



Joint mixability of some integer matrices



Fabio Bellini^{a,b}, Oya Ekin Kardeşan^b, Mustafa Ç. Pınar^{b,*}

^a Department of Quantitative Methods and Statistics, University of Milano, Bicocca, Italy

^b Department of Industrial Engineering, Bilkent University, 06800 Bilkent, Ankara, Turkey

ARTICLE INFO

Article history:

Received 24 July 2015

Received in revised form 25 March 2016

Accepted 26 March 2016

Available online 22 April 2016

MSC:

05B20

15B36

11C20

60C05

90C47

90C59

Keywords:

Jointly mixable matrices

Quantiles

Column permutation

Minimax row sum

ABSTRACT

We study the problem of permuting each column of a given matrix to achieve minimum maximal row sum or maximum minimal row sum, a problem of interest in probability theory and quantitative finance where quantiles of a random variable expressed as the sum of several random variables with unknown dependence structure are estimated. If the minimum maximal row sum is equal to the maximum minimal row sum the matrix has been termed *jointly mixable* (see e.g. Haus (2015), Wang and Wang (2015), Wang et al. (2013)). We show that the lack of joint mixability (the joint mixability gap) is not significant, i.e., the gap between the minimum maximal row sum and the maximum minimal row sum is either zero or one for a class of integer matrices including binary and complete consecutive integers matrices. For integer matrices where all entries are drawn from a given set of discrete values, we show that the gap can be as large as the difference between the maximal and minimal elements of the discrete set. The aforementioned result also leads to a polynomial-time approximation algorithm for matrices with restricted domain. Computing the gap for a $\{0, 1, 2\}$ -matrix is proved to be equivalent to finding column permutations minimizing the difference between the maximum and minimum row sums. A polynomial procedure for computing the optimum difference by solving the maximum flow problem on an appropriate graph is given.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction and background

We consider the problem of permuting each column of a given matrix to achieve minimum maximal row sum or maximum minimal row sum, a problem of recent interest in quantitative finance (see e.g., [1–3]) where quantiles of a random variable expressed as the sum of several random variables with unknown dependence structure are estimated. If the minimum maximal row sum is equal to the maximum minimal row sum the matrix is termed *jointly mixable*, a notion first introduced in [4] for general families of probability distributions. In this paper inspired by the recent work of Haus [2], we develop the study of joint mixability

* Corresponding author.

E-mail addresses: fabio.bellini@unimib.it (F. Bellini), karaslan@bilkent.edu.tr (O.E. Kardeşan), mustafap@bilkent.edu.tr (M.Ç. Pınar).

of some classes of integer matrices in a novel direction: we focus on the lack of mixability. After a brief introduction, we start in Section 2 with a result concerning binary matrices where we establish that the lack of joint mixability is not significant, i.e., the gap between the minimum maximal row sum and the maximum minimum row sum is either zero or one. Since a necessary and sufficient condition for joint mixability of binary matrices is known, one can immediately conclude that the gap is equal to one if the condition fails. Besides, optimal permutations achieving the gap can be obtained in linear time. A similar conclusion holds for complete consecutive integer matrices, a class of integer matrices defined in [2], as we establish in Section 3. In a generalization to integer matrices where all entries are drawn from a given set of discrete values, we show in Section 4 that the gap between the optimized minimum and maximum row sums can be as large as the difference between the minimal and maximal elements of the discrete set. This observation leads to a polynomial time approximation algorithm. For matrices with values from the set $\{0, 1, 2\}$ (termed *two-ary* matrices) we prove in Section 5 that computing the gap is equivalent to finding column permutations minimizing the difference between the maximum and minimum row sums. We also describe a polynomial-time procedure to compute the difference between the minimum maximal row sum and the maximum minimal row sum using so-called *swap operations* that can be implemented within a maximum-flow algorithm for an appropriately defined capacitated network for a given problem instance. Finally, in Section 6 we hint at the challenging nature of the problem by showing that the linear programming bound is trivial.

It is hoped that the present paper will spark renewed interest in the discrete mathematics community for this fascinating problem of considerable importance in statistics and quantitative finance. An attractive feature of the present paper is the elementary nature of proofs which rely on simple combinatorial arguments and involve at most some linear programming.

There is already considerable research activity on the continuous space counterpart of the problems of this paper, see e.g., the recent papers [5–7,4]. The standard reference on joint mixability in probability theory is [7]. In an atomless probability space, where random variables X_i which are distributed according to given univariate distributions F_i , $i = 1, \dots, d$ are given, one asks the question to determine whether a given distribution F is a possible distribution for the sum $S = X_1 + \dots + X_d$. While the problem has a long history (see e.g., the introduction of [7]) Wang and Wang [7] partially answer the above question recently using the theory of *joint mixability*. In their context a vector (X_1, \dots, X_d) is a joint mix if $X_1 + \dots + X_d$ is almost surely a constant. A d -tuple of distributions (F_1, \dots, F_d) is said to be *jointly mixable* if there exists a joint mix with univariate marginal distributions F_1, \dots, F_d . Ref. [7] reformulates the original question above in terms of the equivalent question of determining joint mixability of an n -tuple of distributions.

In the present paper we are concerned with the problem of determining *joint mixability* of matrices, which is described as the following pair of optimization problems: given a matrix $A \in \mathbb{R}^{m \times d}$, (a) find independently a permutation for each column of A such that the maximal row sum of the resulting matrix is minimized and (b) find independently a permutation for each column of A such that the minimal row sum of the resulting matrix is maximized. Let the permutations of the d columns of A be denoted as a *permutation system* $\Pi = (\pi_1, \dots, \pi_d)$, and refer to as A^Π the permuted matrix obtained by permuting column k by permutation matrix π_k , for $k = 1, \dots, d$. Hence the optimization problems we are interested in are:

$$\gamma(A) = \min_{\Pi} \max_{i=1, \dots, m} \left\{ \sum_{j=1}^d A_{ij}^{\Pi} \right\}, \quad (1)$$

and

$$\beta(A) = \max_{\Pi} \min_{i=1, \dots, m} \left\{ \sum_{j=1}^d A_{ij}^{\Pi} \right\}. \quad (2)$$

When $\gamma(A) = \beta(A)$, the matrix A is said to be *jointly mixable*, and the resulting permuted matrix is termed a *joint mix*. Naturally, the matrix problems treated in the present paper are discrete space analogs of the

probabilistic problems studied in the recent Refs. [5–7,4]. We point out these analogies after the relevant result throughout the paper.

Besides being a discrete space counterpart of the continuous space problems briefly described above, the problem of determining joint mixability of matrices is also closely related to the statistical problem of estimating quantiles of random variables obtained by aggregation of several random variables with an unknown dependence structure. More precisely, consider the following problem: let S be an aggregate random variable described as $S = \sum_{i=1}^d X_i$ where the random variables X_i are distributed according F_i . One does not know the dependence among random variables L_i nor the joint distribution function F_S of S . One is typically interested in computing the so-called Value-at-Risk (VaR) $F_S^{-1}(\alpha) = \inf\{x \in \mathbb{R} | F_S(x) \geq \alpha\}$ for $\alpha \in (0, 1)$. Assuming the marginal distributions to be discrete (or discretized according to the procedure described e.g., in [8]), one computes the values $q_r^i = F_i^{-1}(r/N)$ for $r = 1, \dots, N$. Dependence among the constituent random variables X_i is reflected in the matrix A whose i th column is the vector $(q_0^i, q_1^i, \dots, q_N^i)^T$:

$$A = \begin{pmatrix} q_0^1 & \cdots & q_0^d \\ \vdots & & \vdots \\ q_N^1 & \cdots & q_N^d \end{pmatrix}.$$

In order to find the tightest bounds on F_S^{-1} one needs to solve the problem of minimizing the variance of the row sums of A or, equivalently, [9] minimize the difference $\gamma(A) - \beta(A)$ over all arrangements of the F_i 's.

The interested reader is directed to Refs. [2,8] for further details and to [5,1,6,3,4,14] for applications in insurance and finance. A recent paper [10] surveys the main results and open questions on joint mixability (and the related concept of complete mixability). The site <https://sites.google.com/site/rearrangementalgorithm/home> contains pointers to the literature and applications of the problem as well as implementations of algorithms related to the problem. The web site title is in reference to the so-called *rearrangement* algorithm (see [8]) that can compute bounds on $\gamma(A)$ and $\beta(A)$ quite fast. On the other hand, Haus [2] shows that the rearrangement algorithm may terminate with a large error proportional to the largest entry in the matrix; cf. Lemma 4 of [2].

The problem has been studied earlier under the title of assembly line crew scheduling in [11]. In this application, the rows of the matrix A correspond to assembly lines and the columns correspond to operations; the numbers in the j th column indicate the times required by the individual members of the crew performing the j th operation. The objective is to assign the members in each crew so as to minimize the maximum time required to produce an item over the m assembly lines. It was shown in [11] that the classical makespan scheduling problem is a special case of this NP-complete problem. In a related study Coffman and Yannakakis [12] suggested three approximation algorithms to compute a permutation system that minimizes the maximum row sum:

1. “Algorithm D”: Order the first two columns in opposite way. Then order the third oppositely to the sum of the first two, and repeat until the last column is processed. This is very similar to Rüschemdorf’s “rearrangement algorithm” [9], with the difference that in “D” each column is processed only once. In [11] algorithm D was shown to have an approximation bound equal to $2 - 1/m$. We utilize a similar procedure in the proof of [Theorem 4](#).
2. “Algorithm L”: This is a greedy algorithm of the “mark and move” type. The elements in the whole matrix are marked in decreasing order. When an element is marked, it is exchanged in its column with an unmarked one, and goes to the row which has the minimal sum of marked elements.
3. “Algorithm RS”: This algorithm is more complicated than the previous two in that there is a pre-processing phase in which the biggest m elements of the matrix are assigned one for each row, to prevent the sums of two big values. Then the row sums are examined sequentially and improved by a certain fixed pattern of swaps.

They were able to prove that “RS” is better than “L” and “D” in the sense that asymptotically in the worst case its optimal value is $3/2$ the true minimal max row sum.

For $A \in \mathbb{Z}^{m \times 3}$ Haus [2] gives a polynomial 2-approximation algorithm for computing $\gamma(A)$.

1.1. Notation and elementary observations

We refer to the difference $\mathcal{G}(A) = \gamma(A) - \beta(A)$ for a given matrix A as the *joint mixability gap* of A . Let $s_i = \sum_{j=1}^d A_{ij}$ for $i \in \{1, \dots, m\}$ be the i th row sum and s_i^Π be the corresponding sum under permutation Π . Let $s = \sum_{i=1}^m s_i$ and $r = s/m$.

Observation 1. For any integer matrix A , $\gamma(A) \geq \lceil \frac{s}{m} \rceil$ and $\beta(A) \leq \lfloor \frac{s}{m} \rfloor$. In other words, if $m \nmid s$, then $\mathcal{G}(A) \geq 1$.

For $A \in \mathbb{R}^{m \times d}$, consider a permutation such that the maximum difference between the row sums is minimized. Let $\mathcal{K}(A)$ reflect this difference. More formally,

$$\mathcal{K}(A) = \min_{\Pi} \max_{i,j=1,\dots,m} \left\{ \sum_{k=1}^d A_{ik}^\Pi - \sum_{k=1}^d A_{jk}^\Pi \right\}. \tag{3}$$

Let Π^γ denote a permutation with the largest row sum equal to $\gamma(A)$, Π^β a permutation with the smallest row sum equal to $\beta(A)$ and $\Pi^\mathcal{K}$ the permutation with the largest row sum difference equal to $\mathcal{K}(A)$.

Without loss of generality we may assume that $s_1^{\Pi^\gamma} \geq s_2^{\Pi^\gamma} \geq \dots \geq s_m^{\Pi^\gamma}$, $s_1^{\Pi^\beta} \geq s_2^{\Pi^\beta} \geq \dots \geq s_m^{\Pi^\beta}$ and $s_1^{\Pi^\mathcal{K}} \geq s_2^{\Pi^\mathcal{K}} \geq \dots \geq s_m^{\Pi^\mathcal{K}}$.

The following results provide relationships between the numbers \mathcal{G} and \mathcal{K} .

Lemma 1. For any $A \in \mathbb{R}^{m \times d}$, $\mathcal{G}(A) \leq \mathcal{K}(A)$.

Proof. Since $s_1^{\Pi^\mathcal{K}} \geq s_2^{\Pi^\mathcal{K}} \geq \dots \geq s_m^{\Pi^\mathcal{K}}$, $\gamma(A) \leq s_1^{\Pi^\mathcal{K}}$ and $\beta(A) \geq s_m^{\Pi^\mathcal{K}}$ and thus $\mathcal{G}(A) \leq \mathcal{K}(A)$. ■

Lemma 2. For any $A \in \mathbb{R}^{m \times d}$, $\mathcal{G}(A) = 0$ if and only if $\mathcal{K}(A) = 0$.

Proof. Assume $\mathcal{G}(A) = 0$. In other words, $s_1^{\Pi^\gamma} = s_m^{\Pi^\beta}$. Since $\sum_{k=1}^m s_k^{\Pi^\gamma} = \sum_{k=1}^m s_k^{\Pi^\beta}$, we must have $s_i^{\Pi^\gamma} = s_j^{\Pi^\gamma}$ and $s_i^{\Pi^\beta} = s_j^{\Pi^\beta}$ for all $i \neq j$ and thus $\mathcal{K}(A) = 0$. The other direction is a simple corollary of Lemma 1. ■

2. Binary matrices

Recently, Haus [2] proved the following theorem for binary matrices $A \in \mathbb{B}^{m \times d}$, where $\mathbb{B} = \{0, 1\}$.

Theorem 1 (Haus [2]). $A \in \mathbb{B}^{m \times d}$ is jointly mixable if and only if $m \mid s$. The permutation achieving the joint mix can be computed in linear time $\mathcal{O}(m \cdot d)$.

Our first result is the following.

Theorem 2. If $A \in \mathbb{B}^{m \times d}$ is not jointly mixable, then $\gamma(A) = \lceil \frac{s}{m} \rceil$, $\beta(A) = \lfloor \frac{s}{m} \rfloor$, i.e., $\mathcal{G}(A) = 1$. Furthermore, the optimal permutation achieving the gap can be computed in linear time $\mathcal{O}(m \cdot d)$.

Proof. We borrow from the proof of Theorem 1 of [2]. We first deal with the problem (1). Let $\bar{r} = \lceil \frac{s}{m} \rceil$. Define the *defect* $\delta(i)$ of row i as $\delta(i) = \bar{r} - \sum_{j=1}^d A_{ij}$. Let the total absolute defect be $\phi = \sum_{i=1}^m |\delta(i)|$.

Now, consider the following procedure: starting from column $j = 1$, define $S_j = \{i \in \{1, \dots, m\} | \delta(i) < 0, A_{ij} = 1\}$, and $D_j = \{i \in \{1, \dots, m\} | \delta(i) > 0, A_{ij} = 0\}$. If both S_j and D_j are non-empty, let $t_j = \min\{|S_j|, |D_j|\}$, and swap the first t_j entries indexed by S_j and D_j in column j (i.e., if $t_j = 1$, swap the entry in column j indexed by the first element in S_j with the entry in column j indexed by the first entry in D_j ; if $t_j = 2$, then do the previous step with the first respective entries in S_j and D_j , and then for the second respective entries in S_j and D_j , and so on). After the swapping is performed, update the defects δ for the rows involved in the swap. Repeat this step for all $j = 2, 3, \dots, d$.

Clearly, after each swap the defect of rows with positive defect will decrease and the defect of rows with negative defect will increase. Therefore, after each column where at least one swap is performed, the total absolute defect ϕ decreases. Assume that the procedure stops at the last column and there is a row i_2 with a positive defect $\delta(i_2)$ and a row with a negative defect $\delta(i_1)$. Since $s_{i_2} < s_{i_1}$ there must be a column index l such that $A_{i_2,l} = 0$ and $A_{i_1,l} = 1$ since otherwise the rows i_1 and i_2 would have to be identical and would not have defects of opposite sign. Hence, the row i_1 should be in S_l and the row i_2 should be in D_l , and should have been involved in a swap, a contradiction. Therefore, when the procedure terminates, one can only have all non-negative or all non-positive defects. Since $\frac{s}{m} < \lceil \frac{s}{m} \rceil < \frac{s}{m} + 1$, at least one of these defects should be equal to zero.

The proof for the problem (2) is verbatim repetition of the previous arguments with defects defined using $\lfloor \frac{s}{m} \rfloor$. ■

Combining Theorems 1 and 2, we know that for a given binary matrix joint mixability and the joint mixability gap are decided following the result of a division. Besides, optimal permutations are found in linear time.

Corollary 1. For a binary matrix $A \in \mathbb{B}^{m \times d}$, $\mathcal{G}(A) = \mathbb{1}_{m \nmid s}$.

Below in Section 4, we shall obtain the above corollary as a consequence of a more general result.

3. Complete consecutive integers matrices

Let $A \in \mathbb{Z}_{>0}^{m \times d}$ be an $m \times d$ complete consecutive integers matrix [2]. These matrices are characterized by the property that each column is a permutation of the first m natural numbers, that is $\{1, \dots, m\}$. Haus [2] showed how to compute explicitly $\gamma(A)$ and $\beta(A)$ for some instances of such matrices.

We shall show that any complete consecutive integers matrix (for $d \geq 2$) is either jointly mixable or has $\mathcal{G}(A) = 1$. We first need the following observations.

Observation 2. If A is a complete consecutive integers matrix and $d = 1$ then $\gamma(A) = m$ and $\beta(A) = 1$.

Lemma 3. If A is a complete consecutive integers matrix and d is even, then A is jointly mixable.

Proof. It is easy to provide such a permutation. For every two consecutive columns i and $i + 1$ such that i is odd, permute the rows in column i in increasing order, and those in column $i + 1$ in decreasing order. Then, under this permutation Π , $s_k^\Pi = \frac{d}{2}(m + 1)$ for every $k \in \{1, \dots, m\}$. ■

Lemma 4. Let A be an $m \times 3$ complete consecutive integers matrix. Then, A is jointly mixable if and only if m is odd. Moreover, if m is even, then $\mathcal{G}(A) = 1$.

Proof. For any m we shall construct a complete consecutive integers matrix satisfying the property that can easily be attained by independent column permutations.

Case 1: m is odd:

Let

$$A_{i1} = m - i + 1 \quad \forall i \in \{1, \dots, m\}$$

$$A_{i2} = \begin{cases} \frac{i+1}{2} & \text{if } i \text{ is odd} \\ \frac{m+i+1}{2} & \text{if } i \text{ is even} \end{cases} \quad \forall i \in \{1, \dots, m\}$$

and

$$A_{i3} = \begin{cases} \frac{m+i}{2} & \text{if } i \text{ is odd} \\ \frac{i}{2} & \text{if } i \text{ is even} \end{cases} \quad \forall i \in \{1, \dots, m\}.$$

With this permutation, for every i , $s_i = \frac{3}{2}(m+1)$ and $\gamma(A) = \beta(A) = \frac{3}{2}(m+1)$.

Case 2: m is even:

Let

$$A_{i1} = m - i + 1 \quad \forall i \in \{1, \dots, m\}$$

$$A_{i2} = \begin{cases} \frac{i+1}{2} & \text{if } i \text{ is odd} \\ \frac{m+i}{2} & \text{if } i \text{ is even} \end{cases} \quad \forall i \in \{1, \dots, m\}$$

and

$$A_{i3} = \begin{cases} \frac{m+i+1}{2} & \text{if } i \text{ is odd} \\ \frac{i}{2} & \text{if } i \text{ is even} \end{cases} \quad \forall i \in \{1, \dots, m\}.$$

With this permutation, for odd i , $s_i = \frac{3}{2}m + 2$ and for even i , $s_i = \frac{3}{2}m + 1$ and hence $\mathcal{G}(A) = 1$. ■

We can now extend [Theorem 1](#) of Haus [2] to complete consecutive integers matrices. In particular, we have the following.

Theorem 3. A complete consecutive integers matrix $A \in \mathbb{Z}_{>0}^{m \times d}$ for $d \geq 2$ is jointly mixable if and only if $m \mid s$. The permutation achieving the joint mix can be computed in linear time $\mathcal{O}(m \cdot d)$. Moreover, if the condition is not satisfied, then $\mathcal{G}(A) = 1$.

Proof. Note that $s = \frac{m(m+1)d}{2}$ and $m \mid s$ if and only if m is odd or d is even. If d is even, then the permutation in the proof of [Lemma 3](#) leads to a joint mix. If d is odd (say $d > 3$) and m is odd, then starting with the permutation given in the proof of [Lemma 4](#) for the first 3 columns, the remaining $d - 3$ columns could be paired up with alternating permutations of increasing and decreasing orderings as in the proof of [Lemma 3](#) resulting in a permutation with $\gamma(A) = \beta(A) = \frac{(m+1)d}{2}$. If d is odd (say $d > 3$) and m is even, i.e. the only time the condition of the theorem’s statement is violated, then the permutation provided in the proof of [Lemma 4](#) for 3 columns could be augmented with alternating permutations of increasing and decreasing orderings of the remaining $d - 3$ columns resulting in $s_i = \frac{3}{2}m + 2 + \frac{(d-3)(m+1)}{2}$ for an odd row i , $s_i = \frac{3}{2}m + 1 + \frac{(d-3)(m+1)}{2}$ for an even row i and hence $\mathcal{G}(A) = 1$. ■

Notice that, as pointed out by an anonymous referee, an analogous result for continuous distribution has been obtained in Theorem 3.2 of [7], where joint mixability for continuous uniform distribution has been completely characterized.

In the following lemma we collect some elementary properties of joint mixability that will be useful in the sequel; we refer to Proposition 2.3 in [7] for proofs and generalizations.

Lemma 5. *Let A be a real $m \times d$ matrix $[A_{ij}]$. Then, the joint mixability property is preserved with the following perturbations:*

1. *Let $B = [B_{ij}]$ be an $m \times d$ matrix such that $B_{ij} = A_{ij} + u_j$ for some $u_j \in \mathbb{R}$, $j = 1, \dots, d$. Then, B is jointly mixable if and only if A is.*
2. *Let $C = [C_{ij}]$ be an $m \times d$ matrix such that $C_{ij} = tA_{ij}$ for some nonnegative $t \in \mathbb{R}$. Then, C is jointly mixable if and only if A is.*
3. *Let D be an $m \times (d - 1)$ matrix attained from A by deleting a column of identical entries, say $v \in \mathbb{R}$. Then, D is jointly mixable if and only if A is.*

The following matrices generalize complete consecutive integers matrices.

Definition 1. Let $A = [A_{ij}] \in \mathbb{R}^{m \times d}$ be such that $A_{ij} = u_j + ti$ for $i = 1, \dots, m$, $j = 1, \dots, d$, $u_j \in \mathbb{R}$ and $t \in \mathbb{R}_{\geq 0}$. We call such a matrix a complete uniform gap matrix.

In particular, a complete consecutive integers matrix is a complete uniform gap matrix with $u_j = 0$ for $j = 1, \dots, d$ and $t = 1$ and the following result immediately follows from Theorem 3 and Lemma 5.

Corollary 2. *A complete uniform gap matrix A , i.e., a matrix $A \in \mathbb{R}^{m \times d}$ such that $A_{ij} = u_j + ti$ for $i = 1, \dots, m$, $j = 1, \dots, d$, $u_j \in \mathbb{R}$ and $t \in \mathbb{R}_{\geq 0}$, is jointly mixable if and only if $m \mid s$. The permutation achieving the joint mix can be computed in linear time $\mathcal{O}(m \cdot d)$. Moreover, if the condition is not satisfied, then $\mathcal{G}(A) = t$.*

Proof. A can be attained by multiplying each entry of a complete consecutive integers matrix of the same dimensions by t and adding u_j to each cell of column j of the resulting matrix. Clearly, these operations change the gap by the multiplicative factor t . ■

4. Matrices with restricted domain

We now consider matrices whose elements are from a set of discrete values $M = \{v_1, \dots, v_p\}$. Such matrices encompass binary matrices and complete consecutive integers matrices.

Let $M = \{v_1, \dots, v_p\} \subseteq \mathbb{R}$ be a fixed set of values such that $v_1 > v_2 > \dots > v_p$ and $A \in \mathcal{M}^{m \times d}$ denote a matrix with every column coming from the set M .

Theorem 4. *For any matrix $A \in \mathcal{M}^{m \times d}$, $\mathcal{K}(A) \leq v_1 - v_p$.*

Proof. We shall proceed to show this result using induction on d . Clearly, for any $A \in \mathcal{M}^{m \times 1}$, maximum row sum $\leq v_1$ and minimum row sum $\geq v_p$ and thus the difference is at most $v_1 - v_p$. Assume inductively that for any matrix of at most $d - 1$ columns satisfying the statement of the theorem, there exists a permutation such that the difference between the maximum and minimum row sums is at most $v_1 - v_p$. Without loss of generality, assume that s_i is the i th row sum in this permutation and that $s_1 \geq s_2, \dots, \geq s_m$. Now, if the d^{th} column is permuted such that $A_{1d} \leq A_{2d} \leq \dots \leq A_{md}$, then the difference between the i th and j th row

sums is $s_i + A_{id} - (s_j + A_{jd})$ which is clearly at most $v_1 - v_p$ since $s_i - s_j \leq v_1 - v_p$ and $A_{id} - A_{jd} \leq 0$. Note that the procedure employed in this proof is essentially “Algorithm D” of [12]. ■

The probabilistic version of this theorem is Corollary A.3 of [6].

Corollary 3. For any matrix $A \in \mathcal{M}^{m \times d}$, a permutation satisfying $\mathcal{G}(A) \leq v_1 - v_p$ can be achieved in $\mathcal{O}(d \cdot m \log m)$ time complexity.

Proof. In order to permute each column, it is enough to order the elements in this column and the row sums in the previous columns, hence the result follows. ■

In particular, applying this corollary to a binary matrix $A \in \mathbb{B}^{m \times d}$ we obtain Theorem 2 (with a difference in computational complexity) in an alternative fashion. The following result is now immediate from Theorem 4.

Theorem 5. For any matrix $A \in \mathcal{M}^{m \times d}$ there is a $\frac{r+v_1-v_p}{r}$ -approximation algorithm for problem (1) and a $\frac{r}{r-(v_1-v_p)}$ -approximation algorithm for problem (2).

Proof. Given A , let $\Pi^{\mathcal{K}}$ be the permutation achieving the value $\mathcal{K}(A)$ as described in the proof of Theorem 4 and without loss of generality assume that $s_1^{\Pi^{\mathcal{K}}} \geq s_2^{\Pi^{\mathcal{K}}} \geq \dots \geq s_m^{\Pi^{\mathcal{K}}}$. Then,

$$\frac{s_1^{\Pi^{\mathcal{K}}}}{\gamma(A)} \leq \frac{s_m^{\Pi^{\mathcal{K}}} + v_1 - v_p}{\gamma(A)} \leq \frac{r + v_1 - v_p}{\gamma(A)} \leq \frac{r + v_1 - v_p}{r}$$

since $s_m^{\Pi^{\mathcal{K}}} \leq \beta(A) \leq r$ and $\gamma(A) \geq r$. The procedure for computing $\Pi^{\mathcal{K}}$ clearly runs in polynomial time. The result for problem (2) follows similarly. ■

The subclass of integer matrices with entries from $\{0, 1, 2\}$, i.e. *two-ary* matrices, constitute an interesting class of restricted domain matrices. We can use this subclass to show that the bound given in Theorem 4 is tight. Consider the following instance of a 7×4 matrix with entries from the set $\{0, 1, 2\}$:

$$A = \begin{bmatrix} 2 & 0 & 1 & 1 \\ 2 & 1 & 1 & 0 \\ 2 & 2 & 1 & 0 \\ 2 & 2 & 0 & 0 \\ 1 & 2 & 1 & 1 \\ 0 & 2 & 1 & 0 \\ 0 & 2 & 1 & 0 \end{bmatrix}.$$

The matrix has $s = 28$. However, it is not jointly mixable and has the gap $\mathcal{G}(A) = 2$ as can be verified by direct calculation ($\gamma(A) = 5$ and $\beta(A) = 3$). The example also shows that the characterization given by Haus in Theorem 1 for binary matrices is no longer valid.

Other, smaller examples can also be found: e.g.,

$$A = \begin{bmatrix} 2 & 0 & 2 \\ 2 & 2 & 1 \\ 1 & 2 & 1 \\ 0 & 2 & 1 \end{bmatrix}.$$

Lemma 6. Let A be an $m \times d$ integer matrix with elements from the set $\{0, 1, 2\}$. If $m \mid s$ then either $\mathcal{G}(A) = 0$ or $\mathcal{G}(A) = 2$.

Proof. We know that the gap is bounded above by 2. Let $A \in \mathbb{Z}^{m \times d}$ be an arbitrary $\{0, 1, 2\}$ -matrix such that $m \mid s$. Assume the problems (1) and (2) have been solved with respective permutations Π^γ and Π^β and that $\mathcal{G}(A) = 1$. By Observation 1, $\gamma(A) \geq r$ and $\beta(A) \leq r$. If $\gamma(A) = r$ then $s_i^{\Pi^\gamma} = r$ for every $i \in \{1, \dots, m\}$ since $m \mid s$ contradicting the fact that $\mathcal{G}(A) = 1$. If $\gamma(A) = r + 1$ and $\beta(A) = r$, then $s_i^{\Pi^\beta} = r$ for every $i \in \{1, \dots, m\}$ since $m \mid s$ again contradicting the fact that $\mathcal{G}(A) = 1$. The alternative $\gamma(A) \geq r + 2$ is not possible since it implies $\beta(A) > r$. ■

It is tempting to conjecture that $\mathcal{G}(A) = 1$ if and only if $m \nmid s$ for two-ary matrices. However, this is not true as the following example shows:

$$A = \begin{bmatrix} 2 & 0 & 2 \\ 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 2 & 0 \end{bmatrix}.$$

The above result and proof can be repeated, mutatis mutandis, for matrices with elements from an enlarged ground set, e.g., $\{v_1, \dots, v_p\}$. The proof of the following lemma is thus left as an exercise.

Lemma 7. *Let A be an $m \times d$ integer matrix with elements from the set $\{v_1, \dots, v_p\}$. If $m \mid s$ then $\mathcal{G}(A) \neq 1$.*

The converse of the statement in Lemma 6 is not true as the following example shows for a $\{0, 1, 2, 3\}$ matrix:

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \\ 3 & 3 & 1 \\ 2 & 3 & 1 \\ 3 & 3 & 1 \end{bmatrix}$$

where $\mathcal{G}(A) = 2$ with $r = 35/6$.

5. Additional mixability properties of two-ary matrices

In this section we shall show that $\mathcal{G}(A) = \mathcal{K}(A)$ for any $\{0, 1, 2\}$ matrix. We begin with some simple results.

Lemma 8. *For any $A \in \{0, 1, 2\}^{m \times d}$, either $s_1^{\Pi^\mathcal{K}} = s_1^{\Pi^\gamma}$ or $s_m^{\Pi^\mathcal{K}} = s_m^{\Pi^\beta}$.*

Proof. Assume to the contrary that $s_1^{\Pi^\mathcal{K}} > s_1^{\Pi^\gamma}$ and $s_m^{\Pi^\mathcal{K}} < s_m^{\Pi^\beta}$. Say, $s_1^{\Pi^\mathcal{K}} = s_1^{\Pi^\gamma} + t = \gamma(A) + t$ for some $t \geq 1$ and $s_m^{\Pi^\mathcal{K}} = s_m^{\Pi^\beta} - u = \beta(A) - u = \gamma(A) - \mathcal{G}(A) - u$ for some $u \geq 1$. Then, $\mathcal{K}(A) = \gamma(A) + t - \gamma(A) + \mathcal{G}(A) + u$. However, this is only possible if $\mathcal{G}(A) = 0$ and $\mathcal{K}(A) = 2$ which is impossible due to Lemma 2. ■

Theorem 6. *For any $A \in \{0, 1, 2\}^{m \times d}$, $\mathcal{G}(A) = \mathcal{K}(A)$.*

Proof. Based on the results of Lemmas 1 and 2 and Theorem 4, in order to show that $\mathcal{G}(A) = \mathcal{K}(A)$, we need to exclude $\mathcal{G}(A) = 1$ and $\mathcal{K}(A) = 2$ case. With Theorem 4, we may assume without loss of generality that $s_1^{\Pi^\gamma} - s_m^{\Pi^\gamma} = s_1^{\Pi^\beta} - s_m^{\Pi^\beta} = 2$. Assume to the contrary that for some $A \in \{0, 1, 2\}^{m \times d}$, $\mathcal{K}(A) = 2$ but $\mathcal{G}(A) = 1$. In other words, the row sum vector in permutation Π^γ is of the form $(a, \dots, a, a - 1, \dots, a - 1, a - 2, \dots, a - 2)$ and the row sum vector in permutation Π^β is of the form $(a + 1, \dots, a + 1, a, \dots, a, a - 1, \dots, a - 1)$ where

$a = \gamma(A)$. Since any column of identical elements can be removed without any loss of generality by Lemma 5, we assume that neither A nor B has a column consisting of the same element.

Assume without loss of generality that among permutations with the largest row sum equal to a and row sum difference equal to 2, say Π^γ is a permutation where the number of row sums equal to $a - 1$ is maximum. Similarly, let Π^β be a permutation where the smallest row sum is equal to $a - 1$, the row sum difference is equal to 2 and the number of row sums equal to a is maximum. For simplicity, let A and B be the matrices resulting from permutations Π^γ and Π^β , respectively. Then, for any column k , $A_{ik} = 1$ and $A_{jk} = 0$ is not possible for any row i of A with sum equal to a and any row j with sum equal to $a - 2$, since their swap will yield a matrix with a higher number of row sums equal to $a - 1$ violating the choice of Π^γ . Similarly, for such a pair of rows i and j , $A_{ik} = 2$ and $A_{jk} = 1$ is not possible either for any k . So, for any row i with sum equal to a and any row j with sum equal to $a - 2$, there should be some column k such that $A_{ik} = 2$ and $A_{jk} = 0$ in order to lead to a row sum difference of 2. But then $A_{il} = 0$ and $A_{jl} = 1$ for $l \neq k$ is not possible either since by simultaneous swaps of i th and j th rows in columns k and l , one attains a permutation with higher number of row sum values equal to $a - 1$. By the same reasoning, $A_{il} = 1$ and $A_{jl} = 2$ for $l \neq k$ is not possible either. So, we can conclude that for any two rows i and j with row sum difference equal to 2, $|A_{ik} - A_{jk}| = 0$ or 2 for any column k . Moreover, if $A_{ik} = 1$ for some row i with row sum equal to either a or $a - 2$, then $A_{jk} = 1$ for every $j \neq i$ where $s_j(A) = a$ or $a - 2$. The same arguments can be repeated to get the same result for matrix B .

Let $k_1, k_2, m - k_1 - k_2$ be the number of rows with sums equal to $a, a - 1$, and $a - 2$, respectively, in A . Similarly, let $l_1, l_2, m - l_1 - l_2$ be the number of rows with sums equal to $a + 1, a$, and $a - 1$, respectively, in B . In order to have $s_1(A) = s_1(B) - 1$, one must have some column k such that $|A_{1k} - B_{1k}| = 1$. Assume without loss of generality that $B_{1k} = 1$. By the arguments in the preceding paragraph, this implies that $B_{ik} = 1$ for every i with $s_i(B) = a + 1$ or $a - 1$. Moreover, since $A_{1k} \neq 1$, we must have $A_{ik} \neq 1$ for every i with $s_i(A) = a$ or $a - 2$. Since $s(A) = s(B)$ we must have

$$k_1 a + k_2(a - 1) + (m - k_1 - k_2)(a - 2) = l_1(a + 1) + l_2 a + (m - l_1 - l_2)(a - 1),$$

or, equivalently

$$2k_1 + k_2 = 2l_1 + l_2 + m. \tag{4}$$

In matrix A , column k can only have 1's in rows with sums equal to $a - 1$, and thus

$$k_2 \geq l_1 + m - l_1 - l_2 = m - l_2. \tag{5}$$

Since $\mathcal{K}(A) = \mathcal{K}(B) = 2$, we must have the following relations

$$k_1 \geq 1, \quad l_1 \geq 1, \quad m - k_1 - k_2 \geq 1, \quad \text{and} \quad m - l_1 - l_2 \geq 1. \tag{6}$$

However, system (4)–(6) is not feasible. If the first row in matrix A has element 1, inequality (5) would be replaced by $l_1 \geq m - k_2$ and the inequality system will lead to a similar contradiction. ■

Note that $\mathcal{G}(A)$ is not always equal to $\mathcal{K}(A)$ for an integer matrix A . An example of a 3×3 integer matrix for which $\mathcal{G}(A) < \mathcal{K}(A)$ is given in [5].

5.1. A polynomial algorithm for computing $\mathcal{K}(A)$ for $\{0, 1, 2\}$ -Matrices

Through Corollary 3, given a matrix $\{0, 1, 2\}^{m \times d}$ we may assume that we have achieved a permutation with the largest row sum difference at most 2 in time $\mathcal{O}(d \cdot m \log m)$. Let A be the resulting matrix with a row sum pattern of the form

$$\underbrace{(a + 1, \dots, a + 1, a, \dots, a, a - 1, \dots, a - 1)}_{\substack{K \text{ times} \\ L \text{ times}}}$$

We define the following *swap* operations:

1. A *single improving swap* involves two rows i and j and a column p where $A_{ip} - A_{jp} = 1$, $s_i = s_j + 2$ and swaps elements A_{ip} with A_{jp} .
2. A *single neutral swap* involves two rows i and j and a column p where $A_{ip} - A_{jp} = 1$, $s_j \leq s_i \leq s_j + 1$ and swaps elements A_{ip} with A_{jp} .
3. A *double improving swap* involves rows i and j and columns p and q where $s_i = s_j + 2$, $(A_{ip} + A_{iq}) - (A_{jp} + A_{jq}) = 1$ and swaps A_{ip} with A_{jp} and A_{iq} with A_{jq} simultaneously.
4. A *double neutral swap* involves rows i and j and columns p and q where $s_j \leq s_i \leq s_j + 1$, $(A_{ip} + A_{iq}) - (A_{jp} + A_{jq}) = 1$ and swaps A_{ip} with A_{jp} and A_{iq} with A_{jq} simultaneously.
5. An *improving swap chain* involves $k+1$ rows, say i_1, i_2, \dots, i_{k+1} where $s_{i_2} = s_{i_3} = \dots = s_{i_k}$, $s_{i_1} = s_{i_2} + 1$, $s_{i_k} = s_{i_{k+1}} + 1$ and each consecutive row pair i_l and i_{l+1} for $l \in \{1, \dots, k\}$ corresponds to either a single or a double neutral swap and all swaps involve independent columns.

It is easy to see that each of the improving swap operations will reduce $K + L$ by two.

Consider the following two-ary matrix A :

$$A = \begin{bmatrix} 2 & 1 & 2 & 2 & 0 \\ \boxed{2} & \boxed{0} & 1 & 2 & 2 \\ \boxed{0} & \boxed{1} & \boxed{2} & 1 & 2 \\ 0 & 2 & \boxed{1} & \boxed{2} & \boxed{1} \\ 1 & 0 & 2 & \boxed{0} & \boxed{2} \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

The row sums are $(7, 7, 6, 6, 5, 5)$ with $a = 6$ and $K + L = 4$. The boxed entries are subject to an improving swap chain. In particular, rows 2 and 3, columns 1 and 2 correspond to a double neutral swap, rows 3 and 4 and column 3 correspond to a single neutral swap and rows 4 and 5 and columns 4 and 5 again correspond to a double neutral swap. Rows 2 through 5 jointly characterize an improving swap chain. The resulting matrix reduces $K + L$ to 2 and becomes:

$$A = \begin{bmatrix} \boxed{2} & 1 & 2 & 2 & 0 \\ 0 & 1 & 1 & 2 & 2 \\ 2 & 0 & 1 & 1 & 2 \\ 0 & 2 & 2 & 0 & 2 \\ 1 & 0 & 2 & 2 & 1 \\ \boxed{1} & 1 & 1 & 1 & 1 \end{bmatrix}$$

with row sums $(7, 6, 6, 6, 6, 5)$. Rows 1 and 6 and column 1 (similarly, columns 3 and 4) further define a single improving swap and the resulting matrix is declared as jointly mixable.

We shall now describe a procedure which implements all improving exchanges until no more is possible. As is apparent with the proof of [Theorem 6](#), a permutation achieving $\mathcal{K}(A)$ row sum difference is one where $K + L$ value is minimum. In particular, we shall construct a network with two designated source and destination nodes such that any source destination path in this network will correspond to an improving swap operation. Then, finding all possible source destination paths in this network and minimizing $K + L$ value will be accomplished by a polynomial time maximum flow algorithm in an appropriately constructed capacitated network, say $G = (N, \bar{A})$. Let s and t be the source and destination nodes in N . Now, construct the remaining nodes and arcs in G as follows:

1. For each row $i \in \{1, \dots, K\}$, column $p \in \{1, \dots, d\}$ add node (i, p) to N and arc $(s, (i, p))$ to \bar{A} with capacity equal to 1.

2. For each row $i \in \{1, \dots, K\}$, columns $p, q \in \{1, \dots, d\}$ such that $p < q$ add node (i, p, q) to N and arc $(s, (i, p, q))$ to \bar{A} with capacity equal to 1.
3. For each row $i \in \{m - L + 1, \dots, m\}$, column $p \in \{1, \dots, d\}$ add node (i, p) to N and arc $((i, p), t)$ to \bar{A} with capacity equal to 1.
4. For each row $i \in \{m - L + 1, \dots, m\}$, columns $p, q \in \{1, \dots, d\}$ such that $p < q$ add node (i, p, q) to N and arc $((i, p, q), t)$ to \bar{A} with capacity equal to 1.
5. For each row $i \in \{K + 1, \dots, m - L\}$, column $p \in \{1, \dots, d\}$ add two nodes (i, p, I) and (i, p, II) to N .
6. For each row $i \in \{K + 1, \dots, m - L\}$, columns $p, q \in \{1, \dots, d\}$ such that $p < q$ add two nodes (i, p, q, I) and (i, p, q, II) to N .
7. For each node (i, p, II) where $i \in \{K + 1, \dots, m - L\}$ for which there exists some $j \in \{1, \dots, K\}$ such that $A_{jp} - A_{ip} = 1$, add arc $((j, p), (i, p, II))$ to \bar{A} with capacity equal to 1 for every such $j \in \{1, \dots, K\}$.
8. For each node (i, p, q, II) where $i \in \{K + 1, \dots, m - L\}$ for which there exists some $j \in \{1, \dots, K\}$ such that $A_{jp} + A_{jq} - A_{ip} - A_{iq} = 1$, add arc $((j, p, q), (i, p, q, II))$ to \bar{A} with capacity equal to 1 for every such $j \in \{1, \dots, K\}$.
9. For each node (i, p, I) where $i \in \{K + 1, \dots, m - L\}$ for which there exists some $j \in \{m - L + 1, \dots, m\}$ such that $A_{ip} - A_{jp} = 1$, add arc $((i, p, I), (j, p))$ to \bar{A} with capacity equal to 1 for every such $j \in \{m - L + 1, \dots, m\}$.
10. For each node (i, p, q, I) where $i \in \{K + 1, \dots, m - L\}$ for which there exists some $j \in \{m - L + 1, \dots, m\}$ such that $A_{ip} + A_{iq} - A_{jp} - A_{jq} = 1$, add arc $((i, p, q, I), (j, p, q))$ to \bar{A} with capacity equal to 1 for every such $j \in \{m - L + 1, \dots, m\}$.
11. For each $i \in \{K + 1, \dots, m - L\}$, distinct columns p_1, q_1, p_2, q_2 , add arcs $((i, p_1, II), (i, p_2, I))$, $((i, p_1, II), (i, p_2, q_2, I))$, $((i, p_1, q_1, II), (i, p_2, I))$ and $((i, p_1, q_1, II), (i, p_2, q_2, I))$ to \bar{A} all with capacities equal to 1.
12. For distinct $i, j \in \{K + 1, \dots, m - L\}$, add arc $((i, p, I), (j, p, II))$ to \bar{A} with capacity equal to 1 if $A_{ip} - A_{jp} = 1$.
13. For distinct $i, j \in \{K + 1, \dots, m - L\}$, add arc $((i, p, q, I), (j, p, q, II))$ to \bar{A} with capacity equal to 1 if $A_{ip} + A_{iq} - A_{jp} - A_{jq} = 1$.

In this capacitated $s - t$ network G , the first copies of nodes (ending with “I”) mark potential beginning cell(s) and the second copies of nodes (ending with “II”) mark potential ending cell(s) for neutral swap operations within rows $\{K + 1, \dots, m - L\}$. Since there is no need for neutral swap operations within cells in rows $\{1, \dots, K\}$ or in rows $\{m - L + 1, \dots, m\}$, the nodes corresponding to these rows are not duplicated. It is not difficult to see that every $s - t$ directed path in G corresponds to an improving swap (single, double, or chain) and reduces $K + L$ by 2. In particular, in the previous example, the depicted improving swap chain corresponds to the directed path $s \rightarrow (2, 1, 2) \rightarrow (3, 1, 2, II) \rightarrow (3, 3, I) \rightarrow (4, 3, II) \rightarrow (4, 4, 5, I) \rightarrow (5, 4, 5) \rightarrow t$.

We can find all possible such paths by finding the maximum $s - t$ flow in G . In other words, we have the following result.

Theorem 7. *Given an $A \in \{0, 1, 2\}^{m \times d}$, a permutation providing $\mathcal{G}(A) = \mathcal{K}(A)$ can be accomplished in time polynomial in m and d .*

Proof. Finding the maximum flow in a network with unit capacities such as our constructed network $G = (N, \bar{A})$, can be accomplished in $O(\min\{|N|^{\frac{2}{3}}|\bar{A}|, |\bar{A}|^{\frac{2}{3}}\})$ time [13]. Since in our network $|N|$ is $O(d^2)$ and \bar{A} is $O(md^2)$ the result follows. ■

6. Linear relaxations

Interestingly, an approach to the problem via integer programming has not been studied before, to the best of the authors' knowledge, with the exception of [2] which uses an integer program in fixed dimensions for a polynomiality argument and a PTAS (cf. Theorem 4 and Corollary 2 in [2]). In our computational experience with the integer programming formulation of the problem we observed that the problem defies solution in reasonable computation times even for moderately sized matrices with $m \leq 1000$ and $d \leq 500$. More precisely, state-of-the-art integer programming solvers have difficulty closing the optimality gap in the majority of randomly generated instances. An observation reinforcing this challenging aspect of the problem is that the linear programming relaxation bound is quite weak as we show below. This feature obviously invites further research on the structure of the problem.

Consider the linear relaxation of problems (1) and (2) (we denote by $\pi_k(i, j)$ the (i, j) entry of permutation matrix π_k ; the linear relaxations result from relaxing the binary requirements on the variables $\pi_k(i, j)$):

$$\begin{aligned}
 & \min z \\
 & \text{st } z \geq \sum_{k=1}^d \sum_{j=1}^m \pi_k(i, j) A_{jk}, & \forall i = 1, \dots, m \\
 & \sum_{j=1}^m \pi_k(i, j) = 1, & \forall k = 1, \dots, d; \forall i = 1, \dots, m \\
 & \sum_{i=1}^m \pi_k(i, j) = 1, & \forall k = 1, \dots, d; \forall j = 1, \dots, m \\
 & \pi_k(i, j) \geq 0, & \forall k = 1, \dots, d; i, j = 1, \dots, m
 \end{aligned}$$

for problem (1), and

$$\begin{aligned}
 & \max z \\
 & \text{st } z \leq \sum_{k=1}^d \sum_{j=1}^m \pi_k(i, j) A_{jk}, & \forall i = 1, \dots, m \\
 & \sum_{j=1}^m \pi_k(i, j) = 1, & \forall k = 1, \dots, d; \forall i = 1, \dots, m \\
 & \sum_{i=1}^m \pi_k(i, j) = 1, & \forall k = 1, \dots, d; \forall j = 1, \dots, m \\
 & \pi_k(i, j) \geq 0, & \forall k = 1, \dots, d; i, j = 1, \dots, m
 \end{aligned}$$

for problem (2). Let z^γ and z^β denote the optimal values, respectively.

The following observations are easy to prove.

Lemma 9. For $A \in \mathbb{R}^{m \times d}$ we have $z^\gamma = z^\beta = s/m = z(1/m)$ where $z(1/m)$ is the optimal value of the linear programming problem defined over an integer polyhedron:

$$\begin{aligned}
 & \min \frac{1}{m} \sum_{i=1}^m \sum_{k=1}^d \sum_{j=1}^m \pi_k(i, j) A_{ik} \\
 & \text{s.t. } \sum_{j=1}^m \pi_k(i, j) = 1, \quad \forall k = 1, \dots, d, \quad i = 1, \dots, m
 \end{aligned}$$

$$\sum_{i=1}^m \pi_k(i, j) = 1, \quad \forall k = 1, \dots, d, j = 1, \dots, d$$

$$\pi_k(i, j) \geq 0, \quad \forall k, i, j.$$

Proof. Proof of equality between $z(1/m)$ and s/m is by direct calculation. For the rest, let us consider problem (2). Let Π be a permutation system. Clearly we have the following inequality

$$\min_i s_i^\Pi \leq \frac{1}{m} \sum_{i=1}^m s_i^\Pi = s/m,$$

as an upper bound (since the mean of the elements is at least as large as the smallest element in a vector). Therefore, the optimal value of the LP problem above is an upper bound on z^β . However, the matrix with all elements equal to $1/m$ is feasible for the LP relaxation, hence attains the upper bound. The proof for problem (1) is similar. ■

7. Concluding remarks

In this paper we studied the problem of permuting the columns of a matrix to achieve the maximum minimal row sum and the minimum maximal row sum, a problem that received renewed interest from quantitative finance where the two values are desired to be as close as possible. While previous work has concentrated on computational complexity and approximability of problems (1) and (2) and on identifying cases where the equality of the maximum and the minimum is assured (hence, the term jointly mixable) we approached the problem from a novel angle in that we focused on computing the gap (and/or bounding the gap) between the aforementioned maximum and minimum. We were able to quantify the gap for a subset of the integer matrices, and we showed equivalence of the problem of finding the gap to a related problem for which a polynomial time solution procedure was given for matrices with $\{0, 1, 2\}$ entries. We also gave a simple polynomial time approximation algorithm. We are led to believe through our computational experience that some of the results given in the paper remain true for a much larger class of integer (and even real) matrices. We hope these problems will be resolved in the future.

Acknowledgment

We are grateful to an anonymous referee who kindly pointed out some recent references on the problem studied in this paper, and connections to results that are continuous space counterparts of some of our results.

References

- [1] P. Embrechts, G. Puccetti, L. Rüschendorf, Model uncertainty and VaR aggregation, *J. Bank. Finance* 27 (2013) 2750–2764.
- [2] U.-U. Haus, Bounding stochastic dependence, joint mixability of matrices, and multidimensional bottleneck assignment problems, *Oper. Res. Lett.* 43 (2015) 74–79.
- [3] G. Puccetti, L. Rüschendorf, Bounds for joint portfolios of dependent risks, *Statist. Risk Model. Appl. Finance Insur.* 28 (2012) 107–132.
- [4] R. Wang, L. Peng, J. Yang, Bounds for the sum of dependent risks and worst value-at-risk with monotone marginal densities, *Finance Stoch.* 17 (2) (2013) 395–417.
- [5] C. Bernard, X. Jiang, R. Wang, Risk aggregation with dependence uncertainty, *Insurance Math. Econom.* 54 (2014) 93–108.
- [6] P. Embrechts, B. Wang, R. Wang, Aggregation-robustness and model uncertainty of regulatory risk measures, *Finance Stoch.* 19 (4) (2015) 763–790.
- [7] B. Wang, R. Wang, Joint mixability, *Math. Oper. Res.* (2015) forthcoming, SSRN: <http://ssrn.com/abstract=2557067>.
- [8] G. Puccetti, L. Rüschendorf, Computation of sharp bounds on the distribution of a function of dependent risks, *J. Comput. Appl. Math.* 236 (2012) 1833–1840.

- [9] L. Rüschemdorf, Solution of a statistical problem by rearrangement methods, *Metrika* 30 (1983) 55–61.
- [10] R. Wang, Current open questions in complete mixability, Technical report, 2014, <http://arxiv.org/abs/1411.6190>.
- [11] W.-L. Hsu, Approximation algorithms for the assembly line crew scheduling problem, *Math. Oper. Res.* 9 (1984) 376–383.
- [12] E.G. Coffman, M. Yannakakis, Permuting elements within columns of a matrix in order to minimize maximum row sum, *Math. Oper. Res.* 9 (3) (1984) 384–390.
- [13] R.K. Ahuja, T.L. Magnanti, J.B. Orlin, *Network Flows: Theory, Algorithms, and Applications*, Prentice-Hall, New Jersey, 1993.
- [14] P. Embrechts, E. Jakobsons, Dependence uncertainty for aggregate risk: Examples and simple bounds, in: M. Podolskij, et al. (Eds.), *The Fascination of Probability, Statistics and their Applications*, 2015, 395–417.