(127)

NORTH-HOLLAND

# Minimization of Norms and the Spectral Radius of a Sum of Nonnegative Matrices Under Diagonal Equivalence

Daniel Hershkowitz*
*Mathematics Department*
*Technion—Israel Institute of Technology*
*Haifa 32000, Israel*

Wenchao Huang
*Mathematics Department*
*University of Wisconsin–Madison*
*Madison, Wisconsin 53706*

Michael Neumann‡
*Mathematics Department*
*University of Connecticut–Storrs*
*Storrs, Connecticut 06268*

Hans Schneider†
*Mathematics Department*
*University of Wisconsin–Madison*
*Madison, Wisconsin 53706*

Submitted by Ludwig Elsner

ABSTRACT

We generalize in various directions a result of Friedland and Karlin on a lower bound for the spectral radius of a matrix that is positively diagonally equivalent to a

* The research of these authors was supported by their joint grant No. 90-00434 from the United States–Israel Binational Science Foundation, Jerusalem, Israel.

‡ The research of this author was supported in part by NSF Grant DMS-9306357.

† The research of this author was supported in part by NSF Grant DMS-9123318.





0024-3795/96/$15.00
SSDI 0024-3795(95)00521-R

doubly stochastic matrix. The original result characterizes the equality case for two special zero patterns of the doubly stochastic matrix. Here we characterize the equality cases for doubly stochastic matrices of general zero pattern. We further generalize the results to *sums* of matrices that are diagonally equivalent to doubly stochastic matrices. Our claims follow from inequalities we prove on norms of matrices. Finally, we prove the corresponding inequalities (and equalities) for nonnegative matrices that are not sums of matrices diagonally equivalent to doubly stochastic matrices.

---

## 1. INTRODUCTION

As a special case of a theorem on the scaling of irreducible nonnegative matrices, Friedland and Karlin prove the following result [6, Theorem 2.1]: Let $M$ be an $n \times n$ doubly stochastic matrix, and let $D$ be an $n \times n$ diagonal matrix with positive diagonal elements and determinant equal to 1. Then

$$\rho(DM) = \rho(MD) \geqslant 1, \tag{1.1}$$

where $\rho(A)$ denotes the spectral radius of a square matrix $A$. Furthermore, the equality case is proven in [6] for two specific zero patterns of $M$. That is, it is proven that if $M$ has no zero entries, then $\rho(DM) = 1$ if and only if $D$ is equal to the identity matrix $I$, and if $M$ is the matrix representing the simple cycle on $n$ vertices, that is, the matrix given by

$$m_{ij} = \begin{cases} 1, & \text{if } j = i + 1, \text{or if } i = n \text{ and } j = 1 \\ 0, & \text{otherwise} \end{cases} \tag{1.2}$$

then $\rho(DM) = 1$ for every diagonal matrix $D$ with positive diagonal elements and determinant equal to 1.

Let $Y$ and $X$ be positive diagonal matrices such that $\det(YX) = 1$. Since

$$\rho(YMX) + \rho(X(YMX)X^{-1}) = \rho(XYM),$$

(1.1) generalizes to

$$\rho(YMX) \geqslant 1, \tag{1.3}$$

[1, Theorem 4]. The conditions for equality in [6, Theorem 2.1] imply that for positive matrix $M$ we have $\rho(YMX) = 1$ if and only if $Y = X^{-1}$, and for the matrix $M$ defined by (1.2) we have $\rho(YMX) = 1$ for all positive diagonal matrices $Y$ and $X$ such that $\det(YX) = 1$.

In this paper we complete the result of [6] by characterizing the equality cases for doubly stochastic matrices of general zero pattern. We further generalize the result of [6] to *sums* of matrices that are diagonally equivalent to doubly stochastic matrices. Our results follow from inequalities we prove on norms of matrices. Finally, we prove the corresponding inequalities (and equalities) for nonnegative matrices that are not sums of matrices diagonally equivalent to doubly stochastic matrices.

We now describe the paper in some more detail. Section 2 contains notation and preliminaries. We review relations between the algebraic and the geometric means of sequences of positive numbers. Also, we review some definitions and properties of certain norms of matrices.

In Section 3 we discuss sums of matrices that are positively diagonally equivalent to doubly stochastic matrices. We introduce a lower bound for submultiplicative norms of such matrices, and we characterize those cases in which the lower bound is attained. Our results are used to obtain a lower bound for the spectral radius of such sums and to characterize the equality case for the spectral radius.

In Section 4 we introduce a lower bound for submultiplicative norms of sums of nonnegative matrices in terms of norms of certain generalized doubly stochastic matrices. Here too we characterize those cases in which the lower bound is attained. We also obtain a lower bound for the spectral radius of such sums and characterize the corresponding equality case.

## 2. NOTATION AND PRELIMINARIES

2.1. NOTATION. Let $\mathbf{v}$ be a vector. We denote by $G(\mathbf{v})$ the geometric mean of the elements of $\mathbf{v}$.

Let $\mathbf{v}^1, \dots, \mathbf{v}^m$ be $n$-vectors with nonnegative elements. We have

$$G(\mathbf{v}^1 + \cdots + \mathbf{v}^m) \geqslant G(\mathbf{v}^1) + \cdots + G(\mathbf{v}^m), \tag{2.2}$$

where

$$G(\mathbf{v}^1 + \cdots + \mathbf{v}^m) = G(\mathbf{v}^1) + \cdots + G(\mathbf{v}^m)$$

$$\Leftrightarrow \begin{cases} \mathbf{v}^1 + \cdots + \mathbf{v}^m \text{ has a zero component} \\ \text{or} \\ \mathbf{v}^1, \dots, \mathbf{v}^m \text{ are proportional;} \end{cases} \tag{2.3}$$

see (2.7.1) in [8, p. 21].

2.4. NOTATION. Let $\mathbf{v}$ be a vector $(v_i)_{i=1}^n$, and let $r$ be a positive integer. We let $M_r(\mathbf{v}) = ((1/n)\Sigma_{i=1}^n v_i^r)^{1/r}$. Also, we let $M_\infty(\mathbf{v}) = \max_{i \in \{1,\dots,n\}}(v_i)$.

Let $\mathbf{v}$ be a vector with nonnegative elements. Recall that $\mathbf{v}$ satisfies the algebraic–geometric mean inequality, that is,

$$M_1(\mathbf{v}) \geqslant G(v), \tag{2.5}$$

where

$$M_1(\mathbf{v}) = G(\mathbf{v}) \Leftrightarrow \text{all the elements of } \mathbf{v} \text{ are the same.} \tag{2.6}$$

Also, for positive integers $r$ and $s$ (including $s = \infty$) we have

$$s > r \Rightarrow \mathbf{M}_s(\mathbf{v}) \geqslant \mathbf{M}_r(\mathbf{v}); \tag{2.7}$$

see (2.9.1) in [8, p. 26].

2.8. DEFINITION. (i) A norm $\|\cdot\|$ on $C^{nn}$ is called an *operator norm* if there exists a norm $\|\cdot\|$ on $\mathbb{C}^n$ such for every matrix $A$ in $\mathbb{C}^{nn}$ we have $\|A\| = \max_{x \in \mathbb{C}^n,\, x \neq 0}(\|Ax\|/\|x\|)$. The matrix norm is the norm *induced* by the corresponding norm on $\mathbb{C}^n$.

(ii) The *operator $l_p$ norm* $\|\cdot\|_p$ on $\mathbb{C}^{nn}$ is the operator norm induced by the $l_p$ norm on $\mathbb{C}^n$.

(iii) A norm on $\mathbb{C}^{nn}$ is called *submultiplicative* if $\|AB\| \leqslant \|A\|\cdot\|B\|$ for every pair $A$, $B$ of matrices in $\mathbb{C}^{nn}$.

The following lemma consists of well-known statements.

2.9. LEMMA. (i) *Every operator norm on $\mathbb{C}^{nn}$ is submultiplicative.*

(ii) *For every submultiplicative norm $\|\cdot\|$ on $\mathbb{C}^{nn}$ there exists an operator norm $\|\cdot\|^0$ on $\mathbb{C}^{nn}$ such that for every matrix $C$ in $\mathbb{C}^{nn}$ we have $\|C\| \geqslant \|C\|^0$.*

(iii) *For every submultiplicative norm $\|\cdot\|$ on $\mathbb{C}^{nn}$ and every matrix $C$ in $\mathbb{C}^{nn}$ we have $\|C\| \geqslant \rho(C)$.*

*Proof.* (i) See, e.g., Theorem 5.6.2 in [9, p. 293].

(ii) See, e.g., Theorem 5.6.26 in [9, p. 305].

(iii) Since for every operator norm $\|\cdot\|^0$ on $\mathbb{C}^{nn}$ and every matrix $C$ in $\mathbb{C}^{nn}$ we have $\|C\|^0 \geqslant \rho(C)$, the claim now follows from (ii). ■

2.10. DEFINITION. (i) A norm $\|\cdot\|$ on $\mathbb{C}^n$ is said to be *permutation invariant* if for every vector $\mathbf{x}$ in $\mathbb{C}^n$ and every permutation matrix $P$ in $\mathbb{C}^{nn}$ we have $\|P\mathbf{x}\| = \|\mathbf{x}\|$.

(ii) A norm $\|\cdot\|$ on $\mathbb{C}^{nn}$ is said to be *permutation invariant* if for every matrix $B$ in $\mathbb{C}^{nn}$ and every permutation matrix $P$ in $\mathbb{C}^{nn}$ we have $\|BP\| = \|PB\| = \|B\|$.

Note that every operator norm $\|\cdot\|$ on $\mathbb{C}^{nn}$ that is induced by a permutation invariant norm $\|\cdot\|$ on $\mathbb{C}^n$ is permutation invariant, since

$$\|AP\| = \max_{x\in\mathbb{C}^n,\, x\neq 0} \frac{\|APx\|}{\|x\|} = \max_{x\in C^n,\, x\neq 0} \frac{\|APx\|}{\|Px\|} = \|A\|$$

and

$$\|PA\| = \max_{x\in\mathbb{C}^n,\, x\neq 0} \frac{\|PAx\|}{\|x\|} = \max_{x\in\mathbb{C}^n,\, x\neq 0} \frac{\|Ax\|}{\|x\|} = \|A\|.$$

2.11. DEFINITION. A norm $\|\cdot\|$ on $\mathbb{C}^{nn}$ is said to be *unital* if $\|I\| = 1$.

2.12. OBSERVATION. (i) It follows from Definition 2.8 that operator norms are unital.

(ii) Since $I = I^2$, it follows from Definition 2.8 that for a submultiplicative norm $\|\cdot\|$ we always have $\|I\| \geqslant 1$.

(iii) For a permutation invariant norm $\|\cdot\|$ on $\mathbb{C}^{nn}$ and a doubly stochastic matrix $A$ we have $\|A\| \leqslant \|I\|$. To see this observe that by Birkhoff's theorem [2; 10; 11, Theorem 1.7], the matrix $A$ can be written as $A = \sum_{k=1}^{m} \alpha_k P_k$, where $\alpha_1, \ldots, \alpha_m$ are positive numbers satisfying $\sum_{k=1}^{m} \alpha_k = 1$ and where $P_1, \ldots, P_m$ are permutation matrices. Since

$$\|A\| = \left\| \sum_{k=1}^{m} \alpha_k P_k \right\| \leqslant \sum_{k=1}^{m} \alpha_k \|P_k\|,$$

and since for a permutation invariant norm we have $\|P_k\| = \|I\|$, the claim follows.

(iv) Since 1 is an eigenvalue of every doubly stochastic matrix, it follows by Lemma 2.9.ii that for a submultiplicative norm $\|\cdot\|$ and a doubly stochastic matrix $A$ we have $\|A\| \geqslant 1$.

(v) For a permutation invariant submultiplicative unital norm $\|\cdot\|$ and a doubly stochastic matrix $A$ we have $\|A\| = 1$. In particular, this claim applies to permutation invariant operator norms. This assertion follows since by statement (iii) above we have $\|A\| \leqslant 1$ while by statement (iv) we have $\|A\| \geqslant 1$.

## 3. DIAGONAL EQUIVALENCE OF DOUBLY STOCHASTIC MATRICES

3.1. DEFINITION. A nonnegative matrix is said to be *generalized doubly stochastic* if it is a scalar multiple of a doubly stochastic matrix. Note that a square zero matrix is generalized doubly stochastic.

The following theorem is a basic result from which we derive the principal inequalities of this section, viz. Theorems 3.11, 3.13, and 3.15.

3.2. THEOREM. *Let $\|\cdot\|$ be any operator $l_p$ norm on $\mathbb{C}^{nn}$. Let $t$ be a positive integer, let $M_1, \dots, M_t$ be doubly stochastic $n \times n$ matrices, and let $Y_i$ and $X_i$ be positive diagonal matrices, where $\beta_i = \sqrt[n]{\det(Y_i X_i)}$, $i = 1, \dots, t$. Then*

$$\left\| \sum_{i=1}^{t} Y_i M_i X_i \right\| \geqslant \sum_{i=1}^{t} \beta_i. \tag{3.3}$$

*Furthermore, the following are equivalent:*

(i) *We have $\|\sum_{i=1}^{t} Y_i M_i X_i\| = \sum_{i=1}^{t} \beta_i$.*

(ii) *For every $i$, $i \in \{1, \dots, t\}$, we have $Y_i M_i X_i = \beta_i M_i$.*

(iii) *We have $\sum_{i=1}^{t} Y_i M_i X_i = \sum_{i=1}^{t} \beta_i M_i$.*

*Proof.* Let $\mathbf{e}$ be the column $n$-vector all of whose entries are equal to 1. In view of (2.7), (2.5), and (2.2), we have

$$\begin{aligned}\left\| \sum_{i=1}^{t} Y_i M_i X_i \mathbf{e} \right\| &= n^{1/p} M_p\left( \sum_{i=1}^{t} Y_i M_i X_i \mathbf{e} \right) \geqslant n^{1/p} M_1\left( \sum_{i=1}^{t} Y_i M_i X_i \mathbf{e} \right) \\ &\geqslant n^{1/p} G\left( \sum_{i=1}^{t} Y_i M_i X_i \mathbf{e} \right) \geqslant n^{1/p} \sum_{i=1}^{t} G(Y_i M_i X_i \mathbf{e}). \end{aligned} \tag{3.4}$$

Let $i \in \{1, \dots, t\}$. By Birkhoff's theorem, the doubly stochastic matrix $M_i$ can be written as $M_i = \sum_{k=1}^{m} \alpha_k P_k$, where $\alpha_1, \dots, \alpha_m$ are positive numbers satisfying $\sum_{k=1}^{m} \alpha_k = 1$, and where $P_1, \ldots, P_m$ are permutation matrices. Observe that for every $k$, $1 \leqslant k \leqslant m$, we have $Y_i P_k X_i \mathbf{e} = ((X_i)_{\sigma(j)\sigma(j)}(Y_i)_{jj})_{j=1}^{n}$, where $\sigma$ is the permutation corresponding to $P_k$. Therefore, since $\beta_i = \sqrt[n]{\det(Y_i X_i)}$, it follows that

$$G(\alpha_k Y_i P_k X_i \mathbf{e}) = \alpha_k \beta_i. \tag{3.5}$$

By (2.2) it now follows from (3.5) that

$$G(Y_i M_i X_i \mathbf{e}) = G\left(\sum_{k=1}^{m} Y_i \alpha_k P_k X_i \mathbf{e}\right) \geqslant \sum_{k=1}^{m} G(Y_i \alpha_k P_k X_i \mathbf{e}) \geqslant \sum_{k=1}^{m} \alpha_k \beta_i = \beta_i \tag{3.6}$$

and hence it follows from (3.4) that

$$\left\| \sum_{i=1}^{t} Y_i M_i X_i \mathbf{e} \right\| \geqslant n^{1/p} \sum_{i=1}^{t} \beta_i.$$

We now have

$$\left\| \sum_{i=1}^{t} Y_i M_i X_i \right\| \geqslant \frac{\|\sum_{i=1}^{t} Y_i M_i X_i \mathbf{e}\|}{\|\mathbf{e}\|} \geqslant \frac{n^{1/p} \sum_{i=1}^{t} \beta_i}{n^{1/p}} = \sum_{i=1}^{t} \beta_i, \tag{3.7}$$

proving our inequality claim.

We now prove the equality case.

(i) $\Rightarrow$ (ii). If (i) holds then, in view of (3.7), we have $\|\sum_{i=1}^{t} Y_i M_i X_i \mathbf{e}\| = n^{1/p} \sum_{i=1}^{t} \beta_i$, and by (3.4) and (3.6) we must have that

$$M_1\left(\sum_{i=1}^{t} Y_i M_i X_i \mathbf{e}\right) = G\left(\sum_{i=1}^{t} Y_i M_i X_i \mathbf{e}\right), \tag{3.8}$$

$$G\left(\sum_{i=1}^{t} Y_i M_i X_i \mathbf{e}\right) = \sum_{i=1}^{t} G(Y_i M_i X_i \mathbf{e}), \tag{3.9}$$

and for every $i \in \{1, \ldots, t\}$

$$G\left(\sum_{k=1}^{m} \alpha_k Y_i P_k X_i \mathbf{e}\right) = \sum_{k=1}^{m} G(\alpha_k Y_i P_k X_i \mathbf{e}), \tag{3.10}$$

where $M_i = \sum_{k=1}^{m} \alpha_k P_k$ as above. By (2.6), it follows from (3.8) that all elements of $\sum_{i=1}^{t} Y_i M_i X_i \mathbf{e}$ are the same. By (2.3), it follows from (3.9) that $Y_i M_i X_i \mathbf{e}$ $i = 1, \dots, t$, are proportional. It also follows by (2.3) from (3.10) that $Y_i P_k X_i$, $k = 1, \dots, m$, are proportional. Thus, for every $k$, all the elements of $Y_i P_k X_i \mathbf{e} = ((X_i)_{\sigma(j)\sigma(j)}(Y_i)_{jj})_{j=1}^{n}$ are the same. Since $\beta_i = \sqrt[n]{\det(Y_i X_i)}$, it now follows that $(X_i)_{\sigma(j)\sigma(j)}(Y_i)_{jj} = \beta_i$, $j = 1, \dots, n$, implying that $Y_i P_k X_i = \beta_i P_k$, and hence $Y_i M_i X_i = \beta_i M_i$.

(ii) $\Rightarrow$ (iii) is trivial.

(iii) $\Rightarrow$ (i). The matrix $\sum_{i=1}^{t} \beta_i M_i$ is a generalized doubly stochastic matrix, with row sums and column sums all equal to $\sum_{i=1}^{t} \beta_i$. Therefore, it is equal to the scalar $\sum_{i=1}^{t} \beta_i$ times a doubly stochastic matrix. It now follows from (iii) that $\|\sum_{i=1}^{t} Y_i M_i X_i\| = \|\sum_{i=1}^{t} \beta_i M_i\| = \sum_{i=1}^{t} \beta_i$. ■

Let $A$ be an $n \times n$ matrix. As is well known, $A$ can be brought, using an identical permutation on its rows and on its columns, into an upper (or lower) block triangular form, with irreducible square diagonal blocks. Such form is said to be the *Frobenius normal form* of $A$. A diagonal block in the Frobenius normal form of $A$ is said to be a *component* of $A$. The matrix $A$ is said to be *completely reducible* if its Frobenius normal form is block diagonal. Using Theorem 3.2 we can now prove

3.11. THEOREM. *Let $M_1, \dots, M_t$ be doubly stochastic $n \times n$ matrices, and let $Y_i$ and $X_i$ be positive diagonal matrices, where $\beta_i = \sqrt[n]{\det(Y_i X_i)}$, $i = 1, \dots, t$. Then*

$$\rho\left(\sum_{i=1}^{t} Y_i M_i X_i\right) \geqslant \sum_{i=1}^{t} \beta_i. \tag{3.12}$$

*Furthermore, the following are equivalent:*

(i) *We have $\rho(\sum_{i=1}^{t} Y_i M_i X_i) = \sum_{i=1}^{t} \beta_i$.*

(ii) *There exists a positive diagonal matrix $D$ such that for every $i$, $i \in \{1, \dots, t\}$, we have $Y_i M_i X_i = \beta_i D M_i D^{-1}$.*

(iii) *There exists a positive diagonal matrix $D$ such that $\sum_{i=1}^{t} Y_i M_i X_i = D(\sum_{i=1}^{t} \beta_i M_i) D^{-1}$.*

*Proof.* Note that doubly stochastic matrices are completely reducible and that a sum of nonnegative completely reducible matrices is a completely reducible matrix. Therefore, the matrix $\beta = \sum_{i=1}^{t} Y_i M_i X_i$ is completely reducible. Let $D$ be a direct sum of positive diagonal matrices $D^1, \dots, D^q$, whose main diagonals are the Perron vectors corresponding to the components $B^1, \dots, B^q$ of $B$. Without loss of generality we may assume that $B = \mathrm{diag}(B^1, \dots, B^q)$. Then $\|D^{-1}BD\|_\infty = \max_{1 \leqslant i \leqslant q} \|(D^i)^{-1}B^iD^i\|_\infty$. Let $j$, $1 \leqslant j \leqslant q$, be such that $\|D^{-1}BD\|_\infty = \|(D^j)^{-1}B^jD^j\|_\infty$. As in [13], we have $\|(D^j)^{-1}B^jD^j\|_\infty = \rho(B^j)$, and since $\rho(B^j) \leqslant \rho(B) \leqslant \|D^{-1}BD\|_\infty$, it now follows that $\rho(B) = \|D^{-1}BD\|_\infty$. By Theorem 3.2 we now obtain $\rho(B) \geqslant \sum_{i=1}^t \beta_i$.

We next prove the equality case.

(i) $\Rightarrow$ (ii). Let $D$ be the matrix defined above. Since $\|D^{-1}BD\|_\infty = \rho(B) = \sum_{i=1}^t \beta_i$, then, by Theorem 3.2, we have that $D^{-1}Y_iM_iX_iD = \beta_iM_i$.

(ii) $\Leftrightarrow$ (iii) by the corresponding equivalence in Theorem 3.2.

(iii) $\Rightarrow$ (i). If (iii) holds, then clearly $\rho(\sum_{i=1}^t Y_iM_iX_i) = \rho(\sum_{i=1}^t \beta_iM_i)$. Since $\sum_{i=1}^t \beta_iM_i$ is a generalized doubly stochastic matrix with row sums $\sum_{i=1}^t \beta_i$, we have that $\rho(\sum_{i=1}^t \beta_iM_i) = \sum_{i=1}^t \beta_i$, proving our claim. ∎

Theorem 3.11 generalizes [6, Theorem 2.1] by discussing *sums* of matrices diagonally equivalent to doubly stochastic matrices rather than a single such matrix. Also, the equality cases are characterized for general zero patterns rather than the two specific types mentioned in [6]. We further comment that (3.12) generalizes the result in (7) of Theorem 4 of [1] in the case that the matrix $A$ in Theorem 4 of [1] is assumed to be doubly stochastic. Note that (3.12) applies to the sum of diagonal equivalencies of several doubly stochastic matrices, while the result in [1] permits only the sum of diagonal equivalences of a single doubly stochastic matrix.

The following result is related to Theorem 3.11 and follows from it.

3.13. THEOREM. *Let $M_1, \dots, M_t$ be doubly stochastic $n \times n$ matrices, and let $Y_i$ and $X_i$ be positive diagonal matrices, where $\beta_i = \sqrt[n]{\det(Y_iX_i)}$, $i = 1, \dots, t$. Then*

$$\sum_{i=1}^{t} \rho(Y_iM_iX_i) \geqslant \sum_{i=1}^{t} \beta_i. \tag{3.14}$$

*Furthermore, the following are equivalent:*

(i) *We have $\sum_{i=1}^t \rho(Y_iM_iX_i) = \sum_{i=1}^t \beta_i$.*

(ii) *For every $i$, $i \in \{1, \dots, t\}$, we have $\rho(Y_iM_iX_i) = \beta_i$.*

(iii) *For every $i$, $i \in \{1, \dots, t\}$, there exists a positive diagonal matrix $D_i$ such that we have $Y_iM_iX_i = \beta_iD_iM_iD_i^{-1}$.*

*Proof.* In view of (1.3) we have $\rho(Y_iM_iX_i) \geqslant \beta_i$, $i = 1, \dots, t$, implying (3.14). Also, the equivalence of (i) and (ii) follows. The equivalence of (ii) and

(iii) follows from the equivalence of (i) and (ii) in Theorem 3.11, when applied to a single matrix. ■

It should be mentioned that the inequalities (3.12) and (3.14) are independent; that is, there is no simple inequality relating $\rho(\sum_{i=1}^{t} Y_i M_i X_i)$ to $\sum_{i=1}^{t} \rho(Y_i M_i X_i)$. For example, let

$$M_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad M_2 = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}.$$

Note that

$$\rho(M_1 + M_2) = 2 = \rho(M_1) + \rho(M_2)$$

is the minimum of $\rho(\sum_{i=1}^{2} Y_i M_i X_i)$ and of $\sum_{i=1}^{2} \rho(Y_i M_i X_i)$, where $\det(Y_1 X_1) = \det(Y_2 X_2) = 1$, as is asserted by Theorems 3.11 and 3.13. Now, for the matrices

$$B_1 = \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 1/2 & 0 \end{pmatrix}$$

and

$$B_2 = \begin{pmatrix} 1/2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 1/2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1/8 & 1/2 \\ 1/2 & 2 \end{pmatrix},$$

we have

$$\rho(B_1 + B_2) \approx 2.9006 < 3.125 \approx \rho(B_1) + \rho(B_2),$$

while if we replace $B_2$ by

$$B_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix} = \begin{pmatrix} 0 & 1/2 \\ 2 & 0 \end{pmatrix},$$

we obtain

$$\rho(B_1 + B_3) = 2.5 > 2 = \rho(B_1) + \rho(B_3).$$

This example also demonstrates that, in general, equality in (3.14) does not imply equality in (3.12). Note that in the converse direction it follows by Theorems 3.11 and 3.13 that equality in (3.12) does imply equality in (3.14).

Using Theorem 3.11, we can now prove Theorem 3.2 for a wider class of norms.

3.15. THEOREM. *Let* $\|\cdot\|$ *be a submultiplicative norm on* $\mathbb{C}^{nn}$. *Let* $M_1, \dots, M_t$ *be doubly stochastic* $n \times n$ *matrices, and let* $Y_i$ *and* $X_i$ *be positive diagonal matrices, where* $\beta_i = \sqrt[n]{\det(Y_i X_i)}$, $i = 1, \dots, t$. *Then*

$$\left\| \sum_{i=1}^{t} Y_i M_i X_i \right\| \geqslant \sum_{i=1}^{t} \beta_i. \tag{3.16}$$

*Proof.* By Lemma 2.9.iii we have that $\|\sum_{i=1}^t Y_i M_i X_i\| \geqslant \rho(\sum_{i=1}^t Y_i M_i X_i)$. By Theorem 3.11 we have that $\rho(\sum_{i=1}^t Y_i M_i X_i) \geqslant \sum_{i=1}^t \beta_i$, and our assertion follows. ∎

To consider the equality case in Theorem 3.15 we need in addition permutation invariance and unitality requirement on the norm.

3.17. NOTATION. Let $B$ be an $n \times n$ matrix and let $\alpha$ and $\beta$ be nonempty subsets of $\{1, \dots, n\}$. We denote by $B[\alpha \mid \beta]$ the submatrix of $A$ whose rows are indexed by $\alpha$ and whose columns are indexed by $\beta$ in their natural order. We denote by $\beta[\alpha]$ the principal submatrix $B[\alpha \mid \alpha]$.

The following is an immediate consequence of Perron–Frobenius.

3.18. LEMMA. *Let* $A$ *be a doubly stochastic matrix and let* $D$ *be a positive diagonal matrix. Then* $D^{-1}AD$ *is row stochastic or column stochastic if and only if* $D^{-1}AD = A$.

*Proof.* The "if" direction is trivial. Conversely, without loss of generality we can assume that $D^{-1}AD$ is row stochastic (otherwise consider its transpose $DA^TD^{-1}$). Then $D^{-1}AD\mathbf{e} = \mathbf{e}$ and hence $AD\mathbf{e} = D\mathbf{e}$. As doubly stochastic matrices are completely reducible, let $\alpha_1, \dots, \alpha_k$ be the subsets of $\{1, \dots, n\}$ that index the components of the completely reducible matrix $A$. Then

$$A[\alpha_i]D[\alpha_i]\mathbf{e}[\alpha_i] = D[\alpha_i]\mathbf{e}[\alpha_i]. \tag{3.19}$$

Observe that $\mathbf{e}[\alpha_i]$ is a positive eigenvector, corresponding to the Perron root 1 of the row stochastic matrix $A[\alpha_i]$. By the Perron–Frobenius theory, the irreducible matrix $A[\alpha_i]$ has a unique (up to scalar multiple) positive eigenvector, and so it follows from (3.19) that the submatrix $D[\alpha_i]$ is a scalar matrix, $i = 1, \dots, k$. Since $D^{-1}AD = \oplus_{i \in \{1,\dots,k\}} D[\alpha_i]^{-1}A[\alpha_i]D[\alpha_i]$, we have that $D^{-1}AD = A$. ∎

3.20. Theorem. *Let $\|\cdot\|$ be a submultiplicative permutation invariant unital norm on $\mathbb{C}^{nn}$. Let $M_1, \dots, M_t$ be doubly stochastic $n \times n$ matrices, and let $Y_i$ and $X_i$ be positive diagonal matrices with $\beta_i = \sqrt[n]{\det(Y_i X_i)}$, $i = 1, \dots, t$. The following are equivalent:*

(i) *We have $\|\sum_{i=1}^t Y_i M_i X_i\| = \sum_{i=1}^t \beta_i$.*

(ii) *For every $i$, $i \in \{1, \dots, t\}$, we have that $Y_i M_i X_t = \beta_i M_i$.*

(iii) *We have $\sum_{i=1}^t Y_i M_i X_i = \sum_{i=1}^t \beta_i M_i$.*

*Proof.* (i) $\Rightarrow$ (iii). By Lemma 2.9.iii and Theorem 3.11 we have that $\|\sum_{i=1}^t Y_i M_i X_i\| \geqslant \rho(\sum_{i=1}^t Y_i M_i X_i) \geqslant \sum_{i=1}^t \beta_i$. In view of (i) we now have $\rho(\sum_{i=1}^t Y_i M_i X_i) = \sum_{i=1}^t \beta_i$ and it follows from Theorem 3.11 that there exists a positive diagonal matrix $D$ such that $\sum_{i=1}^t Y_i M_i X_i = D(\sum_{i=1}^t \beta_i M_i)D^{-1}$. Denote by $B$ the matrix $\sum_{i=1}^t \beta_i M_i$, and assume that $DBD^{-1} \neq B$. Since $B$ is a generalized doubly stochastic matrix, by Lemma 3.18, the matrix $DBD^{-1}$ is neither row stochastic nor column stochastic. By Theorem 1 in [5], there exists a permutation matrix $P$ such that $\sum_{i=1}^t \beta_i$ is not an eigenvalue of $PDBD^{-1}$. We have $PDBD^{-1} = PDP^T PBD^{-1}$, and so since $PDP^T$ is a diagonal matrix and since $\det(PDP^T D^{-1})$, it follows from Theorem 3.11 that

$$\rho(PDBD^{-1}) \geqslant \sum_{i=1}^t \beta_i. \tag{3.21}$$

Therefore, as $\rho(PDBD^{-1})$ is an eigenvalue of the nonnegative matrix $PDBD^{-1}$, it follows from (3.21) that $\rho(PDBD^{-1}) > \sum_{i=1}^t \beta_i$. As our norm is permutation invariant, we now have that

$$\left\| \sum_{i=1}^t Y_i M_i X_i \right\| = \left\| P \sum_{i=1}^t Y_i M_i X_i \right\| \geqslant \rho\left( P \sum_{i=1}^t Y_i M_i X_i \right)$$

$$= \rho(PDBD^{-1}) > \sum_{i=1}^t \beta_i,$$

in contradiction to (i). Therefore, our assumption that $DBD^{-1} \neq B$ is false and our claim follows.

(ii) $\Leftrightarrow$ (iii) by Theorem 3.2.

(iii) $\Rightarrow$ (i). Since $\sum_{i=1}^{t} \beta_i M_i$ is a generalized doubly stochastic matrix with row sums $\sum_{i=1}^{t} \beta_i$, the implication follows by Observation 2.12.v. ∎

## 4. DIAGONAL EQUIVALENCE OF NONNEGATIVE MATRICES

While in the previous section we discussed lower bounds for norms and spectral radius of matrices that are sums of matrices diagonally equivalent to doubly stochastic matrices, in this section we discuss those bounds for matrices that lie in certain classes of sums of diagonal equivalence of nonnegative matrices.

4.1. DEFINITION. Let $A_1, \dots, A_t$ be $n \times n$ matrices. We define the *restricted diagonal equivalence class* $R(A_1, \dots, A_t)$ by

$$\left\{ \sum_{i=1}^{t} Y_i A_i X_i : Y_i,\ X_i \text{ are positive diagonal matrices} \right.$$

$$\left. \text{satisfying} \det(Y_i X_i) = 1,\ i \in \{1, \dots, t\} \right\}.$$

4.2. DEFINITION. Let $A$ be an $n \times n$ matrix.

(i) A (*generalized*) *diagonal* in $A$ is a set of $n$ positions in $A$, such that every two positions are neither in the same row nor in the same column. A diagonal in $A$ is said to be *strictly nonzero* if all elements of $A$ that lie on that diagonal are nonzero.

(ii) A *cycle* in $A$ is a set of $n$ positions $(i_1, i_2), (i_2, i_3), \dots, (i_t, i_{t+1})$, where $i_1, \dots, i_t$ are distinct and $i_{t+1} = i_1$. A cycle in $A$ is said to be *strictly nonzero* if all elements of $A$ that lie on that cycle are nonzero.

4.3. NOTATION. Let $A$ be an $n \times n$ matrix. We denote by $A^{\#}$ the matrix obtained from $A$ by setting equal to 0 all elements that do not lie on a strictly nonzero diagonal.

4.4. EXAMPLE. Let

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

$A$ has two nonzero diagonals; one consists of the elements in the positions $(1,2)$, $(2,1)$, and $(3,3)$ and another one consists of the elements in the positions $(1,2)$, $(2,3)$, and $(3,1)$. Hence

$$A^{\#} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

4.5. REMARK. Let $A$ be a nonnegative $n \times n$ matrix. It is an immediate consequence of well-known results that $R(A^{\#})$ contains a unique generalized doubly stochastic matrix, e.g., [12].

4.6. THEOREM. *Let $\|\cdot\|$ be a submultiplicative permutation invariant unital norm on $\mathbb{C}^{nn}$, and let $A_1, \dots, A_t$ be nonnegative $n \times n$ matrices. Then*

$$\inf\{\|C\| : C \in R(A_1, \dots, A_t)\} = \sum_{i=1}^{t} \beta_i,$$

*where $\beta_i$ is the row (and column) sum of the (unique) generalized doubly stochastic matrix $M_i$ in $R(A_i^{\#})$, $i \in \{1, \dots, t\}$.*

*Proof.* Let $Y_i$ and $X_i$ be positive diagonal matrices satisfying $\det(Y_i X_i) = 1$, $i \in \{1, \dots, t\}$. By Lemma 2.9.iii we have that

$$\|\Sigma_{i=1}^{t} Y_i A_i X_i\| \geqslant \rho(\Sigma_{i=1}^{t} Y_i A_i X_i).$$

By the Perron–Frobenius spectral theory for nonnegative matrices we have that $\rho(\Sigma_{i=1}^{t} Y_i A_i X_i) \geqslant \rho(\Sigma_{i=1}^{t} Y_i A_i^{\#} X_i)$ and, since $R(A_i^{\#})$ contains the generalized doubly stochastic matrix $M_i$, it follows from Theorem 3.11 that $\rho(\Sigma_{i=1}^{t} Y_i A_i^{\#} X_i) \geqslant \Sigma_{i=1}^{t} \beta_i$. Hence,

$$\left\| \sum_{i=1}^{t} Y_i A_i X_i \right\| \geqslant \rho\left( \sum_{i=1}^{t} Y_i A_i X_i \right) \geqslant \rho\left( \sum_{i=1}^{t} Y_i A_i^{\#} X_i \right) \geqslant \sum_{i=1}^{t} \beta_i,$$

proving that

$$\inf\{\|C\| : C \in R(A_1, \dots, A_t)\} \geqslant \sum_{i=1}^{t} \beta_i. \tag{4.7}$$

We now consider the case of equality in (4.7). Let $i \in \{1, \dots, t\}$. We distinguish between two cases:

(i) $A_i^{\#} = 0$: In this case $A_i$ has no strictly nonzero diagonal. By the Frobenius–König theorem [7; 11, Theorem 1.7.1, p. 97], there exist subsets $\alpha$ and $\gamma$ of $\{1, \dots, n\}$ satisfying $|\alpha| + |\gamma| = m \geqslant n + 1$, such that $A_i[\alpha \mid \gamma] = 0$. Obviously, we have $m \leqslant 2n$. If $m = 2n$ then $A_i = 0$ and there is nothing to prove. So, we may assume that $m < 2n$. Let $\epsilon$ be a positive number, and let $Y_i^\epsilon$ and $X_i^\epsilon$ be the positive diagonal matrices with diagonal elements

$$(Y_i^\epsilon)_{jj} = \begin{cases} \dfrac{1}{\epsilon}, & j \in \alpha \\ \epsilon^{m/(2n-m)}, & j \notin \alpha \end{cases}, \qquad (X_i^\epsilon)_{jj} = \begin{cases} \dfrac{1}{\epsilon}, & j \in \gamma \\ \epsilon^{m/(2n-m)}, & j \notin \gamma \end{cases}.$$

Observe that $\det(Y_i^\epsilon X_i^\epsilon) = 1$. Also, for the matrix $C = Y_i^\epsilon A_i X_i^\epsilon$ we have that $C[\alpha \mid \gamma] = 0$, $C[\alpha \mid \gamma^C] = \epsilon^{(2m-2n)/(2n-m)} A_i[\alpha \mid \gamma^C]$, $C[\alpha^C \mid \gamma] = \epsilon^{(2m-2n)/(2n-m)} A_i[\alpha^C \mid \gamma]$, and $C[\alpha^C \mid \gamma^C] = \epsilon^{2m/(2m-m)} A_i[\alpha^C \mid \gamma^C]$. Therefore, we have $\lim_{\epsilon \to 0}(Y_i^\epsilon A_i X_i^\epsilon) = 0$. Since in this case $M_i = 0$, we have

$$\lim_{\epsilon \to 0} (Y_i^\epsilon A_i X_i^\epsilon) = M_i. \tag{4.8}$$

(ii) $A_i^{\#} \neq 0$: In this case $A_i$ has a strictly nonzero diagonal. Let $P$ be the permutation matrix such that $PA_i$ has a positive main diagonal, and let $Q$ be the permutation matrix such that $E = QPA_iQ^T$ is in Frobenius normal form. Since every strictly nonzero diagonal of $A_i$ permutes to a strictly nonzero diagonal of $E$, it follows that $E^{\#} = QPA_i^{\#}Q^T$. Since $E$ is a completely reducible matrix with positive diagonal elements, it follows that every nonzero element in a component $E_{jj}$ of $E$ lies on a strictly nonzero diagonal in $E$. Also, obviously every nonzero element in an off-diagonal block of $E$ does not lie on a strictly nonzero diagonal in $E$. Therefore, we have that $E^{\#} = \oplus_j E_{jj}$. Let $Y_i$ and $X_i$ be positive diagonal matrices with diagonal blocks $(Y_i)_{jj}$ and $(X_i)_{jj}$, respectively, such that $\det(Y_i X_i) = 1$ and $Y_i E^{\#} X_i$ is the (unique) generalized doubly stochastic matrix in $R(E^{\#})$. Let $\epsilon$ be a positive number, and let $Y_i^\epsilon$ and $X_i^\epsilon$ be the positive diagonal matrices defined by

$$Y_i^\epsilon = \bigoplus_j \frac{1}{\epsilon^{j-1}}(Y_i)_{jj}, \qquad X_i^\epsilon = \bigoplus_j \epsilon^{j-1}(X_i)_{jj}.$$

Then $\det(Y_i^\epsilon X_i^\epsilon) = 1$. Also, observe that while the diagonal blocks of $Y_i^\epsilon E X_i^\epsilon$ are $(Y_i)_{jj}E_{jj}(X_i)_{jj}$, the off-diagonal blocks approach 0 as $\epsilon \to 0$. Therefore, $Y_i^\epsilon E X_i^\epsilon$ approaches the generalized doubly stochastic matrix $Y_i E^{\#} X_i$ as $\epsilon \to 0$. Note that

$$Y_i E^{\#} X_i = Y_i Q P A_i^{\#} Q^T X_i = QP(P^T Q^T Y_i QP) A_i^{\#}(Q^T X_i Q)Q^T. \quad (4.9)$$

Let $\tilde{Y}_i$ and $\tilde{X}_i$ be the diagonal matrices $P^T Q^T Y_i QP$ and $Q^T X_i Q$ respectively. Since $Y_i E^{\#} X_i$ is generalized doubly stochastic, it follows from (4.9) that $\tilde{Y}_i A_i^{\#} \tilde{X}_i$ is generalized doubly stochastic and, as $\det(\tilde{Y}_i \tilde{X}_i) = \det(Y_i X_i) = 1$, the product $\tilde{Y}_i A_i^{\#} \tilde{X}_i$ is equal to the unique generalized doubly stochastic matrix $M_i$ in $R(A_i^{\#})$. Therefore, it follows from (4.9) that

$$\lim_{\epsilon \to 0} (Y_i^\epsilon E X_i^\epsilon) = QPM_i. \quad (4.10)$$

Note that

$$Y_i^\epsilon E X_i^\epsilon = Y_i^\epsilon QPA_i Q^T X_i^\epsilon = QP(P^T Q^T Y_i^\epsilon QP) A_i (Q^T X_i^\epsilon Q)Q^T. \quad (4.11)$$

Let $\tilde{Y}_i^\epsilon$ and $\tilde{X}_i^\epsilon$ be the diagonal matrices $P^T Q^T Y_i^\epsilon QP$ and $Q^T X_i^\epsilon Q$ respectively. Observe that $\det(\tilde{Y}_i^\epsilon \tilde{X}_i^\epsilon) = 1$. By (4.10) and (4.11) we have

$$\lim_{\epsilon \to 0} \left(\tilde{Y}_i^\epsilon A_i \tilde{X}_i^\epsilon\right) = M_i. \quad (4.12)$$

It follows from (4.8) and (4.12) that the generalized doubly stochastic matrix $M = \sum_{i=1}^t M_i$ is on the boundary of $R(A_1, \dots, A_t)$. Since the row (and column) sums of $M$ are all equal to $\sum_{i=1}^t \beta_i$, if follows by Observation 2.12.v that $\|M\| = \sum_{i=1}^t \beta_i$. In view of (4.7), our claim follows. ■

4.13. THEOREM. *Let $A_1, \dots, A_t$ be nonnegative $n \times n$ matrices. Then*

$$\inf\{\rho(C) : C \in R(A_1, \dots, A_t)\} = \sum_{i=1}^t \beta_i,$$

*where $\beta_i$ is the row (and column) sum of the (unique) generalized doubly stochastic matrix $M_i$ in $R(A_i^{\#})$, $i \in \{1, \dots, t\}$.*

*Proof.* By the Perron–Frobenius theorem we have that

$$\inf\{\rho(C) : C \in R(A_1, \dots, A_t)\} \geqslant \inf\{\rho(C) : C \in R(A_1^{\#}, \dots, A_t^{\#})\},$$

and, by Theorem 3.11, we have that $\inf\{\rho(C) : C \in R(A_1^{\#}, \dots, A_t^{\#})\} \geqslant \sum_{i=1}^{t} \beta_i$. Since the generalized doubly stochastic matrix $\sum_{i=1}^{t} M_i$ belongs to $R(A_1^{\#}, \dots, A_t^{\#})$ and, as is proven in the proof of Theorem 4.6, is on the boundary of $R(A_1, \dots, A_t)$, the proof is now done. ■

As a corollary of Theorem 4.13 we obtain

4.14. Corollary. *Let $A_1, \dots, A_t$ be nonnegative $n \times n$ matrices and let $P_1, \dots, P_t$ and $Q_1, \dots, Q_t$ be $n \times n$ permutation matrices. Then*

$$\inf\{\rho(C) : C \therefore R(A_1, \dots, A_t)\}$$

$$= \inf\{\rho(C) : C \in R(P_1 A_1 Q_1, \dots, P_t A_t Q_t)\}.$$

*Proof.* Let $i \in \{1, \dots, t\}$ and let $M_i$ and $\tilde{M}_i$ be the (unique) generalized doubly stochastic matrices in $R(A_i^{\#})$ and $R((P_i A_i Q_i)^{\#})$ respectively. Since for all diagonal matrices $Y$ and $X$ we have that $P(YA_i X)Q = \tilde{Y}(P_i A_i Q_i)\tilde{X}$, where $\tilde{Y}$ and $\tilde{X}$ are the diagonal matrices $P_i Y P_i^T$ and $Q_i^T X Q_i$, respectively, it follows that $\tilde{M}_i$ is equal to the generalized doubly stochastic matrix $PM_iQ$. Our claim now follows by Theorem 4.13. ■

4.15. Remark. Note that since the spectral radius is invariant under diagonal similarity, where the spectral radius is concerned not much is gained by considering a diagonal equivalence $YAX$ rather than a diagonal scaling $AX$. However, the diagonal equivalence approach is essential for the norm results. We demonstrate this observation by the matrices

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad X = \begin{pmatrix} 1/\epsilon & 0 & 0 \\ 0 & \epsilon^2 & 0 \\ 0 & 0 & 1/\epsilon \end{pmatrix}.$$

We have that $\rho(AX) = \sqrt{2\epsilon}$, and so we can make $\rho(AX)$ as small as we wish. However, we cannot make $\|AX\|$ small, since some element of $AX$ is greater than or equal to 1. One does need a diagonal equivalence to reduce the norm of $A$, as is done in the proof of Theorem 4.6.

We now characterize the case in which the infimum in Theorem 4.6 is attained. We begin by stating the following lemma, which constitutes a theorem of Brualdi [3] (cf. also Theorem 3.2.5 in [4, p. 56]), and for which we provide here a simple different proof for the sake of completeness.

4.16. LEMMA. *Let $A$ be an $n \times n$ matrix. There exist permutation matrices $P$ and $Q$ such that $PAQ$ is irreducible if and only if $A$ has at least one nonzero element in each row and in each column.*

*Proof.* The "only if" is immediate. Conversely, let $t$ be the maximal number of nonzero elements of $A$, such that no two positions belong to the same row or same column. There exist permutation matrices $P$ and $Q$ such that $PAQ$ has nonzero elements in positions $(1,2),(2,3),\dots,(t-1,t),(t,1)$. Let $\alpha = \{1,\dots,t\}$. Because of the maximality of $t$ it follows that $A[\alpha^C] = 0$ and hence every column of $A[\alpha \mid \alpha^C]$ and every row of $A[\alpha^C \mid \alpha]$ has a nonzero element. It follows that for every $k$, $t < k \leqslant n$, there exist $i, j \in \{1,\dots,n\}$ such that $(i,k)$ and $(k,j)$ are arcs in the digraph of $PAQ$. Since the digraph of $PAQ[\alpha]$ contains a full cycle, it follows that the digraph of $PAQ$ is strongly connected and hence $PAQ$ is irreducible. ∎

4.17. REMARK. In the statement of Lemma 4.16 we asserted the existence of permutation matrices $P$ and $Q$ such that $PAQ$ is irreducible, while in [3] it is stated that there exists a permutation matrix such that $AQ$ is irreducible. These statements are equivalent as $PAQ$ is irreducible if and only if $P^T(PAQ)P = AQP$ is irreducible.

4.18. THEOREM. *Let $\|\cdot\|$ be a submultiplicative permutation invariant unital norm on $\mathbb{C}^{nn}$, let $A_1,\dots,A_t$ be nonnegative $n \times n$ matrices, and let $\beta_i$ be the row (and column) sum of the (unique) generalized doubly stochastic matrix $M_i$ in $R(A_i^\#)$, $i \in \{1,\dots,t\}$. Then the following are equivalent.*

(i) *There exists a matrix $C$ in $R(A_1,\dots,A_t)\}$ for which $\|C\| = \sum_{i=1}^t \beta_i$.*

(ii) *We have $A_i = A_i^\#$ for all $i \in \{1,\dots,t\}$.*

*Proof.* (i) $\Rightarrow$ (ii). Let $C = \sum_{i=1}^t Y_i A_i X_i$ be a matrix satisfying $\|C\| = \sum_{i=1}^t \beta_i$, where $Y_i$ and $X_i$ are positive diagonal matrices such that

$\det(Y_i X_i) = 1$, $i = 1, \ldots, t$. Assume that for some $j \in \{1, \ldots, t\}$ we have that $A_j \neq A_j^{\#}$. We distinguish between two cases:

(i) $A_i^{\#} = 0$ for all $i \in \{1, \ldots, t\}$: Since $A_j \neq 0$, we choose $P$ and $Q$ to be permutation matrices such that $PA_jQ$ has a positive diagonal element. It follows that

$$\rho(PCQ) > 0 = \sum_{i=1}^{t} \beta_i. \tag{4.19}$$

(ii) $A_i^{\#} \neq 0$ for some $i \in \{1, \ldots, t\}$: Here, the matrix $A_i$ has a strictly nonzero diagonal and so, by Lemma 4.16, there exist permutation matrices $P$ and $Q$ such that $PA_iQ$ is irreducible. It follows that $PCQ$ is an irreducible matrix. Since $A_j \neq A_j^{\#}$ it follows from the Perron–Frobenius theorem that

$$\rho(PCQ) > \rho\left(\sum_{i=1}^{t} PY_i A_i^{\#} X_i Q\right). \tag{4.20}$$

Note that $A_i^{\#} = \overline{Y}_i M_i \overline{X}_i$ for some positive diagonal matrices $\overline{Y}_i$ and $\overline{X}_i$ satisfying that $\det(\overline{Y}_i \overline{X}_i) = 1$. Thus we have $PY_i A_i^{\#} X_i Q = \tilde{Y}_i(PM_iQ)\tilde{X}_i$, where $\tilde{Y}_i$ and $\tilde{X}_i$ are the positive diagonal matrices $PY_i\overline{Y}_iP^T$ and $Q^T\overline{X}_iX_iQ$, respectively, satisfying that $\det(\tilde{Y}_i\tilde{X}_i) = 1$. Since $PM_iQ$ is a generalized doubly stochastic matrix that has the same line sums as $M_i$, it now follows by Theorem 3.11 that

$$\rho\left(\sum_{i=1}^{t} PY_i A_i^{\#} X_i Q\right) = \rho\left(\sum_{i=1}^{t} \tilde{Y}_i(PM_iQ)\tilde{X}_i\right) \geqslant \sum_{i=1}^{t} \beta_i,$$

and so, by (4.20), we obtain that

$$\rho(PCQ) > \sum_{i=1}^{t} \beta_i. \tag{4.21}$$

Since $\|\cdot\|$ is a permutationally invariant norm, it now follows by Lemma 2.9.iii that

$$\|C\| = \|PCQ\| \geqslant \rho(PCQ),$$

and so, in view of (4.19) and (4.21), we have a contradiction to $\|C\| = \sum_{i=1}^{t} \beta_i$.

(ii) $\Rightarrow$ (i). For every $i \in \{1, \ldots, t\}$, let $Y_i$ and $X_i$ be positive diagonal matrices such that $\det(Y_iX_i) = 1$ and $Y_iA_iX_i = M_i$. By Theorem 3.2, the matrix $C = \sum_{i=1}^{t} Y_iA_iX_i = \sum_{i=1}^{t} M_i$ satisfies $\|C\| = \sum_{i=1}^{t} \beta_i$. ■

In order to characterize the case in which the infimum in Theorem 4.13 is attained, we introduce some further notation.

4.22. NOTATION. Let $A$ be an $n \times n$ matrix. We denote by $A^{\$}$ the matrix obtained from $A$ by setting equal to 0 all elements that do not lie on a strictly nonzero cycle.

4.23. EXAMPLE. Let $A$ be the matrix of Example 4.4. The strictly nonzero cycles in $A$ are $(1,2)$, $(1,3)$, $(1,2,3)$, and $(3)$. Since every nonzero element of $A$ lies on at least one of these cycles, it follows that $A^{\$} = A$.

4.24. COMMENT. (i) Let $A$ be an $n \times n$ matrix and let $P$ be a permutation matrix such that $C = PAP^T$ is in Forbenius normal form. It is easy to verify that $C^{\$} = PA^{\$}P^T$ is the matrix obtained from $C$ by setting equal to 0 all the off-diagonal blocks. It thus follows that for all diagonal matrices $Y$ and $X$, the matrices $YAX$ and $YA^{\$}X$ share the same spectrum.

(ii) Let $A$ be an $n \times n$ matrix. It follows from Notation 4.3 and 4.22 that $A^{\$} \geqslant A^{\#}$.

(iii) It is easy to verify that the matrices $A^{\$}$ and $A^{\#}$ are completely reducible.

4.25. PROPOSITION. *Let $A$ and $B$ be completely reducible nonnegative $n \times n$ matrices satisfying $B \geqslant A$, and assume that all components of $A$ share the same spectral radius. Then $\rho(B) = \rho(A)$ if and only if $B = A$.*

*Proof.* Obviously, all we have to prove is that if all components of $A$ share the same spectral radius and if $\rho(B) = \rho(A)$, then $B = A$. Partition $A$ conformably with the Frobenius normal form of $B$. Observe that every component $(B)_{ii}$ of $B$ corresponds to a direct sum $(A)_{ii}$ of components of $A$. Since all components of $A$ share the same spectral radius and since $\rho(B) = \rho(A)$, it follows that

$$\rho((B)_{ii}) \leqslant \rho(B) = \rho(A) = \rho((A)_{ii}). \tag{4.26}$$

Since $(B)_{ii} \geqslant (A)_{ii}$ and since $(B)_{ii}$ is irreducible, we deduce from the Perron–Frobenius spectral theory for nonnegative matrices that

$$\rho((B)_{ii}) \geqslant \rho((A)_{ii}), \quad \text{where equality holds if and only if } (B)_{ii} = (A)_{ii}. \tag{4.27}$$

It now follows from (4.26) and (4.27) that $(B)_{ii} = (A)_{ii}$, and hence $B = A$. ∎

4.28. THEOREM. *Let $A_1, \ldots, A_t$ be nonnegative $n \times n$ matrices and let $\beta_i$ be the row (and column) sum of the (unique) generalized doubly stochastic matrix $M_i$ in $R(A_i^{\#})$, $i \in \{1, \ldots, t\}$. Then the following are equivalent.*

(i) *There exists a matrix $C$ in $R(A_1, \ldots, A_t)$ for which $\rho(C) = \sum_{i=1}^t \beta_i$.*

(ii) *We have $(\sum_{i=1}^t A_i)^{\$} = \sum_{i=1}^t A_i^{\$}$ and $A_i^{\$} = A_i^{\#}$ for all $i \in \{1, \ldots, t\}$.*

*Proof.* (i) $\Rightarrow$ (ii). Let $C = \sum_{i=1}^t Y_i A_i X_i$ be a matrix satisfying $\rho(C) = \sum_{i=1}^t \beta_i$, where $Y_i$ and $X_i$ are positive diagonal matrices such that $\det(Y_i X_i) = 1$, $i = 1, \ldots, t$. Since

$$C \geqslant C^{\$} \geqslant \sum_{i=1}^{t} Y_i A_i^{\$} X_i \geqslant \sum_{i=1}^{t} Y_i A_i^{\#} X_i, \tag{4.29}$$

it follows by the Perron–Frobenius theorem that

$$\rho(C) \geqslant \rho(C^{\$}) \geqslant \rho\left(\sum_{i=1}^{t} Y_i A_i^{\$} X_i\right) \geqslant \rho\left(\sum_{i=1}^{t} Y_i A_i^{\#} X_i\right). \tag{4.30}$$

Note that $A_i^{\#} = \bar{Y}_i M_i \bar{X}_i$ for some positive diagonal matrices $\bar{Y}_i$ and $\bar{X}_i$ satisfying $\det(\bar{Y}_i \bar{X}_i) = 1$. Thus, we have that $Y_i A_i^{\#} X_i = \tilde{Y}_i M_i \tilde{X}_i$, where $\tilde{Y}_i$ and $\tilde{X}_i$ are the positive diagonal matrices $Y_i \bar{Y}_i$ and $\bar{X}_i X_i$ respectively, satisfying $\det(\tilde{Y}_i \tilde{X}_i) = 1$. It now follows by Theorem 3.11 that

$$\rho\left(\sum_{i=1}^{t} Y_i A_i^{\#} X_i\right) \geqslant \sum_{i=1}^{t} \beta_i. \tag{4.31}$$

Since $\rho(C) = \sum_{i=1}^t \beta_i$, it now implies by (4.30) and (4.31) that

$$\rho(C^{\$}) = \rho\left(\sum_{i=1}^{t} Y_i A_i^{\$} X_i\right) = \rho\left(\sum_{i=1}^{t} Y_i A_i^{\#} X_i\right) \tag{4.32}$$

and

$$\rho\left(\sum_{i=1}^{t} Y_i A_i^{\#} X_i\right) = \sum_{i=1}^{t} \beta_i. \tag{4.33}$$

By Theorem 3.11, the equality (4.33) yields that for every $i \in \{1, \ldots, t\}$, $Y_i A_i^{\#} X_i$ is diagonally similar to the generalized doubly stochastic matrix $M_i$.

Hence all components of $Y_i A_i^{\#} X_i$ share the same spectral radius $\beta_i$. In view of (4.29) and since the matrices $C^{\$}$, $\Sigma_{i=1}^{t} Y_i A_i^{\$} X_i$, and $\Sigma_{i=1}^{t} Y_i A_i^{\#} X_i$ are completely reducible, it follows from (4.32) by Proposition 4.25 that $C^{\$} = \Sigma_{i=1}^{t} Y_i A_i^{\$} X_i = \Sigma_{i=1}^{t} Y_i A_i^{\#} X_i$, which clearly implies (ii).

(ii) $\Rightarrow$ (i). For every $i \in \{1, \dots, t\}$ let $Y_i$ and $X_i$ be positive diagonal matrices such that $\det(Y_i X_i) = 1$ and $Y_i A_i^{\#} X_i = M_i$. We now see that

$$\rho\left(\sum_{i=1}^{t} Y_i A_i X_i\right) = \rho\left(\left(\sum_{i=1}^{t} Y_i A_i X_i\right)^{\$}\right) = \rho\left(\sum_{i=1}^{t} Y_i A_i^{\$} X_i\right)$$

$$= \rho\left(\sum_{i=1}^{t} Y_i A_i^{\#} X_i\right) = \rho\left(\sum_{i=1}^{t} M_i\right).$$

By Theorem (3.11), it now follows that $\rho(\Sigma_{i=1}^{t} Y_i A_i X_i) = \Sigma_{i=1}^{t} \beta_i$. ■

In view of Theorems 4.18 and 4.28 it would be interesting to check the relations between the equality cases $A^{\$} = A^{\#}$ and $A^{\#} = A$. By Comment 4.24.ii, the equality $A^{\#} = A$ implies that $A^{\$} = A^{\#}$. However, the converse is, in general, false, as is demonstrated by the following example.

4.34. EXAMPLE. Let

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Since $A$ has no strictly nonzero cycle, it follows that $A^{\$} = A^{\#} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ and so, while $A^{\$} = A^{\#}$, we have that $A^{\#} \neq A$.

*The authors are grateful to Professor Hans Weinberger for his comments, which have helped to improve the paper.*

## REFERENCES


1 R. B. Bapat, Two inequalities for the Perron root, *Linear Algebra Appl.* 85:241–248 (1987).

2 G. Birkhoff, Tres observaciones sobre el algebra linea, *Univ. Nac. Rev. Ser. A* 5:147–150 (1946).

3 R. A. Brualdi, Matrices permutation equivalent to irreducible matrices and applications, *Linear Multilinear Algebra* 7:1–12 (1979).

4 R. A. Brualdi and H. J. Ryser, *Combinatorial Matrix Theory*, Cambridge Univ. Press, Cambridge, 1991.
5 R. A. Brualdi and H. Wielandt, A spectral characterization of stochastic matrices, *Linear Algebra Appl.* 1:65–72 (1968).
6 S. Friedland and S. Karlin, Spectral radius of non-negative matrices, *Duke Math. J.* 62:459–490 (1975).
7 G. Frobenius, *Über zerlegbare Determinanten*, pp. 274–277, Sitzungsber. Preuss, Akad. Wiss., Berlin, 1917.
8 G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, 2nd ed., Cambridge Univ. Press, Cambridge, 1952.
9 R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge Univ. Press, Cambridge, 1985.
10 D. König, Über Graphen und ihre Anwenedung auf Determinanten und Matrizen, *Math. Ann.* 77:453–456 (1916).
11 M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Allyn & Bacon, Boston, 1964.
12 R. Sinkhorn and P. Knopp, Concerning nonnegative matrices and doubly stochastic matrices, *Pacific J. Math.* 21:343–348 (1967).
13 H. Wielandt, Unzerlegbare, nicht negative Matrizen, *Math. Z.* 52:642–645 (1950).

*Received 29 December 1994; final manuscript accepted 22 May 1995*