(\sigma, U)\mathbf{Q})_{ij} (a^T b + \tilde{a}^T \tilde{b})_{ij}) \\
&= \Re(x^* \mathbf{Q}^T R(\sigma, U)\mathbf{Q}y).
\end{aligned}$$

This obviously implies part (b) of the Main Theorem. □

For an explicit example, the reader can check that $x^* = [0_{1\times 5}$ 2.3 0 0 0 2.3 0 0 0 2.3 −0.3−3.4i −3.4+0.2i], $y = [0_{1\times 5}$ 0 0 2.3 0 2.3 0 0 0 2.3 1 i$]^T$, yield positive real part of $x^* \mathbf{Q}^T R(\sigma, U_{\mathrm{Dru}})\mathbf{Q}y$ for all $\sigma \in S_4$. The construction of these vectors was made using numerical results as given in the next section and applying the ideas above.

5. Notes on numerical experiments and decidability of OMC for $n = 4$

We report on numerical experiments documented in [15] (we suspect from his papers, that Drury [6] has similar results) and give a suggestion of how OMC for $n = 4$ can possibly be resolved with current computing facilities.

Recall that by (7) $R(\sigma, U)$ is subjected to a rotation. Similarly the polytopes $\mathscr{P}_{2\mathrm{perm}}$, $\mathscr{P}_{O_{\mathrm{perm}}} = \mathrm{co}\{O(\sigma) \oplus O'(\sigma)\colon\ \sigma \in S_4\}$, $\mathscr{P}_U = \mathrm{co}\{3(\sigma, U) \oplus 2(\sigma, U)\colon\ \sigma \in S_4\}$ are congruent ([8], Article 139 C). The latter two polytopes live in the 13-dimensional space of matrices of $(3 \times 3) \oplus (2 \times 2)$ format; in it we consider also the cones $\mathscr{C}_U = \mathrm{cone}(\mathscr{P}_U)$ and its dual $\mathscr{C}_U^* = \{X\colon$ For all $C \in \mathscr{C}_U$, $\langle X, C\rangle \leqslant 0\}$. The interior of $\mathscr{C}_U^*$ comprises (up to sign) exactly the matrices (normal to hyperplanes) separating O from $\mathscr{P}_U$ – see [8], Article 89F, for more.

a. A question of natural interest is how far lhs(3) can fail to be true? By Corollary 2 this question is closely related to how far the origin (zero matrix) in $M_5(\mathbb{R})$ can be from the polytope $\mathscr{P}_U$. In investigations made for $U = U_{\mathrm{Dru}}$, we found O by several accounts very near to $\mathscr{P}_{U_{\mathrm{Dru}}}$; it is this nearness in comparison with the distance that the eigenvalues of $T'^{-1}T''$ etc. (see Section 2) have from the real axis that led us to formulate the conjecture of Section 1. Indeed, the orthogonal projection of O onto $\mathscr{P}_{U_{\mathrm{Dru}}}$, thus the point of $\mathscr{P}_{U_{\mathrm{Dru}}}$ nearest to O is given by the matrix

$$\mathcal{O}' = \begin{bmatrix} 0.0366 & & \\ & 0.0366 & \\ & & 0.0366 \end{bmatrix} \oplus \begin{bmatrix} & 0.0211 \\ -0.0211 & \end{bmatrix}.$$

The Euclidean norm of this matrix is 0.0701, while the distance between any two vertices of $\mathscr{P}_{U_{\mathrm{Dru}}}$ is always one of the numbers $\sqrt{8} \approx 2.83$ or $\sqrt{12} \approx 3.46$, for these are the distances possible between any two $|C_2(P_\sigma)|$. An observer at the origin will see 13 of the 24 vertices of $\mathscr{P}_{U_{\mathrm{Dru}}}$. These correspond to permutations of ordinal numbers given in [4 8 10 11 12 13 14 17 18 19 21 23 24]. If V is any visible vertex, then one will find that $\angle(\overrightarrow{OV}, \overrightarrow{OO'}) \approx 88°$. The cone $\mathscr{C}_{U_{\mathrm{Dru}}}$ therefore is simplicial; in fact representatives of its extremerays are precisely the linearly independent vectors $\overrightarrow{OV}$, V visible. Not surprisingly from these data, a pure random search of a hyperplane separating O from $\mathscr{C}_{U_{\mathrm{Dru}}}$ – or equivalently a random search of an element (matrix) $X \in \mathscr{C}^*_{U_{\mathrm{Dru}}}$ did not succeed.

b. By the properties of the Frobenius inner product and Corollary 2 we have

$$\begin{aligned} x^* R(\sigma, U) y &= \langle R(\sigma, U), x\bar{y}^{\mathrm{T}} \rangle \\ &= \langle 0_{5\times 5} \oplus \bigoplus_{j=1}^{3} 3(\sigma, U) \oplus 2(\sigma, U), \mathbf{Q}^{\mathrm{T}} x\bar{y}^{\mathrm{T}} \mathbf{Q} \rangle. \end{aligned} \tag{11}$$

Given that OMC fails to extend to an assertion for bilinear forms we work in the sequel only with variables x_i and x_{i+4} =former y_i for $i = 1, 2, 3, 4$. The discussion in Section 1 shows that the vectors relevant for proving OMC can be taken to be $[x_{1234}, x_{234}, \ldots x_{\emptyset}]^{\mathrm{T}}$ and $[x_{\emptyset}, x_5, \ldots, x_{5678}]^{\mathrm{T}}$. If one evaluates rhs(11) one finds that there is a matrix $R(x)$ of format $(3 \times 3) \oplus (2 \times 2)$ whose entries are homogeneous polynomials of degree 4 in $x = (x_1, \ldots, x_8)$ in terms of which OMC($n = 4$) is equivalent to proving for all $x \in \mathbb{C}^8$ and all external $U \in \mathrm{SU}(4)$ that $R(x) \notin \mathrm{interior}(\mathscr{C}^*_U)$. As a measure of how far $R(x)$ fails to lie in the interior of this cone, we tried to minimize the angle between it and $\overrightarrow{OR(x)}$ by means of the MATLAB routine `fmins`. Admitting complex $x \in \mathbb{C}^8$, wherever we started we ended up somewhere between 10° and 11°, for real x the corresponding value was about 16.7°.

c. By results of Brualdi [2], Theorem 2.8, and Cruse [4] the 48-vertex, 13-dimensional polytope of all 6×6 centrosymmetric doubly stochastic matrices must contain $\mathrm{co}\{|C_2(U)|^{\circ 2}\colon U \in \mathrm{SU}(4)\}$, a set with probably infinitely many extreme points. If OMC($n = 4$) is true then some truncation of that polytope will still contain that set, yet be small enough to guarantee for all its finitely many vertices V that (*) $R(x) \notin \mathrm{interior}(\mathscr{C}(V)^*)$. Here $\mathscr{C}(V)$ is obtained by substituting in the construction leading to $\mathscr{C}_U$ the matrix $|C_2(U)|^{\circ 2}$ by V. Since for every U there will be a V such that $\mathscr{C}(V) \subseteq \mathscr{C}_U$ and since dualizing is inclusion reversing, establishing (∗) for all x and V implies OMC($n = 4$). We note that the quantities x_I satisfy many algebraic relations; $x_{3456}x_{1256} - x_{2456}x_{1356} = 0$ is one example of dozens. Separating the real and the imaginary part of the $x\ldots$, such a quadratic relation for complex numbers gives rise to two similar quadratic relations over the reals. In this way we can write down a large number of quadratic relations that real numbers have to satisfy in order to originate candidates for the entries of $R(x)$. Such a matrix is not in (the interior of) $\mathscr{C}(V)^*$, iff there are no $\lambda_i \geqslant 0$ such that $R(x) - \sum_i \lambda_i \text{ext}_i = 0_{5\times 5}$; here the ext_i designate representatives of extremal rays of $\mathscr{C}(V)^*$. Thus OMC($n=4$) would follow by showing the inconsistency of each of a finite number of systems of linear and quadratic relations in $\Re x_I$, $\Im x_I$ and λ_i under nonnegativity constraints for the λs.

Acknowledgements

We thank J.K. Merikoski for suggestions that improved the presentation.

References

[1] N. Bebiano, J.K. Merikoski, J. da Providência, On a conjecture of G.N. de Oliveira on determinants, Linear and Multilinear Algebra 20 (1987) 167–170.
[2] R.A. Brualdi, Convex polytopes of permutation invariant doublystochastic matrices, J. Comb. Theory Ser. B. 23 (1977) 58–67.
[3] R.A. Brualdi, Note on the Birkhoff-algorithm for doubly stochastic matrices, Can. Math. Bull. 25 (1982) 191–199.
[4] A.B. Cruse, Some Combinatorial properties of centrosymmetric matrices, Linear Algebra Appl. 16 (1977) 65–77.
[5] S.W. Drury, A counterexample to a question of Merikoski and Virtanen, Linear Algebra Appl. 168 (1992) 251–257.
[6] S.W. Drury, Unpublished numerical results on OMC.
[7] S.W. Drury, The determinantal conjecture and Hadamard type inequalities, Linear Algebra Appl. 246 (1996) 279–297.
[8] Encyclopedic Dictionary of Mathematics, in 3 + 1 volumes, MIT Press, Cambridge, MA, 1987.
[9] M. Fiedler, Bounds for the determinant of the sum of two Hermitian matrices, Proc. Amer. Math. Soc. 30 (1971) 27–31.
[10] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, New York, 1990.
[11] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, New York, 1991.
[12] R.A. Horn, The Hadamard Product, in: Matrix Theory and Applications, Proceedings of Symposia in Applied Mathematics, vol. 40, American Mathematical Society, Providence, RI, 1990.
[13] G.D. James, The Representation Theory of the Symmetric Group, Lecture Notes in Math., vol. 682, Springer, Berlin, 1978.
[14] R. Keown, An Introduction to Group Representation Theory, Mathematics in Science and Engineering, vol. 116, Academic Press, New York, 1975.

[15] A. Kovačec, Some Numerical Investigations on the Marcus–de Oliveira Conjecture as a Statement about Cones: a documentation, preprint, 1996.
[16] M. Marcus, Derivations, Plücker relations and the numerical range, Indiana Univ. Math. Jour. 22 (1973) 1137–1149.
[17] MATLAB, Numeric Computation Software, The MathWorks, Inc., Natick, MA, 1984.
[18] J.K. Merikoski, A. Virtanen, Some notes on de Oliveira's determinantal conjecture, Lin. Alg. Appl. 121 (1989) 345–352.
[19] J.K. Merikoski, A. Virtanen, N. Bebiano, J. da Providencia, Interdependence of some problems arising in generalizing the Marcus–de Oliveira conjecture, Portugal. Math. 50 (1993) 227–236.
[20] G.N. de Oliveira, Research Problem: Normal matrices, Linear and Multilinear Algebra 12 (1982) 153–154.
[21] B.E. Sagan, The Symmetric Group, Wadsworth, CA, 1991.
[22] J. Saxl, Characters of multiply transitive permutation groups, J. Algebra 34 (1975) 528–539.
[23] G. Styan, Hadamard products and multivariate statistical analysis, Linear Algebra Appl. 6 (1973) 217–240.