On the Geometry of the Birkhoff Polytope
II. The Schatten p-norms

Ludovick Bouthat[1], Javad Mashreghi[1], Frédéric Morneau-Guérin[2*]

[1]Département de mathématiques et de statistique, Université Laval,
2325 Rue de l'Université, Québec, G1V 0A6, Qc, Canada.
[2*]Département Éducation, Université TÉLUQ, 455 Rue du Parvis,
Québec, G1K 9H6, Qc, Canada.

*Corresponding author(s). E-mail(s): frederic.morneau-guerin@teluq.ca;
Contributing authors: ludovick.bouthat.1@ulaval.ca;
javad.mashreghi@mat.ulaval.ca;

Abstract

In the first of this series of two articles, we studied some geometrical aspects of the Birkhoff polytope, the compact convex set of all $n \times n$ doubly stochastic matrices, namely the Chebyshev center, and the Chebyshev radius of the Birkhoff polytope associated with metrics induced by the operator norms from ℓ_n^p to ℓ_n^p for $1 \leq p \leq \infty$. In the present paper, we take another look at those very questions, but for a different family of matrix norms, namely the Schatten p-norms, for $1 \leq p < \infty$. While studying these properties, the intrinsic connection to the minimal trace, which naturally appears in the assignment problem, is also established.

Keywords: Doubly stochastic matrices, Birkhoff polytope, Chebyshev center, Chebyshev radius, Schatten p-norms

MSC Classification: 15B51 46B20 52B12 47B10

1 Introduction

A square matrix is said to be *doubly stochastic* if its entries are non-negative and all its row and column sums are one. It is readily verified that the set of $n \times n$ doubly stochastic matrices is closed under matrix multiplication.

Note that a doubly stochastic matrix can be interpreted as the *adjacency matrix* of a *weighted directed graph* such that at each vertex the sum of the weights of the incoming edges as well as the sum of the weights of the outgoing edges are equal to one [1]. Through this association, doubly stochastic matrices play a key role in networked control problems (NSC) [2, 3].

The connection between such matrices and graph theory in particular, and combinatorics in general does not stop there. The celebrated Birkhoff's theorem asserts that every $n \times n$ doubly stochastic matrix can be written (though not necessarily in a unique way) as a convex combination of $n \times n$ permutation matrices, and the $n \times n$ permutation matrices are precisely the extreme points of Ω_n, the set of $n \times n$ doubly stochastic matrices often referred to as the *Birkhoff polytope*. This foundational result turns out to be logically equivalent to a long list of remarkably powerful theorem in combinatorics which includes the Kőnig–Egerváry theorem, Kőnig's theorem, Menger's theorem, Dilworth's theorem as well as Ford and Fulkerson's *max-flow min-cut* theorem [4].

In [5], the geometry of the Birkhoff polytope is studied. More specifically, the Chebyshev center and the Chebyshev radius of Ω_n relative to the operator norms from ℓ_n^p to ℓ_n^p $(1 \le p \le \infty)$ is determined. These concepts are interesting from a theoretical point of view. However, they also have received attention in some practical settings, notably in two papers from 1998 by Glunt, Hayden and Reams [6] and by Khoury [7]. In the former, while studying numerical simulation of large linear semiconductor circuit networks, the authors naturally came across the following question: *Given a matrix B of order n subject to the constraints $e_1^\intercal D^k e_1 = e_1^\intercal B^k e_1$, $k \ge 1$, where $e_1 = (1, 0, \ldots, 0)^\intercal$, which generalized doubly stochastic matrix D is* closest *to B in the sense of Frobenius?* They were able to give an algorithm to numerically find the solution. In the latter, Khoury [7] independently studied the same question, bar the constraints mentioned above. He showed that $D = WBW + J_n$, where $W = I_n - J_n$ and J_n is the $n \times n$ matrix with every entry uniformly equal to $1/n$. This line of investigation was later pursued by Glunt, Hayden and Reams in [8] and further developed in the more specific case of doubly stochastic matrices (as opposed to generalized doubly stochastic matrices) in [9] by Bai, Chu and Tan in 2007.

The present article continues the line of research pursued in [5]. Therein, in the process of determining the Chebyshev center and radius of Ω_n relative to the operator norm from ℓ_n^2 to ℓ_n^2, the case of the Frobenius norm is also addressed. A natural generalization of this norm is the *Schatten p-norms*. Hence, in this paper, we propose to also determine the Chebyshev center and the Chebyshev radius of Ω_n, but this time relative to the Schatten p-norms $(1 \le p < \infty)$. The case of the Schatten ∞-norm is omitted since it coincide with the operator norm from ℓ_n^2 to ℓ_n^2, which is already covered in [5]. In this paper, we also address the minimal bounding ball's radius for the Birkhoff polytope and the smallest enclosing ball problem.

In Section 2, we establish some preliminary results. More precisely, we recall some basic definitions in Section 2.1 and Section 2.2, we present a few elementary properties of doubly stochastic matrices in Section 2.3, we state some basic facts about the minimal trace of a matrix in 2.4, and we recall some general results on the Chebyshev radius proven in [5]. In Section 3, we determine the minimum and maximum distance of an element of the Birkhoff polytope from the origin. In Section 4, we study the

minimal bounding ball of the Birkhoff polytope. Among other things, we find an explicit formula for the radius of the smallest enclosing ball of Ω_n centered at D in the case of the Frobenius norm, for any doubly stochastic matrix D. In doing so, interesting connections with the assignment problem are highlighted. Finally, in Section 5, we study the Chebyshev center and the Chebyshev radius of Ω_n relative to the Schatten p-norms.

2 Definitions, Properties and Preliminary results

2.1 Basic definitions

In this section, we recall the definition of some basic notions which are essential in this paper. Secondly, we state some well-known properties of doubly stochastic matrices. Thirdly, we define the minimal trace of a $n \times n$ matrix and discuss some of its elementary properties. Finally, we recall some known results about the Chebyshev center and radius of Ω_n relative to a metric induced by a permutation-invariant norm.

Singular values

Given an $m \times n$ matrix A, the i-th singular value of A is given by

$$\sigma_i(A) = \inf_{\substack{V \subseteq \mathbb{C}^n \\ \dim(V)=n-i+1}} \sup_{\substack{x \in V \\ \|x\|_2=1}} \|Ax\|_2, \qquad (i = 1, 2, \ldots, \min\{m, n\}).$$

It is easy to see that

$$\sigma_1(A) \geq \sigma_2(A) \geq \cdots \geq \sigma_{\min\{m,n\}}(A) \geq 0.$$

Moreover, the singular values of A are also given by the square root of the non-negative eigenvalues of A^*A, where A^* denotes the conjugate transpose of A. In this case, the singular values are not necessarily ordered and thus we assume without loss of generality that they are.

Schatten norms

For $p \geq 1$, the Schatten p-norm of a $m \times n$ matrix A, denoted by $\|A\|_{\mathcal{S}_p}$, is defined as the p-norm of its singular values, that is

$$\|A\|_{\mathcal{S}_p} := \left(\sum_{i=1}^{\min\{m,n\}} \sigma_i^p(A) \right)^{1/p}$$

The case $p = 2$ yields the well-known *Frobenius norm*, which admits an easier characterization. Indeed, this latter matrix norm is induced by the inner product $\langle A, B \rangle_{\mathrm{F}} := \mathrm{tr}(AB^*)$, where A and B belong to $M_{m,n}(\mathbb{C})$. Hence, the Schatten 2-norm

of $A = [a_{ij}]_{m \times n}$ is also given by

$$\|A\|_{\mathcal{S}_2} = \|A\|_{\mathrm{F}} = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2}.$$

2.2 Useful properties of the Schatten p-norms

In what follows, we present a few well-known features of the Schatten p-norms that we shall use later.

Monotonicity

For $1 \le p \le q$,
$$\|A\|_{\mathcal{S}_1} \ge \|A\|_{\mathcal{S}_p} \ge \|A\|_{\mathcal{S}_q}.$$
This is an obvious consequence of well-known results on inclusions between ℓ^p spaces.

Permutation-invariance

Recall that any complex $m \times n$ matrix A admits a *singular value decomposition*, that is a factorization of the form
$$A = U \Sigma V^*,$$
where U is an $m \times m$ complex unitary matrix, Σ is an $m \times n$ rectangular diagonal matrix with the singular values of A on the diagonal, and V^* is the conjugate transpose of a $n \times n$ complex unitary matrix V. For all $p \ge 1$, it is easy to verify that

$$\|A\|_{\mathcal{S}_p} = \|\Sigma\|_{\mathcal{S}_p}.$$

Similarly, the Schatten p-norm of any matrix having the same singular values as A is equal to $\|\Sigma\|_{\mathcal{S}_p}$. This means that the Schatten p-norms are unitarily-invariant and, *a fortiori*, permutation-invariant.

Submultiplicativity

For all $p \ge 1$, the Schatten p-norm is *submultiplicative*, that is

$$\|AB\|_{\mathcal{S}_p} \le \|A\|_{\mathcal{S}_p} \|B\|_{\mathcal{S}_p}$$

for all $p \ge 1$ and for every $m \times n$ matrix A and every $n \times k$ matrix B.

To demonstrate this elementary fact, it suffices to show that it holds true for rectangular diagonal matrices Σ_A and Σ_B having for diagonal entries the singular values as A and B, respectively, in decreasing order of modulus. Under these assumptions, we have

$$\|\Sigma_{AB}\|_{\mathcal{S}_p} = \|\Sigma_A \Sigma_B\|_{\mathcal{S}_p}$$
$$= \left(\sum_{i=1}^{n} (\sigma_i(A)\sigma_i(B))^p \right)^{1/p}$$

$$\leq \sigma_1(A) \cdot \left(\sum_{i=1}^{n} \sigma_i(B)^p\right)^{1/p}$$

$$\leq \left(\sum_{i=1}^{n} \sigma_i(A)^p\right)^{1/p} \cdot \left(\sum_{i=1}^{n} \sigma_i(B)^p\right)^{1/p}$$

$$= \|\Sigma_A\|_{\mathcal{S}_p} \|\Sigma_B\|_{\mathcal{S}_p}.$$

2.3 Elementary spectral properties of doubly stochastic matrices

As we mentioned in the first part of this series of articles, the spectrum of a doubly stochastic matrix D is bounded by 1, and it always includes the scalar 1 (associated to the eigenvector $e = (1, 1, \ldots, 1)^{\intercal}$). As for the other eigenvalues, they have modulus 1 if and only if D is a permutation matrix [10, Theorem 5].

Let us state some other useful results relating to doubly stochastic matrices which will be put to use in our studies.

(i) A convex real-valued function on Ω_n attains its maximum at a permutation matrix [11, Corollary 8.7.4].

(ii) $DJ_n = J_nD = J_n$ for every $n \times n$ doubly stochastic matrix D, where J_n is the $n \times n$ matrix where every entry is equal to $1/n$.

(iii) J_n is the uniform convex combination of all the $n \times n$ permutation matrices, i.e., $J_n = \frac{1}{n!}\sum_{P \in \mathcal{P}_n} P$ [5].

2.4 The minimal trace of a matrix

The *minimal trace* of an $n \times n$ matrix A is

$$\operatorname{tr}_{\min}(A) := \min_{P \in \mathcal{P}_n} \operatorname{tr}(AP).$$

Alternatively, it can be defined as the minimal diagonal sum of A, i.e.,

$$\operatorname{tr}_{\min}(A) = \min_{\sigma \in \operatorname{Sym}(n)} \sum_{i=1}^{n} a_{i\sigma(i)},$$

where $\operatorname{Sym}(n)$ denotes the full symmetric group of degree n, that is the group whose elements are the bijections from the set $\{1, 2, \ldots, n\}$ onto itself.

The function $\operatorname{tr}_{\min} : M_n(\mathbb{R}) \to \mathbb{R}$, as well as its restriction to Ω_n, were extensively studied by Wang [12]. It is also closely related to the *assignment problem*, which is a fundamental combinatorial optimization problem [13]. Indeed, finding the value of $\operatorname{tr}_{\min}(A)$ is equivalent to solving an assignment problem associated with the matrix A. This, in turn, can be done with the $O(n^3)$ time algorithm known as the *Hungarian algorithm* [14].

We now state a result, due to Wang [12, Proposition 1.1], which provides interesting upper and lower bounds for the minimal trace of a doubly stochastic matrix. We also provide a complete and new elementary proof.

Lemma 1. *The minimal trace of a doubly stochastic matrix lies on the closed unit interval, i.e., $0 \leqslant \mathrm{tr}_{\min}(D) \leqslant 1$ for all $D \in \Omega_n$. Moreover, both these bounds are sharp. Indeed,*

(i) $\mathrm{tr}_{\min}(D) = 0$ if and only if D has a diagonal consisting entirely of zeroes;

(ii) $\mathrm{tr}_{\min}(D) = 1$ if and only if $D = J_n$.

Proof. The lower bound is an immediate consequence of the non-negativity of all the entries of D. As for the necessary and sufficient condition for equality to hold, it is trivial.

Let us now turn our attention towards the upper bound. Let P be a permutation matrix for which the minimal trace of D is realized and let T_{ij} be the permutation matrix which swap the i-th and j-th columns with respect to *right* multiplication. Observe that if $DP = [a_{ij}]$, the minimality of P implies that

$$\sum_{k=1}^{n} a_{kk} = \mathrm{tr}(DP) \leq \mathrm{tr}(DPT_{ij}) = a_{ij} + a_{ji} + \sum_{\substack{k=1 \\ k \neq i,j}}^{n} a_{kk}$$

for each $1 \leq i, j \leq n$, $i \neq j$. It follows that

$$a_{ii} + a_{jj} = \sum_{k=1}^{n} a_{kk} - \sum_{\substack{k=1 \\ k \neq i,j}}^{n} a_{kk} \leq a_{ij} + a_{ji},$$

which trivially holds even if $i = j$. By summing over all i and j we get, on the one hand,

$$\sum_{i,j=1}^{n} (a_{ii} + a_{jj}) = n \sum_{i=1}^{n} a_{ii} + n \sum_{j=1}^{n} a_{jj} = 2n \, \mathrm{tr}(DP) = 2n \, \mathrm{tr}_{\min}(D),$$

and, on the other hand,

$$\sum_{i,j=1}^{n} (a_{ij} + a_{ji}) = 2 \sum_{i,j=1}^{n} a_{ij} = 2n.$$

Therefore,

$$2n \, \mathrm{tr}_{\min}(D) = \sum_{i,j=1}^{n} (a_{ii} + a_{jj}) \leq \sum_{i,j=1}^{n} (a_{ij} + a_{ji}) = 2n. \tag{1}$$

Hence $\mathrm{tr}_{\min}(D) \leq 1$.

It only remains to discuss the conditions under which this upper bound is attained. One can easily check, by direct verification, that $\mathrm{tr}_{\min}(J_n) = 1$. To complete the proof, suppose that D is a doubly stochastic matrix D such that $\mathrm{tr}_{\min}(D) = 1$ and assume that the minimal trace of D is realized by $P \in \mathcal{P}_n$. Set $DP = [a_{ij}]$. Under

our assumption that $\mathrm{tr}_{\min}(D) = 1$, equality must hold in (1). But this happens if and only if

$$a_{ii} + a_{jj} \;=\; a_{ij} + a_{ji} \tag{2}$$

for all $1 \le i, j \le n$. Summing up over all i, we obtain

$$\sum_{i=1}^{n}(a_{ii} + a_{jj}) \;=\; \mathrm{tr}_{\min}(D) + na_{jj} \;=\; 1 + na_{jj}.$$

on the left-hand side, and

$$\sum_{i=1}^{n}(a_{ij} + a_{ji}) \;=\; 1 + 1 \;=\; 2$$

on the right-hand side. Hence $a_{jj} = 1/n$ for all $1 \le j \le n$.

Now, let $\sigma \in \mathrm{Sym}(n)$ be an arbitrary permutation of $\{1, 2, \dots, n\}$. Then, by (2),

$$2\,\mathrm{tr}_{\min}(D) = \sum_{i=1}^{n}(a_{ii} + a_{\sigma_i \sigma_i}) = \sum_{i=1}^{n}(a_{i\sigma_i} + a_{\sigma_i i}) = \mathrm{tr}(DQ) + \mathrm{tr}(Q^*D) \ge 2\,\mathrm{tr}_{\min}(D),$$

where Q is the permutation matrix corresponding to the permutation σ. It follows that $\mathrm{tr}(DQ) = \mathrm{tr}(Q^*D) = \mathrm{tr}_{\min}(D) = 1$ and thus, for each permutation σ, by the above discussion we must have $a_{i\sigma_i} = 1/n$ for all $1 \le i \le n$. Hence, $a_{ij} = 1/n$, for every $1 \le i, j \le n$. $\qquad\square$

2.5 Geometric notions

The remainder of this section is devoted to the introduction of geometric notions and theorems of metric geometry formulated in a general context which will be put to good use later. Although these notions and results were also presented in the first part of this series of articles, for the sake of completeness we thought it best to present them succinctly here too.

The Minimal Bounding Ball

Given a metric space $(\mathcal{U}, d)$ and a nonempty, closed and bounded $\mathcal{B} \subseteq \mathcal{U}$, then we say that $B(x, r)$, i.e., the closed ball with radius $r > 0$ centered at $x \in \mathcal{U}$, is a *bounding ball* of $\mathcal{B}$ if $\mathcal{B} \subseteq B(x, r)$. For a fixed $x \in U$, the smallest radius r such that $B(x, r)$ is a *bounding ball* of $\mathcal{B}$ is called the *minimal radius of a bounding ball of $\mathcal{B}$ centered at x* and is denoted by $r_d(x)$.

In this paper, the ambient space is always $M_n(\mathbb{R})$ and the nonempty, closed and bounded set $\mathcal{B}$ shall always be the Birkhoff polytope Ω_n. Moreover, the metric will also be the one induced by the Schatten p-norm. Hence, we have

$$r_{\|\cdot\|_{\mathcal{S}_p}}(A) \;=\; \sup_{B \in \Omega_n} \|A - B\|_{\mathcal{S}_p}. \tag{3}$$
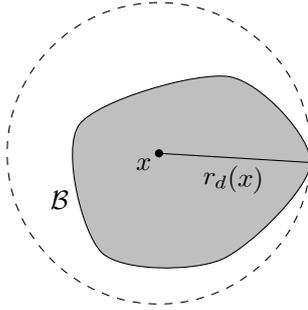
7

Fig. 1 The minimal bounding ball of the nonempty closed bounded set $\mathcal{B}$ centered at x with respect to the metric d.

As a matter of fact, the set Ω_n being compact, the supremum in (3) can be replaced by a maximum, and it follows that the minimal radius of a bounding ball of Ω_n centered at A exists and is attained for every $A \in M_n(\mathbb{R})$.

The Chebyshev radius and center

Given a metric space $(\mathcal{U}, d)$, with a nonempty closed constraint set $\mathcal{R}$ and a nonempty closed, bounded set $\mathcal{B}$, the *Smallest Enclosing Ball Problem*, a classical problem of computational geometry, consists in finding the smallest possible radius $r > 0$ and a point x in the constraint set $\mathcal{R}$ such that the ball $B(x, r)$ contains all of $\mathcal{B}$.

The smallest enclosing ball for $\mathcal{B}$ with respect to the metric space $(\mathcal{U}, d)$ and the constraint $\mathcal{R}$ may not always exist. And if it does, uniqueness is not assured. In cases where it exists, the radius r is known as the *Chebyshev radius* of $\mathcal{B}$ relative to the metric space $(\mathcal{U}, d)$ and the constraint set $\mathcal{R}$, denoted by $R_d(\mathcal{B})$. A point $x \in \mathcal{R}$ that achieves this is called a *Chebyshev center* of $\mathcal{B}$ relative to the metric space $(\mathcal{U}, d)$ and the constraint set $\mathcal{R}$. Using this notation, we have

$$R_d(\mathcal{B}) = \inf_{x \in \mathcal{R}} \sup_{y \in \mathcal{B}} d(x, y) = \inf_{x \in \mathcal{R}} r_d(x).$$

When $\mathcal{R} = \mathcal{U}$, the problem is referred to as *unconstrained*. In this setting, we focus on the metric space $(\mathcal{R}, d)$, and we define $B(x, r)$ as the minimal bounding ball of $\mathcal{B}$ in relation to the metric space $(\mathcal{R}, d)$ (and the same goes for the Chebyshev radius and centers). It is important to note that in these cases, we require $\mathcal{B} \subseteq \mathcal{R}$, unlike the general case where the constraint set $\mathcal{R}$ can be unrelated to $\mathcal{B}$.

To this day, the Smallest Enclosing Ball Problem remains an active area of research (see [15, 16] and the references therein).

Before beginning our study of the Chebyshev centers and the Chebyshev radius of the Birkhoff polytope relative to the Schatten p-norms, we need to recall some known general results.

We begin with some sufficient conditions, due to Mordukhovich, Nguyen Mau and Villalobos [16], guaranteeing the *existence* of a minimal bounding ball of $\mathcal{B}$.

Proposition 2. *[16, Theorem 1] Let $\mathcal{B}$ be a nonempty closed bounded subset in the normed vector space $(V, \|\cdot\|)$. Suppose one of the following holds:*
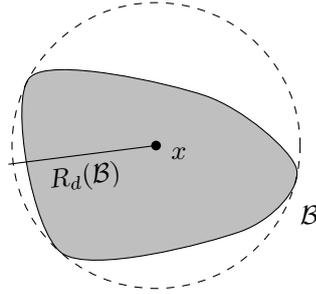
Fig. 2 The smallest enclosing ball of the nonempty closed, bounded set $\mathcal{B}$ with respect to the metric d.

   (i) *The constraint set $\mathcal{R} \subseteq V$ is nonempty and compact.*

   (ii) *The normed vector space $(V, \|\cdot\|)$ is a reflexive Banach space and the constraint set $\mathcal{R} \subseteq V$ is weakly closed.*

Then there exists a smallest enclosing ball of $\mathcal{B}$ relative to $(V, \|\cdot\|)$ and the constraint set $\mathcal{R}$.

    We continue with a sufficient condition guaranteeing the *uniqueness* of the smallest enclosing ball in the context of *strictly convex normed vector spaces*, i.e., a normed vector space where $x \neq y$ imply that

$$\|\lambda x + (1 - \lambda)y\| \;<\; \lambda\|x\| + (1 - \lambda)\|y\|$$

for all $0 < \lambda < 1$.

Proposition 3. *Let $\mathcal{B}$ be a nonempty compact subset in the strictly convex normed vector space $(V, \|\cdot\|)$ and let $\mathcal{R} \subseteq V$ be a nonempty closed, convex constraint set. If there exists a Chebyshev center of $\mathcal{B}$ relative to $(V, \|\cdot\|)$ and the constraint set $\mathcal{R}$, then it is unique.*

    Before moving on to the next result, let us recall that the set $\mathcal{K} \subset M_n$ is said to be *permutation-invariant* if $PKQ \in \mathcal{K}$ for any $K \in \mathcal{K}$ and any permutation matrices $P, Q \in \mathcal{P}_n$. Observe that the sets Ω_n and $M_n(\mathbb{R})$ are both permutation-invariant.

    For the remainder of this section, the results will concern the Chebyshev centers and radius of Ω_n in the context of a metric induced by a permutation-invariant norm and a constraint set which is also permutation-invariant.

Theorem 4. *[5, Theorem 6.4] Let $\mathcal{R} \subseteq M_n(\mathbb{R})$ be a convex permutation-invariant constraint set and let $\|\cdot\|$ be a permutation-invariant norm on $M_n(\mathbb{R})$. If there exist a Chebyshev center A of Ω_n relative to the metric space $(M_n(\mathbb{R}), \|\cdot\|)$ and the constraint set $\mathcal{R}$, then the matrix $J_n A J_n = \left(\frac{1}{n}\sum_{i,j=1}^{n} a_{ij}\right) J_n$ is also a Chebyshev center of Ω_n relative to the aforementioned metric space and constraint set. Moreover, the Chebyshev radius of Ω_n in this setting is given by*

$$R_{\|\cdot\|}(\Omega_n) \;=\; \|J_n A J_n - I_n\| \;=\; \inf_{\substack{\alpha \in \mathbb{R} \\ \alpha J_n \in \mathcal{R}}} \|\alpha J_n - I_n\|$$

and the infimum is attained by $\alpha = \frac{1}{n}\sum_{i,j=1}^{n} a_{ij}$.

9

As a direct consequence of Theorem 4, we also have the following corollary which states that, under mild conditions, any Chebyshev center of Ω_n is equidistant to every permutation matrix.

Corollary 5. *[5, Corollary 6.5] Let $\mathcal{R} \subseteq M_n(\mathbb{R})$ be a convex permutation-invariant constraint set and let $\|\cdot\|$ be a permutation-invariant norm on $M_n(\mathbb{R})$. If there exist a Chebyshev center A of Ω_n relative to the metric space $(M_n(\mathbb{R}), \|\cdot\|)$ and the constraint set $\mathcal{R}$, then $\|A - P\| = R_{\|\cdot\|}(\Omega_n)$ for any permutation matrix $P \in \mathcal{P}_n$.*

The above results are valid even if $\mathcal{R}$ has no direct relation to Ω_n. If we also make the natural assumption that $\mathcal{R} = \Omega_n$, then we obtain the following concrete result.

Corollary 6. *[5, Corollary 6.6] If $\|\cdot\|$ is a permutation-invariant norm on $M_n(\mathbb{R})$, then the special doubly stochastic matrix J_n is a Chebyshev center of Ω_n relative to the metric space $(\Omega_n, \|\cdot\|)$. Moreover, the associated Chebyshev radius is given by $R_{\|\cdot\|}(\Omega_n) = \|J_n - I_n\|$.*

3 The minimum and maximum distance of an element of the Birkhoff polytope from the origin

We begin our discussion by presenting a useful result on the range of the Schatten p-norms $\|\cdot\|_{\mathcal{S}_p}$ ($p \geq 1$) when its entry runs through the Birkhoff polytope Ω_n. Finding lower and upper bounds for this norms (as well as identifying the conditions under which these are optimal) will prove valuable later on.

Proposition 7. *Given $D \in \Omega_n$ and $p \geq 1$, then $1 \leqslant \|D\|_{\mathcal{S}_p} \leqslant n^{1/p}$ with $\|D\|_{\mathcal{S}_p} = 1$ if and only if $D = J_n$ and $\|D\|_{\mathcal{S}_p} = n^{1/p}$ if and only if D is a permutation matrix.*

Proof. Let $D \in \Omega_n$. The matrix D^*D being real symmetric and positive semi-definite, its eigenvalues are non-negative real numbers. Furthermore, D^*D being doubly stochastic, its eigenvalues are in absolute value smaller or equal to 1. Hence the singular values of D lie in the interval $[0, 1]$. Therefore,

$$1 \leq \sum_{i=1}^{n} \sigma_i^p(D) \leq n, \tag{4}$$

and thus $1 \leq \|D\|_{\mathcal{S}_p} \leq n^{1/p}$.

Since 1 is always an eigenvalue of the doubly stochastic matrix D^*D, it follows that $\sigma_1(D) = 1$. So, if $1 = \sum_{i=1}^{n} \sigma_i^p(D) = \|D\|_{\mathcal{S}_p}^p$, then $\sigma_j(D) = 0$ for $2 \leq j \leq n$ and thus $\text{rank}(D) = 1$. But M. Marcus showed that J_n is the only doubly stochastic matrix of rank 1 [17, Lemma 2]. Consequently $D = J_n$.

Now, if $n = \sum_{i=1}^{n} \sigma_i^p(D) = \|D\|_{\mathcal{S}_p}^p$, then $\sigma_i(D) = 1$ for every $1 \leq i \leq n$. Hence, the singular value decomposition of D is given by $D = UI_nV^* = UV^*$, where U and V are unitary matrices. Thus,

$$D^*D = (UV^*)^*(UV^*) = VU^*UV^* = VV^* = I_n.$$

So D is a unitary matrix and its eigenvalues must then all lie on the unit circle. Hence, it follows the first observation made in Section 2.3 that D is a permutation

matrix. Finally, since $\|P\|_{\mathcal{S}_p} = n^{1/p}$ for every $n \times n$ permutation matrix, the conclusion follows. $\qquad\square$

Note that the above result is not original. Wang, for instance, mentions it in part without proof [12].

4 The Minimal Bounding Ball of the Birkhoff polytope

We now turn our attention to the problem of characterizing the minimal radius of an enclosing ball of Ω_n centered at $A \in M_n(\mathbb{R})$ when the ambient space is equipped with the Schatten p-norms. In some cases, we will make the additional assumption that the matrix A is doubly stochastic; this loss of generality will be compensated by some considerably stronger results.

Finding an explicit formula for $r_{\mathcal{S}_p}(x)$ for a general $p \geq 1$ turns out to be a difficult problem. We thus restrict our attention to the case of the Frobenius norm (i.e., the Schatten 2-norm), where the structure arising from the underlying inner product $\langle A, B \rangle_{\mathcal{S}_2} := \operatorname{tr}(AB^*)$ allows us to derive some interesting formulas. In particular, we are able to characterize the minimal radius of a bounding ball of Ω_n centered at A in terms of, among other things, the minimal trace $\operatorname{tr}_{\min}(A)$. These formulas will be instrumental in an upcoming paper to better understand the diameter of Ω_n in the case of the Frobenius norm.

Theorem 8. *The minimal radius of a bounding ball of Ω_n centered at $A \in M_n(\mathbb{R})$ relative to the Schatten 2-norm, denoted by $r_{\mathcal{S}_2}(A)$, is*

$$r_{\mathcal{S}_2}(A) = \left(\|A\|_{\mathcal{S}_2}^2 + n - 2\operatorname{tr}_{\min}(A) \right)^{1/2}.$$

Proof. We know from property (i) of Ω_n (see Section 2.3) that the maximum of a convex real-valued function on Ω_n is attained by a permutation matrix. Hence, since $\|A - D\|_{\mathcal{S}_2}$ is a convex function with respect to the variable $D \in \Omega_n$ for any fixed $A \in M_n(\mathbb{R})$, we have $r_{\mathcal{S}_2}(A) = \max_{D \in \Omega_n} \|A - D\|_{\mathcal{S}_2} = \max_{P \in \mathcal{P}_n} \|A - P\|_{\mathcal{S}_2}$. Moreover, if $P \in \mathcal{P}_n$, Proposition 7 ensures us that

$$
\begin{aligned}
\|A - P\|_{\mathcal{S}_2}^2 &= \langle A - P,\, A - P \rangle_{\mathcal{S}_2} \\
&= \|A\|_{\mathcal{S}_2}^2 + \|P\|_{\mathcal{S}_2}^2 - \langle A, P \rangle_{\mathcal{S}_2} - \langle P, A \rangle_{\mathcal{S}_2} \\
&= \|A\|_{\mathcal{S}_2}^2 + \|P\|_{\mathcal{S}_2}^2 - \operatorname{tr}(AP^*) - \operatorname{tr}(PA^*) \\
&= \|A\|_{\mathcal{S}_2}^2 + n - 2\operatorname{tr}(AP^*).
\end{aligned}
$$

It follows that

$$r_{\mathcal{S}_2}^2(A) = \max_{P \in \mathcal{P}_n} \left(\|A\|_{\mathcal{S}_2}^2 + n - 2\operatorname{tr}(AP^*) \right) = \|A\|_{\mathcal{S}_2}^2 + n - 2\operatorname{tr}_{\min}(A).$$

Taking the square root yield the result we sought. $\qquad\square$

It is well known that there exist a connection between the space Ω_n and the assignment problem [12, 18, 19]. However, to the best of our knowledge, none of these links arose naturally by studying Ω_n. These papers were mostly motivated by the assignment problem and the authors found that by exploiting some of the properties of Ω_n (i.e., the assignment polytope), they could gain some insight into their original questions. Here, we uncover a new geometric connection between these two objects of study.

However precise it may be, Theorem 8 is not practical since the minimal trace $\mathrm{tr}_{\min}(A)$ is relatively hard to compute for large n. If we make the additional assumption that $A \in \Omega_n$, we can derive a more workable result from the estimations (i.e., $0 \leq \mathrm{tr}_{\min}(A) \leq 1$ for all $A \in \Omega_n$) given in Lemma 1.

Corollary 9. *For $D \in \Omega_n$, we have*

$$\left(\|D\|_{\mathcal{S}_2}^2 + n - 2\right)^{\frac{1}{2}} \leqslant r_{\mathcal{S}_2}(D) \leqslant \left(\|D\|_{\mathcal{S}_2}^2 + n\right)^{\frac{1}{2}}. \tag{5}$$

In particular, the ratio $r_{\mathcal{S}_2}^2(D)/\left(\|D\|_{\mathcal{S}_2}^2 + n\right)$ converges uniformly to 1 as $n \to \infty$.

Proof. The lower and upper bounds are obtained directly from Theorem 8 and Lemma 1. For the second part of the proof, remark that the ratio $r_{\mathcal{S}_2}^2(D)/\left(\|D\|_{\mathcal{S}_2}^2 + n\right)$ converges uniformly to 1 as $n \to \infty$ if and only if the ratio $r_{\mathcal{S}_2}^2(D)/\left(\|D\|_{\mathcal{S}_2}^2 + n - 1\right)$ also does. Since the computations in the second one are easier, we instead prove the latter. By squaring (5) and dividing by $\|D\|_{\mathcal{S}_2}^2 + n - 1$, we get

$$1 - \frac{1}{\|D\|_{\mathcal{S}_2}^2 + n - 1} \leq \frac{r_{\mathcal{S}_2}^2(D)}{\|D\|_{\mathcal{S}_2}^2 + n - 1} \leq 1 + \frac{1}{\|D\|_{\mathcal{S}_2}^2 + n - 1}.$$

That is,

$$\left| \frac{r_{\mathcal{S}_2}^2(D)}{\|D\|_{\mathcal{S}_2}^2 + n - 1} - 1 \right| \leq \frac{1}{\|D\|_{\mathcal{S}_2}^2 + n - 1}.$$

Hence, since $\|D\|_{\mathcal{S}_2} \geq 1$ by Proposition 7, we have

$$\left| \frac{r_{\mathcal{S}_2}^2(D)}{\|D\|_{\mathcal{S}_2}^2 + n - 1} - 1 \right| \leq \frac{1}{n}$$

and the conclusion follows. $\qquad\square$

The full force of the inequalities stated will prove instrumental in Section 5.

Note in closing that, using the estimates for $\|\cdot\|_{\mathcal{S}_p}$ as D runs through Ω_n given in Proposition 7, one can deduce upper and lower bounds:

$$n - 1 \leq r_{\mathcal{S}_2}^2(D) \leq 2n.$$

While these bounds are less precise than (5), they are simpler and somehow more elegant.

5 The Chebyshev radius and center of the Birkhoff polytope

We now seek to determine the Chebyshev radius $R_{\mathcal{S}_p}(\Omega_n)$ and the Chebyshev centers of Ω_n relative to the metric space $(M_n(\mathbb{R}), \|\cdot\|_{\mathcal{S}_p})$ and $(\Omega_n, \|\cdot\|_{\mathcal{S}_p})$. In fact, we shall see in due course that the latter case follows almost directly from the first.

Observe that Ω_n is compact and that the metric space $(M_n(\mathbb{R}), \|\cdot\|_{\mathcal{S}_p})$ is a reflexive Banach space for which $M_n(\mathbb{R})$ is obviously (weakly) closed on itself. Hence, Proposition 2 guarantees the existence of a Chebyshev center in both of the aforementioned metric spaces. Moreover, since the Schatten p-norms are permutation-invariant and since $M_n(\mathbb{R})$ is convex, Theorem 4 tells us that the matrix αJ_n is a Chebyshev center for some $\alpha \in \mathbb{R}$ and that the Chebyshev radius is given by $\inf_{\alpha \in \mathbb{R}} \|\alpha J_n - I_n\|_{\mathcal{S}_p}$. Using these properties, we derive the following result.

Theorem 10. *For $1 < p < \infty$, the special doubly stochastic matrix J_n is the unique Chebyshev center of Ω_n relative to the metric space $(M_n(\mathbb{R}), \|\cdot\|_{\mathcal{S}_p})$ and the associated Chebyshev radius is equal to $(n-1)^{1/p}$.*

Proof. We know that, for any $1 < p < \infty$, the Chebyshev radius of Ω_n is given by $\inf_{\alpha \in \mathbb{R}} \|\alpha J_n - I_n\|_{\mathcal{S}_p}$. Now, observe that $(\alpha J_n - I_n)^*(\alpha J_n - I_n) = I_n - (2\alpha - \alpha^2)J_n$. Moreover, a direct computation shows that the eigenvalues of J_n are 1 and 0. Therefore, the singular values of $\alpha J_n - I_n$ are 1 with multiplicity $n-1$, and $\sqrt{1 - 2\alpha + \alpha^2} = |1 - \alpha|$ with multiplicity 1. Hence,

$$
\begin{aligned}
\inf_{\alpha \in \mathbb{R}} \|\alpha J_n - I_n\|_{\mathcal{S}_p} &= \inf_{\alpha \in \mathbb{R}} \left(\sum_{i=1}^n \sigma_i^p(\alpha J_n - I_n) \right)^{1/p} \\
&= \inf_{\alpha \in \mathbb{R}} (n - 1 + |1 - \alpha|^p)^{1/p} \qquad (6) \\
&= (n-1)^{1/p}.
\end{aligned}
$$

Thus, $R_{\mathcal{S}_p}(\Omega_n) = (n-1)^{1/p}$ and since the above minimum is attained when $\alpha = 1$, the matrix J_n is a Chebyshev center. Since the Schatten p-norm are strictly convex for any $1 < p < \infty$ [20], it follows from Proposition 3 that the Chebyshev center is unique. □

As a matter of fact, the proof detailed above remains largely valid for $p = 1$. Indeed, only the uniqueness argument fails as the Schatten 1-norm is not strictly convex. To establish the result for $p = 1$, we need stronger tools, including the following lemma.

Lemma 11. *Let $A \in M_n(\mathbb{R})$. If A commutes with every $n \times n$ permutation matrices, then there exist some $a, b \in \mathbb{R}$ such that $A = aI_n + bJ_n$. Moreover, if $A \in \Omega_n$, then $a, b \geq 0$ and $a + b = 1$.*

Proof. Since $AP = PA$ for every permutation matrices, we have

$$
a_{i\sigma(j)} = (AP)_{ij} = (PA)_{ij} = a_{\sigma^{-1}(i)j}, \qquad 1 \leq i, j \leq n,
$$

where σ is a permutation associated with the permutation matrix P. First consider any permutation matrix P associated with a permutation σ satisfying $\sigma(j) = i$. Then we have

$$a_{ii} = a_{i\sigma(j)} = a_{\sigma^{-1}(i)j} = a_{jj}.$$

Since i and j were arbitrary, each entry in the diagonal of A is identical.

Now consider any permutation matrix P associated with a permutation σ satisfying $\sigma(j) = k$ and $\sigma(l) = i$, where $k \neq i$ and $l \neq j$. It is clear that for any $1 \leq l, k \leq n$ satisfying $k \neq i$ and $l \neq j$, such a permutation exist. We then have

$$a_{ik} = a_{i\sigma(j)} = a_{\sigma^{-1}(i)j} = a_{lj}.$$

Since $k \neq i$ and $l \neq j$, we find that every off-diagonal entry of A is identical. Hence, A has the general form

$$A = \begin{bmatrix} \alpha & \beta & \cdots & \beta \\ \beta & \alpha & \cdots & \beta \\ \vdots & \vdots & \ddots & \vdots \\ \beta & \beta & \cdots & \alpha \end{bmatrix},$$

which can be written as $aI_n + bJ_n$, where $b = n\beta$ and $a = \alpha - \beta$. If A is doubly stochastic, then $a, b \geq 0$ since the coefficients are non-negative and $a + b = 1$ since the row and column sums must be equal to 1. $\qquad\square$

Before addressing the following theorem, let us first recall von Neumann's trace inequality [21], which states that if $B, C \in M_n(\mathbb{R})$, then $|\operatorname{tr}(BC)| \leq \sum_{i=1}^n \sigma_i(B)\sigma_i(C)$, with equality if and only if there exist $n \times n$ unitary matrices U and V such that

$$B = U \operatorname{diag}(\sigma_1(B), \ldots, \sigma_n(B))V^* \quad \& \quad C = U \operatorname{diag}(\sigma_1(C), \ldots, \sigma_n(C))V^*. \quad (7)$$

Theorem 12. *J_n is the unique Chebyshev center of Ω_n relative to the metric space $(M_n(\mathbb{R}), \|\cdot\|_{\mathcal{S}_1})$ and the associated Chebyshev radius is equal to $n - 1$.*

Proof. The proof that the associated Chebyshev radius is equal to $n - 1$ is identical to the proof of Theorem 10. Therefore, we just show that J_n is the unique Chebyshev center.

Let $A \in M_n(\mathbb{R})$ be a Chebyshev center of Ω_n relative to the metric space $(M_n(\mathbb{R}), \|\cdot\|_{\mathcal{S}_1})$. By Corollary 5, we know that

$$\|A - P\|_{\mathcal{S}_1} = n - 1, \qquad P \in \mathcal{P}_n. \quad (8)$$

Moreover, von Neumann's trace inequality shows that $|\operatorname{tr}(B)| \leq \|B\|_{\mathcal{S}_1}$ for every $B \in M_n(\mathbb{R})$. Hence, we have

$$n - 1 = \|A - P^*\|_{\mathcal{S}_1} = \|AP - I_n\|_{\mathcal{S}_1} \geq |\operatorname{tr}(AP - I_n)| = |\operatorname{tr}(AP) - n|. \quad (9)$$

This means in particular that $1 - n \leq \operatorname{tr}(AP) - n$, i.e., that $1 \leq \operatorname{tr}(AP)$ for every permutation matrix P. By summing these inequalities over all permutation matrices,

we get

$$
\begin{aligned}
1 = \frac{1}{n!} \sum_{P \in \mathcal{P}_n} 1 &\leq \frac{1}{n!} \sum_{P \in \mathcal{P}_n} \operatorname{tr}(AP) = \operatorname{tr}\left(A \frac{1}{n!} \sum_{P \in \mathcal{P}_n} P\right) = \operatorname{tr}(AJ_n) \\
&= \frac{1}{n} \sum_{i,j=1}^{n} a_{ij},
\end{aligned} \tag{10}
$$

where the second to last equality is due to Lemma (iii).

Now, using once again (8), we have in particular that $n - 1 = \|A - P^*Q^*\|_{\mathcal{S}_1}$ for every $P, Q \in \mathcal{P}_n$ and it follows from the permutation-invariance of the Schatten norms that $n - 1 = \|PAQ - I_n\|_{\mathcal{S}_1}$ for any permutation matrices P and Q. Consequently, we find with another application of Lemma (iii) that

$$
\begin{aligned}
n - 1 &= \frac{1}{(n!)^2} \sum_{P,Q \in \mathcal{P}_n} \|PAQ - I_n\|_{\mathcal{S}_1} \geq \left\|\frac{1}{(n!)^2} \sum_{P,Q \in \mathcal{P}_n} (PAQ - I_n)\right\|_{\mathcal{S}_1} \\
&= \left\|\frac{1}{(n!)^2} \sum_{P,Q \in \mathcal{P}_n} PAQ - I_n\right\|_{\mathcal{S}_1} = \|J_n A J_n - I_n\|_{\mathcal{S}_1} = \|\alpha J_n - I_n\|_{\mathcal{S}_1} \\
&= n - 1 + |1 - \alpha|,
\end{aligned}
$$

where $\alpha := \frac{1}{n} \sum_{i,j=1}^{n} a_{ij}$ and the last equality is due to (6). Thus, we have $|1 - \alpha| \leq 0$ and it follows that $\frac{1}{n} \sum_{i,j=1}^{n} a_{ij} = \alpha = 1$. Therefore, the inequality in (10) is in fact an equality and this happens if and only if

$$
\operatorname{tr}(AP) = 1, \qquad P \in \mathcal{P}_n. \tag{11}
$$

Moreover, it also follows that von Neumann's trace inequality in (9) is saturated. Hence, by (7) (with $B = AP - I_n$ and $C = I_n$), it follows that there exist unitary matrices U and V such that $AP - I_n = U \operatorname{diag}(\sigma_1(AP - I_n), \ldots, \sigma_n(AP - I_n))V^*$ and $I_n = UV^*$. The latter implies that $U = V$ and thus that $AP - I_n$ is unitarily diagonalizable.

In particular, this means that AP is also unitarily diagonalizable. Hence, AP is normal for any permutation matrix P and it follows that

$$
P^*A^*AP = APP^*A^* = AA^* = A^*A.
$$

In other words, the matrix A^*A commutes with P for any permutation matrix P. It then follows from Lemma 11 that $A^*A = aI_n + bJ_n$ for some real numbers a and b. Moreover, since A^*A is positive semidefinite, the eigenvalues of $A^*A = aI_n + bJ_n$ are real and non-negative. Since its eigenvalues are $a + b$ (simple) and a (multiplicity of $n - 1$), we must have $a \geq 0$ and $b \geq -a$.

We know from the unicity of the decomposition of positive definite matrices that if B is a matrix such that $B^*B = aI_n + bJ_n$, then A must be of the form $A = UB$, where U is some unitary matrix. It is then a matter of simple computation to verify that $B = \sqrt{a}I_n + (\sqrt{a+b} - \sqrt{a})J_n =: \alpha I_n + \beta J_n$ is such a matrix and thus that $A = U(\alpha I_n + \beta J_n)$.

15

Now, use once again the fact that $A^*A = AA^*$ to obtain

$$\begin{aligned}
aI_n + bJ_n &= (\alpha I_n + \beta J_n)(\alpha I_n + \beta J_n) = (\alpha I_n + \beta J_n)U^*U(\alpha I_n + \beta J_n) = A^*A \\
&= AA^* = U(\alpha I_n + \beta J_n)(\alpha I_n + \beta J_n)U^* = U(aI_n + bJ_n)U^* \\
&= aI_n + bUJ_nU^*,
\end{aligned}$$

and thus that $J_n = UJ_nU^*$. Denote by r_i the i-th row sum of U and by c_i the i-th column sum of U. From $J_n = UJ_nU^*$, a direct computation reveal that

$$\frac{1}{n} = (J_n)_{ij} = (UJ_nU^*)_{ij} = \frac{r_ir_j}{n}, \qquad 1 \leq i, j \leq n.$$

In particular, $|r_i|^2 = 1$ and thus, $r_i = \pm 1$ for any $i \in \{1, 2, \ldots, n\}$. Suppose that $r_1 = 1$. Then we also have $\frac{r_j}{n} = \frac{r_1r_j}{n} = \frac{1}{n}$ and thus, $r_j = 1$ for every $j \in \{1, 2, \ldots, n\}$. Similarly, if $r_1 = -1$ then $r_j = -1$ for every $j \in \{1, 2, \ldots, n\}$.

Moreover, from $J_n = UJ_nU^*$ we also find that $J_nU = UJ_n$ which is equivalent to having

$$\frac{r_j}{n} = (J_nU)_{ij} = (UJ_n)_{ij} = \frac{c_i}{n}$$

and thus, $r_j = c_i$ for every $i, j \in \{1, 2, \ldots, n\}$. Consequently, every row and column of U sum to either 1 or -1. Suppose that it sum to 1. Then we know from [22] that U must be of the form $U = \sum_m c_mP_m$, where P_m are permutation matrices and c_m are complex numbers such that $\sum_m c_m = 1$ and $\sum_m |c_m|^2 = 1$. If the rows and columns of U sum to -1, then U is the form $U = -\sum_m c_mP_m$.

From this, it follows that

$$A = U(\alpha I_n + \beta J_n) = \alpha U \pm \beta J_n,$$

since $UJ_n = \pm \sum_m c_mP_mJ_n = \pm \sum_m c_mJ_n = \pm J_n$. Now, recall that we have $\frac{1}{n}\sum_{i,j=1}^n a_{ij} = 1$ and thus

$$1 = \frac{1}{n}\sum_{i,j=1}^n a_{ij} = \frac{1}{n}\alpha\sum_{i,j=1}^n u_{ij} \pm \frac{1}{n}\beta\sum_{i,j=1}^n \frac{1}{n} = \pm(\alpha + \beta),$$

where the last equality stem from the fact that each row of U sum to ± 1. Since $\alpha = \sqrt{a} \geq 0$ and $\beta = \sqrt{a+b} - \sqrt{a} \geq 0$, it follows that the negative case can be ruled out, i.e., that $U = \sum_m c_mP_m$ and thus that $\alpha + \beta = 1$.

Finally, recall (11) which states that $\mathrm{tr}(AP) = 1$ for every permutation matrix P. By the linearity of the trace and the fact that $\mathrm{tr}(J_nP) = \mathrm{tr}(J_n) = 1$, we find that $\mathrm{tr}(UP)$ is also constant for every permutation matrix P, say $\mathrm{tr}(UP) = k$. Therefore, since U is unitary we have

$$n = \mathrm{tr}(I_n) = \mathrm{tr}(UU^*) = \mathrm{tr}\left(U\sum_m \overline{c_m}P_m^*\right) = \sum_m \overline{c_m}\,\mathrm{tr}(UP_m^*) = k\sum_m \overline{c_m} = k.$$

16

Therefore, we must have $\operatorname{tr}(UP) = k = n$ for every permutation matrix P. However, we also have

$$
\begin{aligned}
1 = \operatorname{tr}(AP) = \operatorname{tr}((\alpha U + \beta J_n)P) &= \alpha \operatorname{tr}(UP) + \beta \operatorname{tr}(J_n) = \alpha \operatorname{tr}(UP) + \beta \\
&= \alpha \operatorname{tr}(UP) + 1 - \alpha,
\end{aligned}
$$

which means that $\alpha = \alpha \operatorname{tr}(UP)$ for every permutation matrix P. If $\alpha \neq 0$, then this implies that $1 = \operatorname{tr}(UP) = n$, a contradiction (since we can suppose without any loss of generality that $n > 1$). Therefore, α must be equal to 0 and we thus find that

$$
A = \beta J_n = J_n,
$$

which is what we wanted to show. □

From the previous theorems, the particular case where the constraint set is limited to Ω_n follows almost immediately since the Chebyshev center J belong to $\Omega_n \subseteq M_n(\mathbb{R})$.

Corollary 13. *For $1 \leq p < \infty$, the special doubly stochastic matrix J_n is the unique Chebyshev center of Ω_n relative to the metric space $(\Omega_n, \|\cdot\|_{\mathcal{S}_p})$ and the associated Chebyshev radius is equal to $(n-1)^{1/p}$.*

Proof. Fix $1 \leq p < \infty$ and suppose that $D_0 \in M_n(\mathbb{R})$ is a Chebyshev center of Ω_n relative to the metric space $(M_n(\mathbb{R}), \|\cdot\|_{\mathcal{S}_p})$. If $D_0 \in \Omega_n$, then we have

$$
\begin{aligned}
(n-1)^{\frac{1}{p}} = \inf_{A \in M_n(\mathbb{R})} \sup_{P \in \mathcal{P}_n} \|A - P\|_{\mathcal{S}_p} &\leq \inf_{D \in \Omega_n} \sup_{P \in \mathcal{P}_n} \|D - P\|_{\mathcal{S}_p} \\
&\leq \sup_{P \in \mathcal{P}_n} \|D_0 - P\|_{\mathcal{S}_p} = (n-1)^{\frac{1}{p}}.
\end{aligned}
$$

Therefore, $R_{\mathcal{S}_p}(\Omega_n) = (n-1)^{1/p}$. By Theorems 10 and 12, $J_n \in \Omega_n$ is the only $n \times n$ real matrix satisfying $\sup_{P \in \mathcal{P}_n} \|D - P\|_{\mathcal{S}_p} = (n-1)^{1/p}$ and in particular, it is the only doubly stochastic matrix satisfying this equation. Consequently, J_n is the unique Chebyshev center of Ω_n relative to the metric space $(\Omega_n, \|\cdot\|_{\mathcal{S}_p})$. □

Remark 1. *More generally, suppose that $\mathcal{B}$ be a nonempty closed bounded subset in the metric space $(V, \|\cdot\|)$, and let $\mathcal{R}_1, \mathcal{R}_2 \subseteq V$ be two nonempty closed constraint set. If $\mathcal{R}_1 \subseteq \mathcal{R}_2$ and A is the unique Chebyshev center of $\mathcal{B}$ relative to the metric space $(V, \|\cdot\|)$ and the constraint set $\mathcal{R}_2$, then A is also the unique Chebyshev center of $\mathcal{B}$ relative to the metric space $(V, \|\cdot\|)$ and the constraint set $\mathcal{R}_1$. Moreover, the Chebyshev radius is equal in both settings.*

Remark 2. *The matrix J_n is not the Chebyshev center of Ω_n relative to the metric space $(M_n(\mathbb{R}), \|\cdot\|)$ for every matrix norm $\|\cdot\|$. For example, consider the entrywise matrix norm $\|A\|_1 = \sum_{i,j=1}^n |a_{ij}|$. According to Theorem 4, the Chebyshev radius is*

$$
\begin{aligned}
R_{\|\cdot\|_1}(\Omega_n) = \inf_{\alpha \in \mathbb{R}} \|\alpha J_n - I_n\|_1 &= \inf_{\alpha \in \mathbb{R}} \left(n \left| \frac{\alpha}{n} - 1 \right| + (n^2 - n) \left| \frac{\alpha}{n} \right| \right) \\
&= \inf_{\alpha \in \mathbb{R}} (|\alpha - n| + (n-1)|\alpha|)
\end{aligned}
$$

which is attained uniquely *at $\alpha = 0$. Thus, the matrix 0 is a Chebyshev center of Ω_n relative to the metric space $(M_n(\mathbb{R}), \|\cdot\|_1)$, while J_n is not.*

6 Concluding Remarks

To conclude this paper, we state a few remarks and set out some open questions that appear to be of interest.

(i) Can we find formulas for the minimal radius of a bounding ball of Ω_n centered at $D \in \Omega_n$ relative to the Schatten p-norm for $p \neq 2$, and in particular for $p = 1$ and $p = \infty$?

(ii) The problem of determining $\mathrm{tr}_{\min}(D)$ is an important one in the area of combinatorial optimization. Is it possible to find an efficient algorithm to compute $\mathrm{tr}_{\min}(D)$ for $D \in \Omega_n$ by using the fact that $\mathrm{tr}_{\min}(D) = (\|D\|_{\mathcal{S}_2}^2 + n - r_{\mathcal{S}_2}^2(D))/2$?

(iii) A theorem of von Neumann asserts that every unitarily invariant matrix norm (i.e., a matrix norm such that $\|UAV\| = \|A\|$ for any unitary matrix U and V) is a symmetric gauge function applied to the singular values [11, Thm. 7.4.7.2]. Can the results of the present paper be extended to this family of norms?

Declaration of Conflict of Interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

Acknowledgement

References

[1] Gharesifard, B., Cortés, J.: When does a digraph admit a doubly stochastic adjacency matrix? In: Proceedings of the 2010 American Control Conference, pp. 2440–2445 (2010). IEEE

[2] Bullo, F., Cortés, J., Martínez, S.: Distributed Control of Robotic Networks. Princeton Series in Applied Mathematics, p. 320. Princeton University Press, Princeton (2009). https://doi.org/10.1515/9781400831470

[3] Ren, W., Beard, R.W.: Distributed Consensus in Multi-vehicle Cooperative Control, p. 319. Springer, London (2008). https://doi.org/10.1007/978-1-84800-015-5

[4] Reichmeider, P.F.: The Equivalence of Some Combinatorial Matching Theorems, p. 123. Polygonal Publ. House, Washington (1984)

[5] Bouthat, L., Mashreghi, J., Morneau-Guérin, F.: On the Geometry of the Birkhoff Polytope. I. The operator ℓ^p-norms. Acta Scientiarum Mathematicarum (2024, Submitted)

[6] Glunt, W., Hayden, T.L., Reams, R.: The nearest 'doubly stochastic' matrix to a real matrix with the same first moment. Numer. Linear Algebra Appl. **5**(6), 475–4821999 (1998) https://doi.org/10.1002/(SICI)1099-1506(199811/12)5:6⟨475::AID-NLA155⟩3.3.CO;2-X

[7] Khoury, R.N.: Closest matrices in the space of generalized doubly stochastic matrices. J. Math. Anal. Appl. **222**(2), 562–568 (1998) https://doi.org/10.1006/jmaa.1998.5970

[8] Glunt, W., Hayden, T.L., Reams, R.: The nearest generalized doubly stochastic matrix to a real matrix with the same first and second moments. Comput. Appl. Math. **27**(2), 201–210 (2008)

[9] Bai, Z.-J., Chu, D., Tan, R.C.E.: Computing the nearest doubly stochastic matrix with a prescribed entry. SIAM J. Sci. Comput. **29**(2), 635–655 (2007) https://doi.org/10.1137/050639831

[10] Perfect, H., Mirsky, L.: Spectral properties of doubly-stochastic matrices. Monatsh. Math. **69**, 35–57 (1965) https://doi.org/10.1007/BF01313442

[11] Horn, R.A., Johnson, C.R.: Matrix Analysis, 2nd edn., p. 643. Cambridge University Press, Cambridge (2013)

[12] Wang, E.T.-H.: Maximum and minimum diagonal sums of doubly stochastic matrices. Linear Algebra Appl. **8**(6), 483–505 (1974) https://doi.org/10.1016/0024-3795(74)90083-4

[13] Burkard, R., Dell'Amico, M., Martello, S.: Assignment Problems. Society for Industrial and Applied Mathematics, Philadelphia (2012). https://doi.org/10.1137/1.9781611972238

[14] Edmonds, J., Karp, R.M.: Theoretical improvements in algorithmic efficiency for network flow problems. In: Combinatorial Structures and Their Applications (Proc. Calgary Internat. Conf., Calgary, Alta., 1969), pp. 93–96. Gordon and Breach, New York (1970)

[15] Cheng, D., Hu, X., Martin, C.: On the smallest enclosing balls. Commun. Inf. Syst. **6**(2), 137–160 (2006)

[16] Mordukhovich, B., Nam, N.M., Villalobos, C.: The smallest enclosing ball problem and the smallest intersecting ball problem: existence and uniqueness of solutions. Optim. Lett. **7**(5), 839–853 (2013) https://doi.org/10.1007/s11590-012-0483-7

19

[17] Marcus, M.: On subdeterminants of doubly stochastic matrices. Illinois J. Math. **1**, 583–590 (1957)

[18] Burkard, R.E.: Selected topics on assignment problems. Discrete Applied Mathematics **123**(1), 257–302 (2002) https://doi.org/10.1016/S0166-218X(01)00343-2

[19] Mehlum, M.: Doubly stochastic matrices and the assignment problem. Master's thesis, University of Oslo (2012)

[20] Aziznejad, S., Unser, M.: Duality mapping for Schatten matrix norms. Numer. Funct. Anal. Optim. **42**(6), 679–695 (2021) https://doi.org/10.1080/01630563.2021.1922438

[21] Neumann, J.: Some matrix-inequalities and metrization of matrix-space. Tomsk Univ. Rev. **1**, 286–300 (1937)

[22] De Baerdemacker, S., De Vos, A., Chen, L., Yu, L.: The Birkhoff theorem for unitary matrices of arbitrary dimensions. Linear Algebra Appl. **514**, 151–164 (2017) https://doi.org/10.1016/j.laa.2016.10.028