Doubly stochastic matrices and the Bruhat order

Article *in* Czechoslovak Mathematical Journal · October 2016

DOI: 10.1007/s10587-016-0286-6

Richard A. Brualdi
University of Wisconsin–Madison

Geir Dahl
University of Oslo

Czechoslovak Mathematical Journal, 66 (141) (2016), 681–700

DOUBLY STOCHASTIC MATRICES AND THE BRUHAT ORDER

Richard A. Brualdi, Madison, Geir Dahl, Oslo,
Eliseu Fritscher, Rio de Janeiro

(Received September 26, 2015)

This article is dedicated to the memory of Professor Miroslav Fiedler

Abstract. The Bruhat order is defined in terms of an interchange operation on the set of permutation matrices of order n which corresponds to the transposition of a pair of elements in a permutation. We introduce an extension of this partial order, which we call the stochastic Bruhat order, for the larger class Ω_n of doubly stochastic matrices (convex hull of $n \times n$ permutation matrices). An alternative description of this partial order is given. We define a class of special faces of Ω_n induced by permutation matrices, which we call Bruhat faces. Several examples of Bruhat faces are given and several results are presented.

Keywords: Bruhat order; doubly stochastic matrix; face

MSC 2010: 05B20, 06A07, 15B51

1. Introduction

Let $\mathcal{S}_n$ denote the symmetric group of order n consisting of all permutations of $\{1, 2, \ldots, n\}$. With each permutation $\sigma \in \mathcal{S}_n$, there is a corresponding $n \times n$ permutation matrix $P = [p_{ij}]$, where $p_{ij} = 1$ if and only if $j = \sigma(i)$. Let $\mathcal{P}_n$ denote the set of all $n \times n$ permutation matrices. The *Bruhat order* on $\mathcal{S}_n$ in terms of $\mathcal{P}_n$ is the partial order $\preceq_B$ defined as $P \preceq_B Q$ provided that P can be obtained from Q by a sequence of *backward interchanges*, that is, replacing 2×2 submatrices equal to L_2 with I_2 as shown below:

$$L_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \longrightarrow I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Eliseu Fritscher is partially supported by CNPq (Brazil)—Grant 150521/2015-4.

It follows that the identity matrix I_n is the unique minimal element (no backward interchanges possible) and the anti-identity matrix L_n is the unique maximal element ($\binom{n}{2}$ backward interchanges possible) of the Bruhat order on $\mathcal{P}_n$.

For an $m \times n$ matrix $A = [a_{ij}]$ we define the $m \times n$ matrix

$$\Sigma(A) = [\sigma_{ij}(A)], \quad \text{where } \sigma_{ij}(A) = \sum_{1 \leqslant k \leqslant i,\, 1 \leqslant l \leqslant j} a_{kl}, \quad 1 \leqslant i \leqslant m, \text{ and } 1 \leqslant j \leqslant n.$$

The Bruhat order on $\mathcal{P}_n$ may be characterized as follows. For $m \times n$ matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ we write $A \geqslant B$ (or $B \leqslant A$) to denote entrywise inequality.

The following result is known; see Theorem 2.1.5 in [1] or Lemma 7 of [7].

Theorem 1.1 ([1], [7])**.** *Let $P, Q \in \mathcal{P}_n$. Then $P \preceq_B Q$ if and only if $\Sigma(P) \geqslant \Sigma(Q)$.*

An improved version of this characterization was shown in [2]. The Bruhat order for the class of $(0,1)$-matrices with given row and column sums was investigated in [5], [6].

Recall that a square matrix is *doubly stochastic* provided it is nonnegative and each row and column sum is 1. We let Ω_n denote the set of doubly stochastic matrices of order n. Then Ω_n is a convex polytope of dimension $(n-1)^2$, often called the *Birkhoff polytope*, whose set of vertices is $\mathcal{P}_n$. Let $A_1, A_2 \in \Omega_n$. If $\Sigma(A_1) \geqslant \Sigma(A_2)$, we write $A_1 \preceq_B A_2$. This is a partial order on Ω_n, which we call the *stochastic Bruhat order*. Due to Theorem 1.1, the stochastic Bruhat order on Ω_n, when restricted to $\mathcal{P}_n$, reduces to the Bruhat order on $\mathcal{P}_n$.

The goal of this paper is to investigate properties of the stochastic Bruhat order and related subpolytopes of Ω_n.

A vector $x = (x_1, x_2, \ldots, x_n)$ is *non-decreasing* if $x_1 \leqslant x_2 \leqslant \ldots \leqslant x_n$. The *support* of an $m \times n$ matrix $A = [a_{ij}]$ is the set $\operatorname{supp} A = \{(i,j)\colon\ a_{ij} \neq 0\}$. An $n \times n$ matrix A has *total support* if each of its nonzero elements lies in a nonzero diagonal of A (a permutation set of places occupied by nonzeros of A). The *convex hull* of a set S is denoted by $\operatorname{conv} S$. We recall some notions and results from [4]. Let $P = [p_{ij}]$ be a permutation matrix of order n corresponding to a permutation $\sigma = (i_1, i_2, \ldots, i_n)$ of $\{1, 2, \ldots, n\}$. The *Bruhat shadow* $\mathcal{S}(P)$ of P is the $(0,1)$-matrix of order n whose support equals the union of the supports of all permutation matrices Q satisfying $Q \preceq_B P$, i.e., $\mathcal{S}(P)$ is the Boolean sum of these matrices. Define the *left-sequence*[1] of P as $l(P) = l_1, l_2, \ldots, l_n$, where l_k is the largest integer in the set $\{i_1, i_2, \ldots, i_k\}$ of integers ($k = 1, 2, \ldots, n$). Similarly, we define the *right-sequence* $r(P) = r_1, r_2, \ldots, r_n$

[1] The terminology left- and right- is due to the first k positions and last k positions, respectively, in the sequence σ.

of P, where r_k is the smallest integer in the set $\{i_k, i_{k+1}, \dots, i_n\}$. Then $r_k \leqslant k \leqslant l_k$ and $r_k \leqslant i_k \leqslant l_k$ for $k = 1, 2, \dots, n$.

Theorem 1.2 ([4]). *Let P be a permutation matrix of order n. Then its Bruhat shadow $\mathcal{S}(P) = [s_{kj}]$ is given by*

$$s_{kj} = \begin{cases} 1 & \textit{if } r_k \leqslant j \leqslant l_k, \\ 0 & \textit{otherwise}, \end{cases} \qquad 1 \leqslant k \leqslant n, \text{ and } 1 \leqslant j \leqslant n.$$

The matrix $\mathcal{S}(P)$ has total support.

The definition of the left- and right-sequences implies that the matrix $\mathcal{S}(P)$ has a staircase pattern with $I_n \leqslant \mathcal{S}(P)$ and $P \leqslant \mathcal{S}(P)$. Here by a *staircase pattern* we mean that the 1's in each row and column are consecutive where the first (last) 1 in a row is in the same or earlier (later) column than the first (last) 1 in the following row. For example, if $\sigma = (5, 7, 1, 3, 2, 6, 4)$, we have $l(P) = 5, 7, 7, 7, 7, 7, 7$ and $r(P) = 1, 1, 1, 2, 2, 4, 4$, so

$$\mathcal{S}(P) = \begin{bmatrix} 1 & 1 & 1 & 1 & \mathbf{1} & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & \mathbf{1} \\ \mathbf{1} & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & \mathbf{1} & 1 & 1 & 1 & 1 \\ 0 & \mathbf{1} & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & \mathbf{1} & 1 \\ 0 & 0 & 0 & \mathbf{1} & 1 & 1 & 1 \end{bmatrix},$$

where the 1's of the permutation matrix corresponding to σ are in boldface.

2. Doubly stochastic matrices

Given a permutation matrix $Q \in \mathcal{P}_n$, let

$$(\preceq_B Q) = \{P \in \mathcal{P}_n \colon\ P \preceq_B Q\}.$$

Then $(\preceq_B Q)$ is a principal ideal of the Bruhat order on $\mathcal{P}_n$. Let

$$\Omega_n(\preceq_B Q) = \operatorname{conv}(\preceq_B Q)$$

be the convex hull of $(\preceq_B Q)$, which is a subpolytope of Ω_n. Moreover, we define

$$\Omega_n\ (\geqslant \Sigma(Q)) = \{A \in \Omega_n \colon\ \Sigma(A) \geqslant \Sigma(Q)\}$$

and this set coincides with $\{A \in \Omega_n \colon\ A \preceq_B Q\}$.

Any $(0,1)$-matrix C of order n having total support induces a face of the Birkhoff polytope Ω_n as

$$\Omega_n^C := \{A \in \Omega_n \colon\ A \leqslant C\}.$$

In addition, any face of Ω_n arises from such a unique C in this way (see [3]). In particular, when Q is a permutation matrix, $\mathcal{S}(Q)$ has total support, so $\Omega_n^{\mathcal{S}(Q)}$ is a face of Ω_n.

Proposition 2.1. *Let $Q = [q_{ij}]$ be a permutation matrix of order n corresponding to the permutation $(i_1, i_2, \ldots, i_n)$. Then*

$$\Omega_n(\preceq_B Q) \subseteq \Omega_n\ (\geqslant \Sigma(Q)) \subseteq \Omega_n^{\mathcal{S}(Q)} \tag{1}$$

and all these sets are polytopes.

Proof. We have that $\Omega_n\ (\geqslant \Sigma(Q))$ is a polytope, as it is a bounded polyhedron defined by the n^2 linear inequalities from $\Sigma(A) \geqslant \Sigma(Q)$ and the linear equations/inequalities defining the Birkhoff polytope. Since $\Omega_n\ (\geqslant \Sigma(Q))$ contains each $P \in \mathcal{P}_n$ satisfying $P \preceq_B Q$, the first inclusion in (1) follows from convexity.

Next, we show that $\Omega_n\ (\geqslant \Sigma(Q)) \subseteq \Omega_n^{\mathcal{S}(Q)}$. Let $A = [a_{ij}] \in \Omega_n\ (\geqslant \Sigma(Q))$ and $1 \leqslant k \leqslant n$. Since the ones in rows $1, 2, \ldots, k$ of Q are in columns $i_1, i_2, \ldots, i_k$, $\sigma_{kl_k}(A) \geqslant k$, where $l_k = \max\{i_1, i_2, \ldots, i_k\}$. But $\sigma_{kn}(A) = k$, so we conclude that $a_{kj} = 0$ for $j > l_k$. Similarly, consider column k of A and let l be the largest index of the row that contains a one within columns $1, 2, \ldots, k$. The staircase pattern of $\mathcal{S}(Q)$ now implies that all the ones in columns $1, 2, \ldots, k$ of Q are in rows $1, 2, \ldots, l$, so $\sigma_{lk}(A) \geqslant \sigma_{lk}(Q) = k$. Therefore $a_{ik} = 0$ for $i > l$. This shows that $A \leqslant \mathcal{S}(Q)$, so $A \in \Omega_n^{\mathcal{S}(Q)}$. $\square$

Note that if $A \in \Omega_n$, then the entries in the last row and the last column of $\Sigma(A)$ are $1, 2, \ldots, n$.

Example 1. In this example we show that the first containment in Proposition 2.1 may be strict. Let

$$Q = \left[\begin{array}{c|c|c|c|c|c} & & & 1 & & \\ \hline & 1 & & & & \\ \hline & & & & 1 & \\ \hline 1 & & & & & \\ \hline & & & & & 1 \\ \hline & & 1 & & & \end{array}\right].$$

Then

$$\Sigma(Q) = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 & 2 & 2 \\ 0 & 1 & 1 & 2 & 3 & 3 \\ 1 & 2 & 2 & 3 & 4 & 4 \\ 1 & 2 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{bmatrix}.$$

Let $A = [a_{ij}] \in \Omega_6$. If $\Sigma(A) \geqslant \Sigma(Q)$, then by Proposition 2.1 A has zeros as shown below:

$$A = \left[\begin{array}{c|c|c|c|c|c} & & & & 0 & 0 \\ \hline & & & & 0 & 0 \\ \hline & & & & & 0 \\ \hline & & & & & 0 \\ \hline 0 & 0 & & & & \\ \hline 0 & 0 & & & & \end{array}\right].$$

Since A is assumed to be doubly stochastic, the only inequality in $\Sigma(A) \geqslant \Sigma(Q)$ that does not follow from the form of A is

$$\sigma_{22}(A) \geqslant \sigma_{22}(Q), \quad \text{that is } a_{11} + a_{12} + a_{21} + a_{22} \geqslant 1.$$

Let

$$A = \frac{1}{4}\left[\begin{array}{c|c|c|c|c|c} 3 & & 1 & & & \\ \hline & 3 & & 1 & & \\ \hline & 1 & 3 & & & \\ \hline 1 & & & 3 & & \\ \hline & & & & 3 & 1 \\ \hline & & & & 1 & 3 \end{array}\right].$$

Since $\sigma_{22}(A) = 3/2 \geqslant 1$, it follows that A satisfes $\Sigma(A) \geqslant \Sigma(Q)$ and hence that $A \preceq_B Q$. However, A is not in the convex hull of $(\preceq_B Q)$ because any permutation matrix with a one in position $(1,3)$ whose support is a subset of the support of A is of the form

$$P = \left[\begin{array}{c|c|c|c|c|c} & & 1 & & & \\ \hline & & & 1 & & \\ \hline & 1 & & & & \\ \hline 1 & & & & & \\ \hline & & & & * & * \\ \hline & & & & * & * \end{array}\right]$$

and is not in $(\preceq_B Q)$.

The previous example leads to the following question concerning a weaker property. Since $\Omega_n(\preceq_B)$ may not equal $\Omega_n\ (\geqslant \Sigma(Q))$, a weaker property is that $A \in \Omega_n(\geqslant \Sigma(Q))$ implies that there exists a permutation matrix P with $P \preceq_B Q$ and $\operatorname{supp} P \subseteq \operatorname{supp} A$. But even this may not be true as the following example shows.

Example 2. Consider the following permutation matrix Q where the zeros shown are those of the Bruhat shadow:

$$Q = \left[\begin{array}{c|c|c|c|c|c|c|c} & & & & & & 1 & 0 \\ \hline & & & & 1 & & & 0 \\ \hline & & & & & 1 & & 0 \\ \hline & & & & & & & 1 \\ \hline & & & 1 & & & & \\ \hline & 1 & & & & & & \\ \hline & & 1 & & & & & \\ \hline 1 & & & & & & & \end{array}\right].$$

Let

$$A = \left[\begin{array}{c|c|c|c|c|c|c|c} & & & 1/2 & & 1/2 & & \\ \hline & & & & 1/2 & & 1/2 & \\ \hline & & & & 1/2 & 1/2 & & \\ \hline & & & & & & 1/2 & 1/2 \\ \hline 1/2 & & & & & & & 1/2 \\ \hline & 1/2 & 1/2 & & & & & \\ \hline 1/2 & & 1/2 & & & & & \\ \hline & 1/2 & & 1/2 & & & & \end{array}\right].$$

Clearly $A \in \Omega_8^{\mathcal{S}(Q)}$, and one can verify that $A \in \Omega_8\ (\geqslant \Sigma(Q))$. Consider the permutation matrices

$$P_1 = \left[\begin{array}{c|c|c|c|c|c|c|c} & & & 0 & & 1 & & \\ \hline & & & & 0 & & 1 & \\ \hline & & & & 1 & 0 & & \\ \hline & & & & & & 0 & 1 \\ \hline 1 & & & & & & & 0 \\ \hline & 1 & 0 & & & & & \\ \hline 0 & & 1 & & & & & \\ \hline & 0 & & 1 & & & & \end{array}\right], \quad P_2 = \left[\begin{array}{c|c|c|c|c|c|c|c} & & & 1 & & 0 & & \\ \hline & & & & 1 & & 0 & \\ \hline & & & & 0 & 1 & & \\ \hline & & & & & & 1 & 0 \\ \hline 0 & & & & & & & 1 \\ \hline & 0 & 1 & & & & & \\ \hline 1 & & 0 & & & & & \\ \hline & 1 & & 0 & & & & \end{array}\right].$$

Here $P_1, P_2 \in \Omega_8^{\mathcal{S}(Q)}$ but $P_1 \npreceq_B Q$ and $P_2 \npreceq_B Q$ as

$$0 = \sigma_{25}(P_1) < \sigma_{25}(Q) = 1 \quad \text{and} \quad 0 = \sigma_{62}(P_2) < \sigma_{62}(Q) = 1.$$

Now, $A = \frac{1}{2}P_1 + \frac{1}{2}P_2$, and the only permutation matrices P satisfying $\operatorname{supp} P \subseteq \operatorname{supp} A$ are P_1 and P_2. The last fact is easy to check directly. In fact, P_1 and P_2

have disjoint support and their union corresponds to a single cycle in the bipartite graph representation of the permutation matrices.

In the previous example, $P_1, P_2 \in \Omega_8^{S(Q)}$ but $P_1, P_2 \not\preceq_B Q$, and hence by Theorem 1.2, P_1, P_2 are not in Ω_n $(\geqslant \Sigma(Q))$, which shows that the second containment in Proposition 2.1 can be proper, even with respect to permutation matrices.

Example 3. Let

$$Q = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

There are four permutation matrices in $(\preceq_B Q)$, namely the 3×3 permutation matrices with a zero in position $(3,1)$. Hence it follows that

$$\Omega_3(\preceq_B Q) = \Omega_3^{S(Q)} = \left\{ \begin{bmatrix} b+d & c & a \\ a+c & d & b \\ 0 & a+b & c+d \end{bmatrix} : a,b,c,d \geqslant 0,\ a+b+c+d = 1 \right\}.$$

Let $A = [a_{ij}] \in \Omega_3$ satisfy

$$\Sigma(A) \geqslant \Sigma(Q) = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 3 \end{bmatrix}.$$

Then $a_{11} + a_{21} \geqslant 1$, and hence $a_{11} + a_{21} = 1$ and $a_{31} = 0$. Thus, A is a convex combination of 3×3 permutation matrices with entry $(3,1)$ equal to 0, that is, A is in $\Omega_3(\preceq_B Q)$. Thus, in this case

$$\Omega_3(\preceq_B Q) = \Omega_3\ (\geqslant \Sigma(Q)) = \Omega^{S(Q)}.$$

$\square$

Let $A_1, A_2 \in \Omega_n$. Our goal is to obtain a better understanding of the *stochastic Bruhat order* (recall $A_1 \preceq_B A_2$ provided that $\Sigma(A_1) \geqslant \Sigma(A_2)$).

Let $A = [a_{ij}] \in \Omega_n$. A *backward ε-interchange* of A is a replacement of a 2×2 submatrix of A with another 2×2 matrix as indicated below:

$$\begin{bmatrix} a_{ij} & a_{il} \\ a_{kj} & a_{kl} \end{bmatrix} \longrightarrow \begin{bmatrix} a_{ij}+\varepsilon & a_{il}-\varepsilon \\ a_{kj}-\varepsilon & a_{kl}+\varepsilon \end{bmatrix}.$$

A *forward ε-interchange* is defined by

$$\begin{bmatrix} a_{ij} & a_{il} \\ a_{kj} & a_{kl} \end{bmatrix} \longrightarrow \begin{bmatrix} a_{ij}-\varepsilon & a_{il}+\varepsilon \\ a_{kj}+\varepsilon & a_{kl}-\varepsilon \end{bmatrix}.$$

Here ε is assumed to satisfy $0 < \varepsilon \leqslant a_{il}, a_{kj}$ in the backward case, and $0 < \varepsilon \leqslant a_{ij}, a_{kl}$ in the forward case. If A' results from a doubly stochastic matrix by a backward ε-interchange in rows i_0, i_1 and columns j_0, j_1, then $\Sigma(A')$ is given by

$$\sigma_{ij}(A') = \begin{cases} \sigma_{ij}(A) + \varepsilon & \text{if } i_0 \leqslant i < i_1 \text{ and } j_0 \leqslant j < j_1, \\ \sigma_{ij}(A) & \text{otherwise.} \end{cases}$$

Thus, if A' results from $A \in \Omega_n$ ($\geqslant \Sigma(Q)$) by a sequence of backward ε-interchanges, then also $A' \in \Omega_n$ ($\geqslant \Sigma(Q)$). Applying a forward ε-interchange, $\sigma_{ij}(A)+\varepsilon$ is replaced by $\sigma_{ij}(A)-\varepsilon$ in the expression above. Note that forward and backward ε-interchanges are inverse operations of each other.

Theorem 2.2. *Let $A_1, A_2 \in \Omega_n$. Then the following statements are equivalent:*

(i) $A_1 \preceq_B A_2$,

(ii) *A_1 can be obtained from A_2 by a finite sequence of backward ε-interchanges; equivalently, A_2 can be obtained from A_1 by a finite sequence of forward ε-interchanges.*

Proof. As shown above, (ii) implies (i), so we only need to prove that (i) implies (ii). Assume $A_1 \preceq_B A_2$, where $A_1 = [a_{ij}]$, $A_2 = [a'_{ij}]$. If $A_1 = A_2$, then there is nothing to be proved.

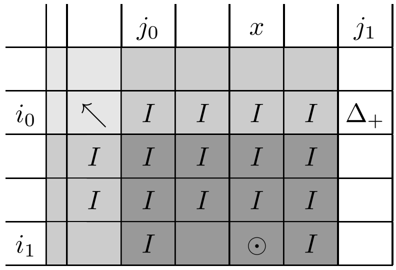
If $A_1 \neq A_2$, then there is at least one entry (i,j) such that $a_{ij} \neq a'_{ij}$. We define the sets of positions

$$\Delta_+ = \{(i,j)\colon\ a_{ij} < a'_{ij}\},$$
$$I = \{(i,j)\colon\ \sigma_{ij}(A_1) > \sigma_{ij}(A_2)\}.$$

Let i_0 be the first row in which A_1 and A_2 differ, and let j_1 be the largest j with $a_{i_0 j} \neq a'_{i_0 j}$. Clearly, $j_1 > 1$ because otherwise row i_0 of A_1 or A_2 would not have sum one. In the arguments that follow we use that $\sigma_{i_0 n}(A_1) = \sigma_{i_0 n}(A_2) = i_0$. We have $(i_0, j_1) \notin I$ because $a_{ij} = a'_{ij}$ for $i \leqslant i_0$ and $j > j_1$, so $\sigma_{i_0 j_1}(A_1) = \sigma_{i_0 j_1}(A_2)$. Since $\sigma_{i_0 j_1 - 1}(A_1) \geqslant \sigma_{i_0 j_1 - 1}(A_2)$, we conclude that $a_{i_0 j_1} < a'_{i_0 j_1}$ and hence $(i_0, j_1) \in \Delta_+$. Note that $(i_0, j_1 - 1) \in I$ because $(i_0, j_1) \in \Delta_+$.

Let $j_0 < j_1$ be the smallest index such that $(i_0, j) \in I$ for all $j_0 \leqslant j < j_1$ (j_0 exists, $j_1 - 1$ is one candidate). Now let $i_1 > i_0$ be the largest index such that $(i,j) \in I$ for all $j_0 \leqslant j < j_1$ and $i_0 \leqslant i < i_1$ (note that i_1 exists, $i_0 + 1$ is a candidate and there is no element in row n belonging to I). In row i_1, there is a column $j_0 \leqslant x < j_1$ such that $(i_1, x) \notin I$, otherwise i_1 would be bigger.

For the contradiction, let us suppose that $a_{ij} \geqslant a'_{ij}$ is in the rectangle given by $i_0 < i \leqslant i_1$ and $j_0 \leqslant j < j_1$.



Using the minimality of j_0 we have $\sigma_{i_0 j_0-1}(A_1) = \sigma_{i_0 j_0-1}(A_2)$, of course if $j_0 \geqslant 2$ (if $j_0 = 1$, then we disregard rectangles with column $j_0 - 1$). Also recall that $\sigma_{i_1 j_0-1}(A_1) \geqslant \sigma_{i_1 j_0-1}(A_2)$ in general, and $\sigma_{i_0 x}(A_1) > \sigma_{i_0 x}(A_2)$ since $(i_0, x) \in I$. Then, by the above assumption, we have

$$\begin{aligned}
\sigma_{i_1 x}(A_1) &= \sigma_{i_1 j_0-1}(A_1) + \sigma_{i_0 x}(A_1) - \sigma_{i_0 j_0-1}(A_1) + \sum_{\substack{i_0 < i \leqslant i_1 \\ j_0 \leqslant j \leqslant x}} a_{ij} \\
&> \sigma_{i_1 j_0-1}(A_2) + \sigma_{i_0 x}(A_2) - \sigma_{i_0 j_0-1}(A_2) + \sum_{\substack{i_0 < i \leqslant i_1 \\ j_0 \leqslant j \leqslant x}} a'_{ij} \\
&= \sigma_{i_1 x}(A_2),
\end{aligned}$$

a contradiction since $(i_1, x) \notin I$. Thus, there is a $(i_*, j_*) \in \Delta_+$ contained in the rectangle defined by $i_0 < i_* \leqslant i_1$ and $j_0 \leqslant j_* < j_1$.

Now we can apply a backward ε-interchange to A_2 by adding the matrix

$$\begin{bmatrix} \varepsilon & -\varepsilon \\ -\varepsilon & \varepsilon \end{bmatrix}$$

to the submatrix $A_2[i_0, i_* | j_*, j_1]$ determined by rows i_0 and i_*, and columns j_* and j_1 with

$$\varepsilon = \min\{a'_{i_0 j_1} - a_{i_0 j_1}; a'_{i_* j_*} - a_{i_* j_*}; \sigma(A_1)_{ij} - \sigma(A_2)_{ij} \text{ for } i_0 \leqslant i < i_* \text{ and } j_* \leqslant j < j_1\}.$$

This operation creates a matrix A^* such that $\Sigma(A_1) \geqslant \Sigma(A^*) \geqslant \Sigma(A_2)$ with at least one entry of $\Sigma(A^*)$ strictly bigger than the corresponding entry of $\Sigma(A_2)$. Therefore $A_1 \preceq_B A^* \preceq_B A_2$.

Case 1: If $\varepsilon = \min\{\sigma(A_1)_{ij} - \sigma(A_2)_{ij}\}$, then we have strictly increased the number of entries where $\Sigma(A_1)$ and $\Sigma(A_2)$ agree, that is, some entries of I are removed.

Case 2: If $\varepsilon = a'_{i_0 j_1} - a_{i_0 j_1}$ (upper right corner), then (i_0, j_1) is no longer in Δ_+. In the next step, we will take a new element of Δ_+ in column $j_2 < j_1$, or there will be no more elements in row i_0 in Δ_+. In any case, the position $(i_0, j_1 - 1)$ will no longer belong to the set I, and again we strictly increased the number of entries where $\Sigma(A_1)$ and $\Sigma(A_2)$ agree.

Case 3: If $\varepsilon = a'_{i_*j_*} - a_{i_*j_*}$, then (i_*, j_*) is no longer in Δ_+. But there could be another position $(i_{**}, j_{**}) \in \Delta_+$ in the rectangle $i_0 < i \leqslant i_1$ and $j_0 \leqslant j < j_1$. We repeat applying backwards ε-interchanges until (i_0, j_1) is eliminated from Δ_+, which eliminates $(i_0, j_1 - 1)$ from I.

Since each backward ε-interchange brings A_2 closer to A_1 by decreasing $|I|$, eventually we will have $I = \emptyset$ and then A_1 is reached, as desired.

A_2 can be obtained from A_1 by forward ε-interchanges in the reverse order. $\square$

The previous proof gives an algorithm for bringing A_2 to A_1 when $A_1 \preceq_B A_2$ holds. We illustrate this algorithm by an example.

Example 4. Consider the matrices A_1 and A_2 below such as $A_1 \preceq_B A_2$ and I and Δ_+ are shown schematically:

$$A_1 = \frac{1}{10}\begin{bmatrix} 3 & 1 & 2 & 3 & 1 \\ 3 & 2 & 4 & 0 & 1 \\ 2 & 3 & 4 & 1 & 0 \\ 2 & 1 & 0 & 3 & 4 \\ 0 & 3 & 0 & 3 & 4 \end{bmatrix}, \qquad A_2 = \frac{1}{10}\begin{bmatrix} 2 & 1 & 3 & 0 & 4 \\ 4 & 2 & 2 & 2 & 0 \\ 2 & 2 & 5 & 1 & 0 \\ 1 & 1 & 0 & 4 & 4 \\ 1 & 4 & 0 & 3 & 2 \end{bmatrix};$$

$$I\Delta_+ = \left[\begin{array}{c|c|c|c|c} I & I & \Delta & I & \Delta \\ \hline \Delta & & I & I^{\Delta} & \\ \hline & I & I^{\Delta} & I & \\ \hline I & I & I & I^{\Delta} & \\ \hline \Delta & \Delta & & & \end{array}\right].$$

The first modification to bring A_2 to A_1 consists of a backward ε-interchange using position $(i_0, j_1) = (1, 5) \in \Delta_+$. There are two positions in the rectangle $1 < i \leqslant 5$ and $4 \leqslant j < 5$ belonging to Δ_+. We choose $(2, 4) \in \Delta_+$ and apply $\varepsilon(1, 2 : 4, 5 : 2/10)$, the backward ε-interchange in rows 1 and 2, and columns 4 and 5, for $\varepsilon = \frac{2}{10} = \frac{1}{10}\min\{3, 2, 3\}$. This leads us to the matrix

$$B_1 = \frac{1}{10}\begin{bmatrix} 2 & 1 & 3 & 2 & 2 \\ 4 & 2 & 2 & 0 & 2 \\ 2 & 2 & 5 & 1 & 0 \\ 1 & 1 & 0 & 4 & 4 \\ 1 & 4 & 0 & 3 & 2 \end{bmatrix}; \qquad \left[\begin{array}{c|c|c|c|c} I & I & \Delta & I & \Delta \\ \hline \Delta & & I & I & \Delta \\ \hline & I & I^{\Delta} & I & \\ \hline I & I & I & I^{\Delta} & \\ \hline \Delta & \Delta & & & \end{array}\right]$$

with the property $A_1 \preceq_B B_1 \preceq_B A_2$. This operation does not get us closer to A_1 (Case 3 of the proof) in the sense that I remains the same. So we choose the next position in the rectangle that belongs to Δ_+, so $(i_0, j_1) = (4, 4)$. We apply $\varepsilon(1, 4|4, 5|1/10)$ and obtain

$$B_2 = \frac{1}{10}\begin{bmatrix} 2 & 1 & 3 & 3 & 1 \\ 4 & 2 & 2 & 0 & 2 \\ 2 & 2 & 5 & 1 & 0 \\ 1 & 1 & 0 & 3 & 5 \\ 1 & 4 & 0 & 3 & 2 \end{bmatrix}; \quad \left[\begin{array}{c|c|c|c|c} I & I & \Delta & & \\ \hline \Delta & & I & I & \Delta \\ \hline & I & I^{\Delta} & I & \\ \hline I & I & I & I & \Delta \\ \hline \Delta & \Delta & & & \end{array}\right].$$

This operation (Cases 1 and 2 in the proof) decreases $|I|$ by one. Next we have $(i_0, j_1) = (1, 3) \in \Delta_+$, and $j_0 = 1$ and $i_1 = 2$. We apply $\varepsilon(1, 2|1, 3|1/10)$ and obtain

$$B_3 = \frac{1}{10}\begin{bmatrix} 3 & 1 & 2 & 3 & 1 \\ 3 & 2 & 3 & 0 & 2 \\ 2 & 2 & 5 & 1 & 0 \\ 1 & 1 & 0 & 3 & 5 \\ 1 & 4 & 0 & 3 & 2 \end{bmatrix}; \quad \left[\begin{array}{c|c|c|c|c} & & & & \\ \hline & & I & I & \Delta \\ \hline & I & I^{\Delta} & I & \\ \hline I & I & I & I & \Delta \\ \hline \Delta & \Delta & & & \end{array}\right].$$

Next we have $(i_0, j_1) = (2, 5) \in \Delta_+$ and choose the unique position of Δ_+ in the rectangle $2 < i \leqslant 5$ and $3 \leqslant j < 5$ and obtain

$$B_3 = \frac{1}{10}\begin{bmatrix} 3 & 1 & 2 & 3 & 1 \\ 3 & 2 & 3 & 0 & 2 \\ 2 & 2 & 5 & 1 & 0 \\ 1 & 1 & 0 & 3 & 5 \\ 1 & 4 & 0 & 3 & 2 \end{bmatrix}; \quad \left[\begin{array}{c|c|c|c|c} & & & & \\ \hline & & I & I & \Delta \\ \hline & I & I^{\Delta} & I & \\ \hline I & I & I & I & \Delta \\ \hline \Delta & \Delta & & & \end{array}\right].$$

We apply $\varepsilon(2, 3|3, 5|1/10)$ and obtain

$$B_4 = \frac{1}{10}\begin{bmatrix} 3 & 1 & 2 & 3 & 1 \\ 3 & 2 & 4 & 0 & 1 \\ 2 & 2 & 4 & 1 & 1 \\ 1 & 1 & 0 & 3 & 5 \\ 1 & 4 & 0 & 3 & 2 \end{bmatrix}; \quad \left[\begin{array}{c|c|c|c|c} & & & & \\ \hline & & & & \\ \hline & I & I & I & \Delta \\ \hline I & I & I & I & \Delta \\ \hline \Delta & \Delta & & & \end{array}\right].$$

Next we have $(i_0, j_1) = (3,5) \in \Delta_+$ and apply $\varepsilon(3,5|2,5|1/10)$ obtaining

$$B_5 = \frac{1}{10}\begin{bmatrix} 3 & 1 & 2 & 3 & 1 \\ 3 & 2 & 4 & 0 & 1 \\ 2 & 3 & 4 & 1 & 0 \\ 1 & 1 & 0 & 3 & 5 \\ 1 & 3 & 0 & 3 & 3 \end{bmatrix}; \quad \left[\begin{array}{c|c|c|c|c} & & & & \\ \hline & & & & \\ \hline & & & & \\ \hline I & I & I & I & \Delta \\ \hline \Delta & & & & \end{array}\right].$$

After that we have $(i_0, j_1) = (4,5) \in \Delta_+$ and apply $\varepsilon(4,5|1,5|1/10)$ to finally reach

$$A_1 = \frac{1}{10}\begin{bmatrix} 3 & 1 & 2 & 3 & 1 \\ 3 & 2 & 4 & 0 & 1 \\ 2 & 3 & 4 & 1 & 0 \\ 2 & 1 & 0 & 3 & 4 \\ 0 & 3 & 0 & 3 & 4 \end{bmatrix}.$$

We remark that one can see from Example 4 that Case 3 in the proof of Theorem 2.2 is in fact needed. The first step allows us to choose (i_*, j_*) as $(2,4)$ or $(4,4)$, but neither of these single choices will decrease set I.

3. Bruhat faces

A *Bruhat face* of Ω_n is a face $\mathcal{F}$ of Ω_n for which there exists a permutation matrix Q such that the set of vertices of $\mathcal{F}$ is $(\preceq_B Q)$; equivalently,

$$\{P \in \mathcal{P}_n \colon\ P \leqslant \mathcal{S}(Q)\} = (\preceq_B Q).$$

We then write $\mathcal{F} = \mathcal{F}(Q)$ and say that Q *induces or generates the Bruhat face* $\mathcal{F}(Q)$. If $\mathcal{F}(Q)$ is a Bruhat face, then the $(0,1)$-matrix determining that face is the shadow $\mathcal{S}(Q)$ of Q. Thus, for a Bruhat face $\mathcal{F}$ the three sets in Proposition 2.1 coincide.

Following [4] we define the *Bruhat convex hull* of a $(0,1)$-matrix $A = [a_{ij}]$ as the $(0,1)$-matrix whose support is the union of all sets $\{(r,s)\colon\ i' \leqslant r \leqslant i \ \text{ and } j \leqslant s \leqslant j'\}$ such that $a_{ij} = a_{i'j'} = 1$ with $i' < i$ and $j < j'$. Let B be a matrix with staircase pattern and let S be its support. Let $(i,j) \in S$ and let B' be the Bruhat convex hull of the matrix with support $S \setminus \{(i,j)\}$. Then (i,j) is in an *extreme position* in B if $B \neq B'$. One might think that if each 1 in Q is in an extreme position, then Q induces a Bruhat face. However, this is not the case as the following example shows.

Example 5. Consider the permutation matrix

$$Q = \begin{bmatrix} & & 1 & & & \\ & & & & 1 & \\ 1 & & & & & \\ & & & & & 1 \\ & 1 & & & & \\ & & & 1 & & \end{bmatrix} \quad \text{with } \mathcal{S}(Q) = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}.$$

Q does not induce a Bruhat face. To see this, consider the following permutation matrix P which lies in $\Omega_n^{\mathcal{S}(Q)}$:

$$P = \begin{bmatrix} 1 & & & & & \\ & & & & 1 & \\ & & & 1 & & \\ & & 1 & & & \\ & 1 & & & & \\ & & & & & 1 \end{bmatrix}.$$

Here $P \not\preceq_B Q$ as $\sigma_{33}(P) = 1 < 2 = \sigma_{33}(Q)$. Moreover, $Q \not\preceq_B P$. Actually, both P and Q are maximal elements in the Bruhat order among permutation matrices in the face $\Omega_n^{\mathcal{S}(Q)}$.

We now consider which permutation matrices Q generate Bruhat faces. If $Q \in \mathcal{P}_n$ induces a Bruhat face, then no other permutation matrix induces the same Bruhat face. This is because if $Q' \in \mathcal{P}_n$ induces the same Bruhat face, then $\mathcal{S}(Q) = \mathcal{S}(Q')$, $Q' \preceq_B Q$ and $Q \preceq_B Q'$, so $Q = Q'$.

Clearly, if $Q \in \mathcal{P}_n$ and $Q' \in \mathcal{P}_m$ each induces a Bruhat face, then the direct sum $Q \oplus Q'$ induces a Bruhat face.

For a nonnegative $n \times n$ matrix $A = [a_{ij}]$ with non-decreasing rows and columns let

$$\Delta(A) = \{(i,j)\colon\ a_{ij} > \max\{a_{i-1,j}, a_{i,j-1}\},$$

where we let $a_{0i} = a_{i0} = 0$, $1 \leqslant i \leqslant n$. Define for $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant n$

$$\gamma_{ij}(Q) = \min\{\sigma_{ij}(P)\colon\ P \leqslant \mathcal{S}(Q),\ P \text{ a permutation matrix}\},$$

so $\gamma_{ij}(Q) \leqslant \sigma_{ij}(Q)$. Let $\Gamma(Q) = [\gamma_{ij}(Q)]$ be the corresponding $n \times n$ matrix with these numbers as its entries. Below we give a simple and efficient method for computing these numbers. $\Gamma(Q)$ is nonnegative and has non-decreasing rows and columns. This is also the case for matrix $\Sigma(Q)$. The *term rank* of a $(0,1)$-matrix A is the maximum cardinality of a set of ones in A such that no two are in the same row or column.

Theorem 3.1. *Let Q be a permutation matrix of order n. Then the following statements are equivalent:*

(i) *Q induces a Bruhat face.*

(ii) $\Gamma(Q) = \Sigma(Q)$.

(iii) $\Delta(\Gamma(Q)) = \Delta(\Sigma(Q))$.

(iv) *For each $i, j \leqslant n$ the term rank of the matrix obtained from $\mathcal{S}(Q)$ by replacing its leading $i \times j$ submatrix with a zero matrix is $n - \sigma_{ij}(Q)$.*

Proof. If Q induces a Bruhat face, then $\Gamma(Q) = \Sigma(Q)$ (for if $\gamma_{ij}(Q) < \sigma_{ij}(Q)$ for some i, j, then there would exist a $P \leqslant \mathcal{S}(Q)$ with $P \not\leqslant_B Q$). Conversely, if $\Gamma(Q) = \Sigma(Q)$, then every permutation matrix P with $P \leqslant \mathcal{S}(Q)$ also satisfies $\Sigma(P) \geqslant \Sigma(Q)$. This shows the equivalence of (i) and (ii).

Clearly, (ii) implies (iii). Next, assume (iii) holds. In each of matrices $\Gamma(Q)$ and $\Sigma(Q)$ the first row consists of a sequence of zeros followed by a sequence of ones. Moreover, the transition from 0 to 1 occurs in the same column j; this follows from the assumption $\Delta(\Gamma(Q)) = \Delta(\Sigma(Q))$ because this set contains a unique element $(1, j)$ for some $j \leqslant n$. The second row of $\Gamma(Q)$ and $\Sigma(Q)$ consists of a sequence of 0's followed by a sequence of 1's and finally a sequence of 2's. Using again assumption (iii) and the fact that the first row of $\Gamma(Q)$ and $\Sigma(Q)$ coincide, we conclude that the second row of these two matrices coincide. We may proceed by induction and conclude that $\Gamma(Q) = \Sigma(Q)$ holds.

Finally, (ii) and (iv) are equivalent as (iv) means that the minimum number of ones in the leading $i \times j$ submatrix of a permutation matrix $P \leqslant \mathcal{S}(Q)$ is $\sigma_{ij}(Q)$. $\square$

Recall that a backward interchange in a permutation matrix P is replacing a 2×2 submatrix which is equal to L_2 by I_2. Note that the resulting matrix P' satisfies $\mathcal{S}(P') \leqslant \mathcal{S}(P)$. Let $1 \leqslant i, j \leqslant n$, and let k be such that $\max\{i + j - n, 0\} \leqslant k \leqslant \min\{i, j\}$. Define the $n \times n$ permutation matrix

$$P^{(i,j,k,n)} = \begin{bmatrix} I_k & O & O & O \\ O & O_{i-k,j-k} & I_{i-k} & O \\ O & I_{j-k} & O & O \\ O & O & O & I_{n-i-j+k} \end{bmatrix}. \tag{2}$$

Theorem 3.2. *Let Q be a permutation matrix and $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant n$. Let $k = \gamma_{ij}(Q)$. Then $P^{(i,j,k,n)} \leqslant \mathcal{S}(Q)$ and*

$$\sigma_{ij}(P^{(i,j,k,n)}) = \gamma_{ij}(Q),$$

and thus $P^{(i,j,k,n)}$ minimizes $\sigma_{ij}(P)$ among all permutation matrices P satisfying $P \leqslant \mathcal{S}(Q)$.

P r o o f. Let $P=[p_{rs}]\in\mathcal{P}_n$ be such that $P\leqslant\mathcal{S}(Q)$ and $\sigma_{ij}(P)=\gamma_{ij}(Q)=k$. Assume that $k\geqslant 1$. If $p_{l1}=1$ with $l>i$, choose (r,s) with $p_{rs}=1$ and $r\leqslant i$, $s\leqslant j$. Then make a backward interchange for rows r, l and columns 1, s. The new matrix, still called P for simplicity, also has k ones in the leading $i\times j$ submatrix. If $r=1$, we now have $p_{11}=1$. Otherwise, when $p_{r1}=1$ for some $r>1$, make a backward interchange involving positions $(r,1)$ and the position of the unique 1 in row 1. After this, the new updated matrix P satisfies $p_{11}=1$. We may now delete the first row and column, and repeat this procedure for the remaining $k-1$ ones in the leading $i\times j$ submatrix. After this we have

$$p_{11}=p_{22}=\ldots=p_{kk}=1.$$

So, even if $k=0$, the leading $i\times j$ submatrix of P now coincides with that of $P^{(i,j,k,n)}$, and P has the following structure

$$P=\begin{bmatrix} I_k & O & O \\ O & O_{i-k,j-k} & A_{23} \\ O & A_{32} & A_{33}\end{bmatrix}.$$

Each column of A_{32} contains a 1 and with backward interchanges we may assure that each 1 in this submatrix is to the right of each 1 in its previous rows. This is possible due to the staircase structure and does not affect the number of ones in the leading $i\times j$ submatrix of P. Moreover, for each row in A_{32} which is zero, there must be a 1 in the same row in A_{33}. This fact makes it possible to perform backward interchanges until the leading $(j-k)\times(j-k)$ submatrix of A_{32} equals I_{j-k}. After this we have

$$P=\begin{bmatrix} I_k & O & O \\ O & O_{i-k,j-k} & A_{24} \\ O & I_{j-k} & O \\ O & O & A_{44}\end{bmatrix}.$$

Now, each row of A_{24} contains a 1 and with backward interchanges involving A_{24} and A_{44} we may assure that each 1 in this submatrix A_{24} is to the right of each 1 in its previous rows. Moreover, for each column in A_{24} which is zero, there must be a 1 in the same column in A_{44}. We may then use backward interchanges, so that the leading $(i-k)\times(i-k)$ submatrix of A_{24} equals I_{i-k}. Now backward interchanges on the lower right submatrix get us to $P=P^{(i,j,k,n)}$ as desired. $\square$

Corollary 3.3. *For every $Q\in\mathcal{P}_n$ and $1\leqslant i\leqslant n$, $1\leqslant j\leqslant n$*

$$\gamma_{ij}(Q)=\min\{k\colon\ P^{(i,j,k,n)}\leqslant\mathcal{S}(Q)\}.$$

This corollary leads to a simple and efficient algorithm for computing $\gamma_{ij}(Q)$ for given i,j and $Q \in \mathcal{P}_n$: start with $k = \max\{i+j-n, 0\}$ and increase k by 1 until $P^{(i,j,k,n)} \leqslant \mathcal{S}(Q)$; then $k = \gamma_{ij}(Q)$. Combining this with Theorem 3.1 (ii) or (iii) we obtain a simple, and polynomial-time, algorithm for deciding if Q induces a Bruhat face. By (iv) of Theorem 3.1, the usual matching algorithm for bipartite graphs also gives a polynomial-time algorithm.

Example 6. Consider again Example 5, and let $i = j = 3$. Then

$$P^{(3,3,0,6)} = \left[\begin{array}{c|c|c|c|c|c} & & & 1 & & \\ \hline & & & & 1 & \\ \hline & & & & & 1 \\ \hline 1 & & & & & \\ \hline & 1 & & & & \\ \hline & & 1 & & & \end{array}\right], \quad P^{(3,3,1,6)} = \left[\begin{array}{c|c|c|c|c|c} 1 & & & & & \\ \hline & & & 1 & & \\ \hline & & & & 1 & \\ \hline & 1 & & & & \\ \hline & & 1 & & & \\ \hline & & & & & 1 \end{array}\right].$$

As $P^{(3,3,0,6)} \nleqslant \mathcal{S}(Q)$, we conclude that $\gamma_{33}(Q) = 1$. As noted before, $\sigma_{33}(Q) = 2$, so Q does not induce a Bruhat face.

Define the *backward direct sum* $P_1 \oplus_b P_2$ of two square matrices P_1 and P_2 as the matrix

$$P_1 \oplus_b P_2 = \begin{bmatrix} O & P_1 \\ P_2 & O \end{bmatrix}.$$

More generally, for k square matrices P_i, $1 \leqslant i \leqslant k$, we define

$$P_1 \oplus_b \ldots \oplus_b P_k = (P_1 \oplus_b \ldots \oplus_b P_{k-1}) \oplus_b P_k.$$

Corollary 3.4. *Let r, s, t be nonnegative integers such as $r + s + t = n$. Then the permutation matrix*

$$Q = I_r \oplus_b L_s \oplus_b I_t \tag{3}$$

induces a Bruhat face whose shadow is given by $r_i = 1$ for $1 \leqslant i \leqslant r + s + 1$, $r_i = i - r - s$ for $r + s + 1 < i \leqslant n$ and $l_i = r + s + i$ for $1 \leqslant i < r$, $l_i = n$ for $r \leqslant i \leqslant n$.

P r o o f. By Theorem 1.2, the shadow of Q is as described in the statement of the corollary and thus is the $n \times n$ $(0,1)$-matrix which has zeros in its upper triangular right corner where I_r has zeros, zeros in its lower triangular left corner where I_t has zeros, and ones everywhere else. Using this characterization of the shadow of Q, the following calculations are straighforward to verify.

Let $1 \leqslant i, j \leqslant n$. We prove that $\sigma_{ij}(Q) = \gamma_{ij}(Q)$, and discuss different cases:

Case 1: $1 \leqslant i \leqslant r$, $1 \leqslant j \leqslant n-r$. Then $\sigma_{ij}(Q) = 0 = \gamma_{ij}(Q)$.

Case 2: $1 \leqslant i \leqslant r$, $n-r < j \leqslant n$. Then $\sigma_{ij}(Q) = \min\{i, j-n+r\}$ and due to the staircase pattern of $\mathcal{S}(Q)$, this coincides with $\gamma_{ij}(Q)$.

Case 3: $1 \leqslant i \leqslant n-t$, $1 \leqslant j \leqslant t$. Then $\sigma_{ij}(Q) = 0 = \gamma_{ij}(Q)$.

Case 4: $n-t < i \leqslant n$, $1 \leqslant j \leqslant t$. Then $\sigma_{ij}(Q) = \min\{i-n+t, j\}$ and due to the staircase pattern of $\mathcal{S}(Q)$, this coincides with $\gamma_{ij}(Q)$.

Case 5: $r < i \leqslant n$, $t < j \leqslant n$. Then $\sigma_{ij}(Q) = \max\{i+j-n, 0\}$. On the other hand, any $P \in \mathcal{P}_n$ contains at most $n-i$ ones in rows $i+1, i+2, \ldots, n$ and at most $n-j$ ones in columns $j+1, j+2, \ldots, n$. Therefore such P contains at least $n-(n-i)-(n-j) = i+j-n$ in its leading $i \times j$ submatrix. So $\gamma_{ij}(Q) \geqslant i+j-n$. Since $\gamma_{ij}(Q) \geqslant 0$, this shows that $\sigma_{ij}(Q) = \max\{i+j-n, 0\} \leqslant \gamma_{ij}(Q)$, but then equality must hold here (as the opposite inequality holds by definition of $\gamma_{ij}(Q)$).

This proves that $\Sigma(Q) = \Gamma(Q)$, and the theorem follows. □

Using Corollary 3.4 and the fact that the property of inducing a Bruhat face is preserved under taking direct sums, one may construct several permutation matrices that induce Bruhat faces, as illustrated in the next example.

Example 7. (i) I_n and L_n induce Bruhat faces; see Corollary 3.4 with $r = n$, $s = t = 0$ and $s = n$, $r = t = 0$, respectively. Therefore the direct sum $I_s \oplus L_r \oplus I_t$ also induces a Bruhat face.

(ii) $Q = I_r \oplus_b I_t$, where $r + t = n$, induces a Bruhat face ($s = 0$ in the corollary). In particular, with $t = 1$ one obtains a Hessenberg matrix, for instance

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

(iii) The matrix $P^{(i,j,k,n)}$ in (2) induces a Bruhat face because

$$P^{(i,j,k,n)} = I_k \oplus (I_{i-k} \oplus_b I_{j-k}) \oplus I_{n-i-j+k},$$

so it is the direct sum of identity matrices and the matrix in (3) with $s = 0$.

For $n \leqslant 3$ one can check that every permutation matrix induces a Bruhat face (since it can be obtained from I_r and L_s (with $r + s = n$) by taking direct sum or backward direct sum). An example of a matrix which is not obtained using the constructions above, but still induces a Bruhat face is

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

For $n = 4$ the only permutation matrix that does not induce a Bruhat face is

$$P = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

But this is for the obvious reason that P is obtained by an "internal" backward interchange from L_4, and this does not change the Bruhat shadow. We say that a permutation matrix is *shadow-maximal* if it allows no forward interchange within its Bruhat shadow (replacing a submatrix I_2 by L_2). Clearly, a necessary condition for a matrix to induce a Bruhat face is that it is shadow-maximal. Therefore a permutation of the form $I_r \oplus_b P \oplus_b I_s$ induces a Bruhat face if and only if P is the L permutation. But this condition (being shadow-maximal) is not sufficient. The matrix

$$P = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

is shadow-maximal, but it does not induce a Bruhat face. Indeed, the matrix $Q = L_4 \oplus I_1$ is whitin the Bruhat shadow of P, but we have $P \not\preceq_B Q$ and $Q \not\preceq_B P$. Also, in general, backward direct sums of permutations do not induce Bruhat faces, see the 4×4 matrix $I_1 \oplus_b I_2 \oplus_b I_1$ shown above.

A class of permutation matrices that induce Bruhat faces is discussed next. Let $\pi = (i_1, i_2, \ldots, i_n)$ be a permutation of $\{1, 2, \ldots, n\}$. Then i_k, i_{k+1} is a *descent* of π if $i_k > i_{k+1}$; we also say that a descent occurs at position k. Here $1 \leqslant k \leqslant n-1$. A permutation is a *grassmanian* provided it has exactly one descent. We say that a permutation matrix Q is a grassmanian when its corresponding permutation is a grassmanian; if the unique descent of the permutation occurs at position k, then Q has a unique descent at row k. For example, with $n = 12$, $\sigma = (3, 6, 7, 9, 10, 1, 2, 4, 5, 8, 11, 12)$ is a grassmanian whose unique descent occurs at $k = 5$. Another example is the matrix $P^{(i,j,k,n)}$ defined in (2). The permutation matrix corresponding to σ also with the zeros defining its shadow is:

$$Q = \left[\begin{array}{c|c|c|c|c|c|c|c|c|c|c|c}
 & & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline
 & & & & & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline
 & & & & & & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline
 & & & & & & & & 1 & 0 & 0 & 0 \\ \hline
 & & & & & & & & & 1 & 0 & 0 \\ \hline
1 & & & & & & & & & & & \\ \hline
0 & 1 & & & & & & & & & & \\ \hline
0 & 0 & 0 & 1 & & & & & & & & \\ \hline
0 & 0 & 0 & 0 & 1 & & & & & & & \\ \hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & & & & \\ \hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \\ \hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}\right].$$

Thus, the permutation matrices P with $P \leqslant \mathcal{S}(Q)$ are those whose 1's are 1's of Q or are in the empty positions. In the proof of the next theorem it may be helpful to refer to this example.

Theorem 3.5. *Let $Q = [q_{ij}]$ be an $n \times n$ grassmanian permutation matrix. Then Q induces a Bruhat face, so*

$$\Omega_n(\preceq_B Q) = \Omega_n\ (\geqslant \Sigma(Q)) = \Omega_n^{S(Q)}.$$

P r o o f. Let P be a permutation matrix with $P \preceq_B Q$. Then by definition, $P \leqslant \mathcal{S}(Q)$. Now suppose that $P \leqslant \mathcal{S}(Q)$ and P corresponds to the permutation $(i_1, i_2, \ldots, i_n)$. To complete the proof we show that $P \preceq_B Q$ or equivalently, by Theorem 1.1, that $\Sigma(P) \geqslant \Sigma(Q)$.

Since Q is a grassmanian permutation matrix, it has a unique descent, say at row k. Since P is a permutation matrix and $P \leqslant \mathcal{S}(Q)$, it follows that

$$\sigma_{ij}(P) \geqslant \sigma_{ij}(Q) \quad \text{if either } 1 \leqslant i \leqslant k \ \text{ or } 1 \leqslant j < i_1.$$

Now assume that $i > k$ and $j \geqslant i_1$. We claim that the term rank of the matrix $\mathcal{S}(Q)_{ij}$ obtained from $\mathcal{S}(Q)$ by replacing its leading $i \times j$ submatrix with a zero matrix is at most $n - \sigma_{ij}(Q)$, that is, n minus the number of 1's of Q in its leading $i \times j$ submatrix. This follows from the assumption that Q is grassmanian, since we can then cover all the 1's of $\mathcal{S}(Q)_{ij}$ with $(n-j)$ columns $j+1, j+2, \ldots, n$ (so each containing a 1 of Q) and $(j - \sigma_{ij}(Q))$ rows $u > i$ which contain a 1 in columns $1, 2, \ldots, j$. Thus, $\mathcal{S}(Q)_{ij}$ has term rank $n - \sigma_{ij}(Q)$ proving the claim. Hence, any permutation matrix $P \leqslant \mathcal{S}(Q)$ contains $\sigma_{ij}(Q)$ 1's in its leading $i \times j$ submatrix. Then (see also (iv) in Theorem 3.1) we conclude that $\Sigma(P) \geqslant \Sigma(Q)$ and hence $P \preceq_B Q$. $\square$

We note that if a permutation matrix has more than one descent, it may, or may not, induce a Bruhat face. For instance, the permutation matrix in Example 5 does not induce a Bruhat face and it has two descents. The permutation matrix

$$L_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

has two descents and induces a Bruhat face, namely Ω_3 itself.

Finally, we mention that the Bruhat order may be extended to the class $\mathcal{N}(R, S)$ of nonnegative matrices with row sum vector R and column sum vector S, the class of transportation matrices, and a study of this partial order is ongoing work. An interesting topic is to study the convex hull of $(\preceq Q)$, where Q is an extreme point, by linear constraints and determine its extreme points.

Acknowledgement. We are indebted to a referee who read this paper thoroughly and commented in substantial detail.

References

[1] *A. Björner, F. Brenti*: Combinatorics of Coxeter Groups. Graduate Texts in Mathematics 231, Springer, New York, 2005. zbl MR

[2] *A. Björner, F. Brenti*: An improved tableau criterion for Bruhat order. Electron. J. Comb. *3* (1996), Research paper R22, 5 pages; printed version J. Comb. *3* (1996), 311–315. zbl MR

[3] *R. A. Brualdi*: Combinatorial Matrix Classes. Encyclopedia of Mathematics and Its Applications 108, Cambridge University Press, Cambridge, 2006. zbl MR

[4] *R. A. Brualdi, G. Dahl*: The Bruhat shadow of a permutation matrix. Mathematical Papers in Honour of Eduardo Marques de Sá. Textos de Matemática. Série B 39, Universidade de Coimbra, Coimbra, 2006, pp. 25–38. zbl MR

[5] *R. A. Brualdi, L. Deaett*: More on the Bruhat order for $(0,1)$-matrices. Linear Algebra Appl. *421* (2007), 219–232. zbl MR

[6] *R. A. Brualdi, S.-G. Hwang*: A Bruhat order for the class of $(0,1)$-matrices with row sum vector R and column sum vector S. Electron. J. Linear Algebra (electronic only) *12* (2004/2005), 6–16. zbl MR

[7] *P. Magyar*: Bruhat order for two flags and a line. J. Algebr. Comb. *21* (2005), 71–101. zbl MR

Authors' addresses: Richard A. Brualdi, Department of Mathematics, 480 Lincoln Drive, University of Wisconsin, Madison, Wisconsin 53706, USA, e-mail: `brualdi@math.wisc.edu`; Geir Dahl, Department of Mathematics, University of Oslo, Moltke Moes vei 35, 0851 Oslo, Norway, e-mail: `geird@math.uio.no`; Eliseu Fritscher, Programa de Engenharia de Produção/COPPE, Av. Athos da Silveira Ramos, 149, Centro de Tecnologia, Universidade Federal do Rio de Janeiro, CEP: 21941-909, Rio de Janeiro, RJ, Brazil, e-mail: `eliseu.fritscher@ufrgs.br`.