# Minimum permanents on two faces of the polytope of doubly stochastic matrices*

**Kyle Pula**[a] **Seok-Zun Song**[b] **Ian M. Wanless**[c]

[a] Department of Mathematics
University of Denver, Denver, CO 80208 USA

jpula@math.du.edu

[b] Department of Mathematics
Jeju National University, Jeju 690-756 Republic of Korea

szsong@jejunu.ac.kr

[c] School of Mathematical Sciences
Monash University, Vic 3800 Australia

ian.wanless@monash.edu

**Abstract**

We consider the minimum permanents and minimising matrices on the faces of the polytope of doubly stochastic matrices whose nonzero entries coincide with those of, respectively,

$$U_{m,n} = \begin{bmatrix} I_n & J_{n,m} \\ J_{m,n} & 0_m \end{bmatrix} \qquad \text{and} \qquad V_{m,n} = \begin{bmatrix} I_n & J_{n,m} \\ J_{m,n} & J_{m,m} \end{bmatrix}.$$

We conjecture that $V_{m,n}$ is cohesive but not barycentric for $1 < n < m + \sqrt{m}$ and that it is not cohesive for $n \geqslant m + \sqrt{m}$. We prove that it is cohesive for $1 < n < m + \sqrt{m}$ and not cohesive for $n \geqslant 2m$ and confirm the conjecture computationally for $n < 2m \leqslant 200$. We also show that $U_{m,n}$ is barycentric.

*Keywords:* permanent, doubly stochastic, cohesive, barycentric.

*AMS Subject Classification:* 15A15.

## 1 Introduction and preliminaries

Let $\Omega_n$ be the polytope of $n \times n$ doubly stochastic matrices, that is, the $n \times n$ nonnegative matrices whose row and column sums are all equal to 1. The *permanent*

*This research was undertaken while the first and second authors were visiting Monash University. The first author's visit was supported by the Australian Academy of Science, the Australian-American Fulbright Commission, and the National Science Foundation. The second author's visit was supported by the Korean Science and Engineering Foundation and the Australian Academy of Science. The third author's research is supported by ARC discovery grant DP1093320.

of an $n \times n$ matrix $A = [a_{ij}]$ is defined by

$$\operatorname{per} A = \sum_{\sigma} a_{1\sigma(1)} \cdots a_{n\sigma(n)}$$

where $\sigma$ runs over all permutations of $\{1, 2, \ldots, n\}$.

Let $D = [d_{ij}]$ be an $n \times n$ nonnegative matrix with $\operatorname{per} D > 0$, and let

$$\Omega(D) = \big\{ [x_{ij}] \in \Omega_n \mid x_{ij} = 0 \text{ whenever } d_{ij} = 0 \big\}.$$

Then $\Omega(D)$ is a face of $\Omega_n$, and since it is non-empty and compact, $\Omega(D)$ contains at least one *minimising matrix* $Y$ such that $\operatorname{per} Y \leqslant \operatorname{per} X$ for all $X \in \Omega(D)$.

Let $J_{r,s}$ denote the $r \times s$ matrix all of whose entries are 1, $I_n$ the identity matrix of order $n$ and $0_m$ the $m \times m$ zero matrix. In this paper, we study minimising matrices on the faces $\Omega(U_{m,n})$ and $\Omega(V_{m,n})$, where

$$U_{m,n} = \begin{bmatrix} I_n & J_{n,m} \\ J_{m,n} & 0_m \end{bmatrix} \qquad \text{and} \qquad V_{m,n} = \begin{bmatrix} I_n & J_{n,m} \\ J_{m,n} & J_{m,m} \end{bmatrix}.$$

If $\operatorname{per} D > 0$ then the *barycenter* $b(D)$ of $\Omega(D)$ is given by

$$b(D) = \frac{1}{\operatorname{per} D} \sum_{P \leqslant D} P$$

where the summation extends over the set of all permutation matrices $P$ with $P \leqslant D$. Brualdi [1] defined an $n$-square $(0,1)$ matrix $D$ to be *cohesive* if there is a matrix $Z$ in the interior of $\Omega(D)$ for which $\operatorname{per} Z = \min\big\{ \operatorname{per} X \mid X \in \Omega(D) \big\}$. He defined an $n$-square $(0,1)$ matrix $D$ to be *barycentric* if $\operatorname{per} b(D) = \min\big\{ \operatorname{per} X \mid X \in \Omega(D) \big\}$.

Since $b(D)$ always falls in the interior of $\Omega(D)$, being barycentric is a stronger property than being cohesive. We will consider the question of which values of parameters $m, n$ make $U_{m,n}$ and $V_{m,n}$ barycentric, cohesive, or neither. In §2 we prove that $V_{m,n}$ is cohesive for $1 < n < m + \sqrt{m}$ and not cohesive for $n \geqslant 2m$. In §3 we show that $U_{m,n}$ is barycentric.

Our results contribute towards solution of two problems from Minc's well-known catalogue of unsolved problems on permanents (see [2] for the most recent update). Problems 14 and 15 in Minc's list ask for a characterisation of cohesive and barycentric matrices respectively. These problems were originally posed by Brualdi [1], who determined the minimising matrix on $\Omega(V_{1,n})$. Minc had resolved the face $\Omega(V_{m,2})$ in [4]. Song determined the minimum permanent on $\Omega(V_{m,3})$ in [5, 7], while the faces $\Omega(V_{2,n})$ and $\Omega(V_{3,n})$ were resolved by Song in [6] and Song et al in [8, 9], respectively. Taken together, the prior literature determines the minimising matrices of $\Omega(V_{m,n})$ for $m < 4$ or $n < 4$.

Recall that an $n \times n$ nonnegative matrix is said to be *fully indecomposable* if it contains no $k \times (n-k)$ zero submatrix for $1 \leqslant k < n$. We will use the following well-known Lemma from [3].

**Lemma 1.1** *Let $D = [d_{ij}]$ be an $n \times n$ fully indecomposable $(0, 1)$ matrix, and suppose $Y = [y_{ij}]$ is a minimising matrix on $\Omega(D)$. Then $Y$ is fully indecomposable and*

$$\begin{aligned} \operatorname{per} Y(i \mid j) &= \operatorname{per} Y \quad \text{if } d_{ij} = 1 \text{ and } y_{ij} > 0, \\ \operatorname{per} Y(i \mid j) &\geqslant \operatorname{per} Y \quad \text{if } d_{ij} = 1 \text{ and } y_{ij} = 0. \end{aligned}$$

As usual, for any matrix $M$ and lists $L_1$ and $L_2$ of row and column indices respectively, $M(L_1 \mid L_2)$ denotes the submatrix formed by omitting the rows $L_1$ and columns $L_2$ from $M$.

Terms of the form $0^0$ occurring in our calculations should always be interpreted as 1.

# 2 The Minimising Matrices of $\Omega(V_{m,n})$

In this section, we consider the minimum permanents and minimising matrices on the faces $\Omega(V_{m,n})$. Throughout this section, we assume that $m, n \geqslant 2$.

Let $Y$ be a minimising matrix on $\Omega(V_{m,n})$. Since the last $m$ rows and last $m$ columns of $V_{m,n}$ are the same, we can use the averaging method on those rows and columns of $Y$ (by Theorem 1 in Minc [4]). Without loss of generality, we may therefore assume $Y = [y_{ij}]$ is a minimising matrix of the form:

$$y_{ij} = \begin{cases} x_i & \text{if } i = j \leqslant n, \\ 0 & \text{if } i \leqslant n \text{ and } j \leqslant n \text{ and } i \neq j, \\ a_i & \text{if } i \leqslant n \text{ and } j > n, \\ a_j & \text{if } i > n \text{ and } j \leqslant n, \\ x & \text{if } i > n \text{ and } j > n. \end{cases} \tag{1}$$

Note that $Y$ is doubly stochastic so $x_i = 1 - ma_i$ for $1 \leqslant i \leqslant n$. Since $V_{m,n}$ is fully indecomposable, it follows from Lemma 1.1 that $Y$ is also fully indecomposable. In particular $a_i > 0$ for all $i$, although it is plausible that $x_i = 0$ for some $i$ or that $x = 0$.

We next consider the possible choices of $\{a_i\}$ in (1), i.e. those that minimise $\operatorname{per} Y$.

**Theorem 2.1** *The minimising matrix $Y$ has $a_1 = a_2 = \cdots = a_n$.*

**Proof.** Without loss of generality we assume that $a_1 \leqslant a_2 \leqslant \cdots \leqslant a_n$ and hence $x_1 \geqslant x_2 \geqslant \cdots \geqslant x_n$. Aiming for a contradiction, assume that $a_n > a_1$. Let

$$\begin{aligned} p_0 &= \operatorname{per} Y(1, n \mid 1, n), \\ p_1 &= \operatorname{per} Y(1, n, n+1 \mid 1, n, n+1) \\ p_2 &= \operatorname{per} Y(1, n, n+1, n+2 \mid 1, n, n+1, n+2) \end{aligned}$$

If $x_n > 0$ then by Lemma 1.1, we have that

$$x_1 p_0 + m^2 a_1^2 p_1 = \operatorname{per} Y(n \mid n) = \operatorname{per} Y = \operatorname{per} Y(1 \mid 1) = x_n p_0 + m^2 a_n^2 p_1$$

and hence

$$p_0 = \frac{m^2(a_n^2 - a_1^2)}{x_1 - x_n} p_1 = m(a_1 + a_n)p_1. \tag{2}$$

Also, as $a_1 > 0$,

$$\begin{aligned} \operatorname{per} Y &= \operatorname{per} Y(1 \mid n+1) \\ &= ma_1 \operatorname{per} Y(1, n+1 \mid 1, n+1) \\ &= ma_1\big(x_n p_1 + (m-1)^2 a_n^2 p_2\big). \end{aligned}$$

and similarly $\operatorname{per} Y = ma_n\big(x_1 p_1 + (m-1)^2 a_1^2 p_2\big)$, which leads to

$$(m-1)^2 p_2 = \frac{a_n x_1 - a_1 x_n}{a_1 a_n(a_n - a_1)} p_1 = \frac{p_1}{a_1 a_n}.$$

However, expanding $\operatorname{per} Y$ along the first and $n$-th rows we find

$$\begin{aligned} \operatorname{per} Y &= x_1 x_n p_0 + m^2 a_n^2 x_1 p_1 + m^2 a_1^2 x_n p_1 + m^2 (m-1)^2 a_1^2 a_n^2 p_2 \qquad (3) \\ &= (1 - ma_1)(1 - ma_n) m(a_1 + a_n) p_1 + m^2 a_n^2 (1 - ma_1) p_1 \\ &\quad + m^2 a_1^2 (1 - ma_n) p_1 + m^2 a_1 a_n p_1 \\ &= mp_1(a_1 + a_n - ma_1 a_n). \qquad (4) \end{aligned}$$

Note that although (2) is only valid for $x_n > 0$ we are free to substitute it in (3) in the case $x_n = 0$ as well, since in that case $p_0$ is being multiplied by 0. Examining (4) we see that by varying $a_1, a_n$ while preserving $a_1 + a_n$ we could decrease $\operatorname{per} Y$ unless $a_1 = a_n$. By assumption, $Y$ is a minimising matrix so $a_1 = a_n$, from which the result follows. ∎

In light of Theorem 2.1, for any given values of $m$ and $n$, we are left with a one variable optimisation to find the minimising matrix in $\Omega(V_{m,n})$ since $Y$ is now determined by the value of $x$ in (1). Let $A_x = [a_{ij}]$ be the $(n+m) \times (n+m)$ matrix defined by

$$a_{ij} = \begin{cases} \frac{1}{n}(n - m + m^2 x) & \text{if } i = j \leqslant n, \\ 0 & \text{if } i \leqslant n \text{ and } j \leqslant n \text{ and } i \neq j, \\ x & \text{if } i > n \text{ and } j > n, \\ \frac{1}{n}(1 - mx) & \text{otherwise.} \end{cases} \tag{5}$$

Note that $A_x$ is doubly stochastic provided $x \in [0, \frac{1}{m}]$ when $m \leqslant n$, or $x \in [\frac{m-n}{m^2}, \frac{1}{m}]$ in the case $m > n$. We next consider the problem of finding $x$ in the stated range that minimises $\operatorname{per} A_x$ (and hence satisfies $\operatorname{per} A_x = \operatorname{per} Y$).

**Theorem 2.2** *For $n \geqslant 2m$ and $x > 0$ we have $\operatorname{per} A_x > \operatorname{per} A_0$ and thus $A_0$ is the unique minimising matrix in $\Omega(V_{m,n})$. In contrast, for $n < m + \sqrt{m}$ it is never the case that $A_0$ is a minimising matrix in $\Omega(V_{m,n})$.*

**Proof.** If $m > n$ then $A_0$ is not even doubly stochastic, and if $m = n$ then $A_0$ is not fully indecomposable. So by Lemma 1.1 we may assume that $n > m$. From (5) we have

$$\begin{aligned}\operatorname{per} A_x &= \sum_{i=0}^{m} \binom{m}{i}^2 \frac{i!n!(m-i)!}{(n-m+i)!} x^i \left(\tfrac{1}{n}(1-mx)\right)^{2m-2i} \left(\tfrac{1}{n}(n-m+m^2x)\right)^{n-m+i} \\ &= \sum_{i=0}^{m} \binom{m}{i} \frac{n!m!}{(n-m+i)!\,n^{n+m-i}} x^i (1-mx)^{2m-2i} (n-m+m^2x)^{n-m+i}. \qquad (6)\end{aligned}$$

In particular,

$$\operatorname{per} A_0 = \frac{n!\,m!\,(n-m)^{n-m}}{(n-m)!\,n^{n+m}}, \qquad (7)$$

and hence

$$\begin{aligned}\frac{\operatorname{per} A_x}{\operatorname{per} A_0} &= \sum_{i=0}^{m} \binom{m}{i} \frac{n^i(n-m)!}{(n-m+i)!} x^i (1-mx)^{2m-2i} \left(1 + \frac{m^2x}{n-m}\right)^{n-m} (n-m+m^2x)^i \\ &\geqslant \left(1 + \frac{m^2x}{n-m}\right)^{n-m} \sum_{i=0}^{m} \binom{m}{i} (1-mx)^{2m-2i} \left(nx - mx + m^2x^2\right)^i \\ &= \left(\left(1 + \frac{m^2x}{n-m}\right)^{\frac{n-m}{m}}\right)^m \left((1-mx)^2 + nx - mx + m^2x^2\right)^m \\ &\geqslant (1+mx)^m (1-mx+2m^2x^2)^m \\ &= (1 + m^2x^2 + 2m^3x^3)^m,\end{aligned}$$

whenever $n \geqslant 2m$. The first statement of the theorem follows.

Next, consider $x \to 0$ in (6), where

$$\begin{aligned}\operatorname{per} A_x &= \frac{n!m!(1-mx)^{2m}(n-m+m^2x)^{n-m}}{(n-m)!n^{n+m}} + \frac{m^2n!(m-1)!(n-m)^{n-m+1}}{(n-m+1)!n^{n+m-1}} x + O(x^2) \\ &= \frac{n!m!(n-m)^{n-m}}{(n-m)!n^{n+m}} \left[1 - m^2x + \frac{mn(n-m)}{(n-m+1)} x\right] + O(x^2).\end{aligned}$$

It follows that when $n(n-m) < (n-m+1)m$ (or in other words, $n < m + \sqrt{m}$), $\operatorname{per} A_x < \operatorname{per} A_0$ for small positive $x$. This proves the second statement in the Theorem. ■

For $m \leqslant n$ we know $x \in [0, 1/m]$. Having examined the situation at the lower end of that interval, we now turn our attention to the upper end.

**Theorem 2.3** *$A_{1/m}$ is not a minimising matrix of $\Omega(V_{m,n})$ for $m \leqslant n$.*

**Proof.** From (6) there is a polynomial $q(x)$ such that

$$\operatorname{per} A_x = m!\,n^{-n} x^m (n-m+m^2x)^n + (1-mx)^2\,q(x).$$

Thus the derivative of $\operatorname{per} A_x$ at $x = 1/m$ is $2m!\,m^{2-m}$. Not only is $\operatorname{per} A_x$ increasing at $x = 1/m$, but in fact the rate of increase depends only on $m$. ■

Next we consider similar questions for the case $m \geqslant n$.

**Theorem 2.4** *$V_{m,n}$ is cohesive for $m \geqslant n$.*

**Proof.** From (5), we have

$$\operatorname{per} A_x = \frac{m!^2}{n^{2n}} \sum_{i=0}^{n} \binom{n}{i} \frac{n^i}{(m-n+i)!} (n-m+m^2x)^i (1-mx)^{(2n-2i)} x^{m-n+i} \tag{8}$$

for $x \in [\frac{m-n}{m^2}, \frac{1}{m}]$. In particular, we have

$$\operatorname{per} A_{(m-n)/m^2} = \frac{m!^2 (m-n)^{m-n}}{(m-n)! m^{2m}} \quad \text{and} \quad \operatorname{per} A_{1/m} = \frac{m!}{m^m}.$$

Therefore,

$$\frac{\operatorname{per} A_{1/m}}{\operatorname{per} A_{(m-n)/m^2}} = \frac{m^m (m-n)!}{m!(m-n)^{m-n}} > 1$$

for $0 < n \leqslant m$, and $A_{1/m}$ cannot be a minimising matrix.

Now consider that

$$\begin{aligned} \operatorname{per} A_x = {} & \frac{m!}{n^{2n}} \left( \frac{(1-mx)^{2n} x^{m-n}}{(m-n)!} + \frac{n^2 (1-mx)^{2n-2} x^{m-n+1}}{(m-n+1)!} (n-m+m^2x) \right) \\ & + (n-m+m^2x)^2 \, r(x), \end{aligned}$$

for some polynomial $r(x)$. Therefore the derivative of $\operatorname{per} A_x$ at $x = (m-n)/m^2$ is

$$-\frac{m!}{(m-n+1)!} (m-n)^{m-n} m^{2-2m}.$$

In particular, it is negative so $A_{(m-n)/m^2}$ is not a minimising matrix either. ■

Combining Theorems 2.2, 2.3 and 2.4, we have:

**Corollary 2.5** *$V_{m,n}$ is cohesive for $n < m + \sqrt{m}$ but not for $n \geqslant 2m$.*

For the cases not covered by this corollary, i.e. $m + \sqrt{m} \leqslant n < 2m$, we have demonstrated that $\operatorname{per} A_x$ is increasing at both end points of the interval $[\frac{m-n}{m^2}, \frac{1}{m}]$ but it remains to be determined whether the minimum actually occurs at $x = (m-n)/m^2$. As reported below, we have investigated this question computationally for $m \leqslant 100$.

For $n \geqslant m$, the barycenter of $V_{m,n}$ is located at $A_\beta$ where

$$\beta = \frac{\sum_{i=1}^{m} \binom{n}{n-m+i} \binom{m-1}{i-1}^2 (i-1)!(m-i)!^2}{\sum_{i=0}^{m} \binom{n}{n-m+i} \binom{m}{i}^2 (i-1)!(m-i)!^2} = \frac{\sum_{i=0}^{m} i \big((n-m+i)!(m-i)!i!\big)^{-1}}{m^2 \sum_{i=0}^{m} \big((n-m+i)!(m-i)!i!\big)^{-1}}.$$

The value of $\operatorname{per} b(V_{m,n})$ can then be calculated from (6). For $n < m$, the barycenter is located at $A_{\beta'}$ where

$$\beta' = \frac{m-n}{m^2} + \frac{\sum_{i=0}^{n} i \big((m-n+i)!(n-i)!i!\big)^{-1}}{m^2 \sum_{i=0}^{n} \big((m-n+i)!(n-i)!i!\big)^{-1}}$$

and the value of $\operatorname{per} b(V_{m,n})$ can then be calculated from (8).

In order to investigate small cases not covered by the preceding theory, two of the authors independently wrote programs for the computer algebra systems Maple and Mathematica. The results of their computations agreed and are as follows.

Let $P(x) = \operatorname{per} A_x$ and let $P'(x)$ denote its derivative. For $2 \leqslant n < m \leqslant 100$, we found that $P'(x)$ has no rational roots in the interval $[(m-n)/m^2, \infty)$ and that $P(x)$ is increasing at the barycenter. Either of these facts shows that $V_{m,n}$ is not barycentric in these cases although we know from Corollary 2.5 that it is cohesive.

For $2 \leqslant m \leqslant n < 2m \leqslant 200$, we found that $P(x)$ is increasing at the barycenter and that $P'(x)$ has no non-negative rational roots when $n \neq m + \sqrt{m}$. If $n = m + \sqrt{m}$ the only non-negative rational root is $x = 0$. Again, either fact shows that $V_{m,n}$ is not barycentric.

For $2 < m + \sqrt{m} \leqslant n < 2m \leqslant 200$, $P(x)$ is monotone increasing throughout the interval $[0, 1/m]$. Once again, the only case in which $P'(x)$ has a root in this interval is for $n = m + \sqrt{m}$ and this root occurs at $x = 0$.

Taken together with Corollary 2.5, this data suggests the following conjecture.

**Conjecture 2.1** *$V_{m,n}$ is cohesive but not barycentric for $1 < n < m + \sqrt{m}$, while for $n \geqslant m + \sqrt{m}$, $V_{m,n}$ is not cohesive and $A_0$ is a minimising matrix.*

# 3 The face $\Omega(U_{m,n})$

We finish by determining the minimum permanent and minimising matrix on the face $\Omega(U_{m,n})$. Note that for $\Omega(U_{m,n})$ to be non-empty we require $n \geqslant m$.

Relying on Brualdi [1] for the case $m = 1$ and using a proof identical to that of Theorem 2.1 for $m \geqslant 2$ we get:

**Theorem 3.1** *For any $n \geqslant m$ the unique minimising matrix in $\Omega(U_{m,n})$ is $A_0$.*

By symmetry it is obvious that $b(U_{m,n}) = A_0$ and thus we also have:

**Corollary 3.2** *$U_{m,n}$ is barycentric for any $n \geqslant m$.*

The minimum permanent is given by (7).

**Corollary 3.3** *For any $n \geqslant m$ the minimum permanent in $U_{m,n}$ is*

$$\operatorname{per} A_0 = \operatorname{per} b(U_{m,n}) = \frac{n!\, m!\, (n-m)^{n-m}}{(n-m)!\, n^{n+m}}.$$

For example, the minimum permanent on $\Omega(U_{4,n})$ is

$$\operatorname{per} b(U_{4,n}) = 4! \cdot \frac{(n-1)(n-2)(n-3)(n-4)^{n-4}}{n^{n+3}},$$

which is also the minimum permanent on $\Omega(V_{4,n})$ for $n \geqslant 6$.

# References


[1] R. A. Brualdi, An interesting face of the polytope of doubly stochastic matrices *Linear and Multilinear Algebra* **17** (1985), 5–18.

[2] G.-S. Cheon and I. M. Wanless, An update on Minc's survey of open problems involving permanents, *Linear Algebra Appl.* **403** (2005), 314–342.

[3] T. H. Foregger, On the minimum value of the permanent of a nearly decomposable doubly stochastic matrix, *Linear Algebra Appl.* **32** (1980), 75–85.

[4] H. Minc, Minimum permanents of doubly stochastic matrices with prescribed zero entries, *Linear and Multilinear Algebra* **15** (1984), 225–243.

[5] S.-Z. Song, Minimum permanents on certain faces of matrices containing an identity submatrix, *Linear Algebra Appl.* **108** (1988), 263–280.

[6] S.-Z. Song, Minimum permanents on certain doubly stochastic matrices, *Linear Algebra Appl.* **143** (1991), 49–56.

[7] S.-Z. Song, Minimum permanents and minimizing matrices on certain doubly stochastic matrices, *Linear Algebra Appl.* **162-164** (1992), 786–790.

[8] S.-Z. Song, Minimum permanents on a face of the polytope of doubly stochastic matrices, *Linear Algebra Appl.* **259** (1997), 169–182.

[9] S.-Z. Song, S.-G. Hwang, G.-S. Cheon, Minimum permanents on a face of the polytope of doubly stochastic matrices II, *Linear and Multilinear Algebra* **50** (2002), 101–111.