POSITIVE SEMIDEFINITE RANK

HAMZA FAWZI, JOÃO GOUVEIA, PABLO A. PARRILO, RICHARD Z. ROBINSON,
AND REKHA R. THOMAS

ABSTRACT. Let $M \in \mathbb{R}^{p \times q}$ be a nonnegative matrix. The positive semidefinite rank (psd rank) of M is the smallest integer k for which there exist positive semidefinite matrices A_i, B_j of size $k \times k$ such that $M_{ij} = \text{trace}(A_i B_j)$. The psd rank has many appealing geometric interpretations, including semidefinite representations of polyhedra and information-theoretic applications. In this paper we develop and survey the main mathematical properties of psd rank, including its geometry, relationships with other rank notions, and computational and algorithmic aspects.

1. INTRODUCTION

Matrix factorizations (or more generally, factorizations of linear maps) are a classical and important topic in applied mathematics. For instance, in the standard low-rank matrix factorization problem, given a matrix $M \in \mathbb{R}^{p \times q}$ one constructs matrices $A \in \mathbb{R}^{p \times k}$ and $B \in \mathbb{R}^{k \times q}$ such that $M = AB$, where the intermediate dimension k is as small as possible. Letting a_i and b_j be the rows of A and columns of B, respectively, finding such a factorization can be interpreted as a *realizability* problem, where we want to produce vectors in $\mathbb{R}^k$ that realize the inner products given by $M_{ij} = \langle a_i, b_j \rangle$. The smallest such k is, of course, the usual rank of the matrix M.

In applications, low-rank factorizations often have appealing interpretations (e.g., reduced-order or "simple" models), since they provide a decomposition of a linear map $\mathbb{R}^q \to \mathbb{R}^p$ in terms of mappings $\mathbb{R}^q \to \mathbb{R}^k \to \mathbb{R}^p$ through a "small" subspace. Many classical and successful methods in systems theory (e.g., realization theory, model order reduction), or statistics and machine learning (e.g., principal component analysis, factor analysis) are based on these techniques; e.g. [Kal63, Moo81, Jol02].

In many situations, however, one often requires additional conditions on the possible factors. A well-known example is the case of *nonnegative factorizations* [CR93], where M is a given nonnegative matrix and the factors A, B are also required to be nonnegative (here and throughout the paper, a *nonnegative* matrix is a matrix where all the entries are nonnegative). These requirements often arise from probabilistic interpretations (if M corresponds to a joint distribution, in which case the factors can be interpreted in terms of conditional independence; see e.g. [MSvS03]), or modeling choices (additive representations in terms of features and latent variables; see e.g., [LS99]). Another well-known case is when the factors A, B are required to be "small" with respect to a given matrix norm. This is a situation that has been well-studied in contexts such as Banach space theory, communication complexity and machine learning, where *factorization norms* are used to capture this notion [LMSS07, LS09].

Over the last couple of years, a new and intriguing class of matrix factorization problems has been introduced, by considering *conic factorizations* of nonnegative linear maps through a *convex cone* K, i.e., mappings $\mathbb{R}^q_+ \to K \to \mathbb{R}^p_+$ (nonnegative factorizations correspond to the case when K

is the nonnegative orthant). A particularly interesting case, which is the focus of this paper, occurs when K is the cone of positive semidefinite matrices [GPT13].

More concretely, a *positive semidefinite factorization* of a nonnegative matrix $M \in \mathbb{R}^{p \times q}$ is a collection of symmetric $k \times k$ positive semidefinite matrices $\{A_1, \ldots, A_p\}$ and $\{B_1, \ldots, B_q\}$ such that

$$M_{ij} = \text{trace}(A_i B_j), \qquad i = 1, \ldots, p, \quad j = 1, \ldots, q.$$

The *positive semidefinite rank* (*psd rank*) of M can then be defined as the smallest k for which such a factorization exists [GPT13]. As we explain in Section 3, a natural and important source of these factorizations is their relationship with representability of polytopes by semidefinite programming. These results extend the connections, first explored by Yannakakis in the context of polytopes and linear programming [Yan91], between "algebraic" factorizations of the slack matrix and the "geometric" question of existence of extended formulations. Since this quantity exactly characterizes semidefinite representability, positive semidefinite rank is an essential component of the burgeoning area of convex algebraic geometry [BPT12].

Besides these geometric and complexity-theoretic considerations, however, there are many other reasons to study these natural factorizations and ranks as independent mathematical objects, and this is the viewpoint we emphasize in this paper. Our main goal is to study the positive semidefinite rank of a matrix from the algebraic-geometric perspective, as well as survey and collect the existing results to date.

1.1. **Paper outline.** In Section 2 we present the formal definition and basic properties of psd rank. We analyze its behavior under natural matrix operations, its continuity properties, as well as its dependence on the underlying field. Throughout, we present numerous examples illustrating these notions.

Section 3 presents several complementary interpretations and motivations for psd rank. We discuss extensively its main geometric interpretation in terms of the "complexity" of a convex body that is contained in between two polytopes, a topic initially studied in [GPT13] and further elaborated in [FMP+12]. We also discuss a quantum analogue of the well-known probabilistic interpretations of nonnegative factorizations in terms of conditional independence, where the psd rank characterizes the minimum amount of quantum information that must be shared between two parties to generate samples of a correlated random variable [JSWZ13].

Like nonnegative rank, psd rank can be computationally challenging, although in some situations it can be nicely characterized. In Section 3 we show that the case of psd rank equal to 2 can be decided using convex optimization (in particular, semidefinite programming). The positive semidefinite rank of a matrix has natural relations with other rank notions, such as the usual (or "standard") rank and the nonnegative rank; we discuss these in Section 4. We also present the related notion of "square root rank," a refinement of psd rank to the case of rank-one factors. We show that in general, these rank notions are incomparable (Table 1), and provide explicit examples/counterexamples for all pairwise comparisons between them.

In Section 5 we analyze the situation where additional properties are imposed on the factor matrices A_i, B_j. We show how to guarantee uniform bounds on the factors for different norms (trace, spectral norm), as well as upper bounds on their ranks.

Since psd factorizations are not unique, it is also of interest to study the space of all possible factorizations. In Section 6 we study the topological properties of the space of factorizations, and show that in certain cases it is connected, in contrast to the case of nonnegative factorizations. This geometric insight also allows a better understanding on the rank of possible factors; see e.g., Example 6.8.

In Section 7 we specialize the situation to symmetric matrices with symmetric factorizations, and discuss the connections with some classical matrix cones (completely positive, doubly nonnegative).

We conclude in Section 8 with open problems, and questions for future research.

2. Definition and interpretations

For a positive integer k, let $\mathcal{S}_+^k$ denote the cone of $k \times k$ real symmetric positive semidefinite (psd) matrices. Recall that $\mathcal{S}_+^k$ is a closed convex cone in the vector space $\mathcal{S}^k$ of all $k \times k$ real symmetric matrices. We equip $\mathcal{S}^k$ with the standard inner product defined by:

$$\langle A, B \rangle = \text{trace}(AB) = \sum_{1 \leq i,j \leq k} A_{ij} B_{ij}.$$

The inner product of any two psd matrices is nonnegative, i.e., if $A, B \in \mathcal{S}_+^k$, then $\langle A, B \rangle \geq 0$. In fact the cone $\mathcal{S}_+^k$ is *self-dual*, meaning that:

$$A \in \mathcal{S}_+^k \iff \langle A, X \rangle \geq 0 \quad \forall X \in \mathcal{S}_+^k.$$

The following well-known fact about orthogonal matrices in $\mathcal{S}_+^k$ will be useful later:

Proposition 2.1. *If $A, B \in \mathcal{S}_+^k$ are such that $\langle A, B \rangle = 0$, then $AB = 0$.*

Proof. If $A = \sum_i a_i a_i^T$ and $B = \sum_j b_j b_j^T$ then $\langle A, B \rangle = \sum_{i,j} (a_i^T b_j)^2$. Thus since $\langle A, B \rangle = 0$ we get $a_i^T b_j = 0$ for all i, j. Hence this means that $AB = \sum_{i,j} (a_i^T b_j) a_i b_j^T = 0$. $\qquad\square$

2.1. **Psd rank.** We now give the formal definitions of *psd factorizations* and *psd rank*, which are the main objects of study in this paper:

Definition 2.2 ([GPT13])**.** Given a nonnegative matrix $M \in \mathbb{R}_+^{p \times q}$, a *psd factorization* of M of size k is a collection of psd matrices $A_1, \ldots, A_p \in \mathcal{S}_+^k$ and $B_1, \ldots, B_q \in \mathcal{S}_+^k$ such that $M_{ij} = \langle A_i, B_j \rangle$ for all $i = 1, \ldots, p$ and $j = 1, \ldots, q$. The *psd rank* of M, denoted $\text{rank}_{\text{psd}}(M)$, is the smallest integer k for which M admits a psd factorization of size k.

Remark 2.3. A psd factorization of M is equivalent to the existence of linear maps $\mathbb{R}_+^q \to \mathcal{S}_+^k \to \mathbb{R}_+^p$. Indeed, given a psd factorization, the linear maps $x \mapsto \sum_{j=1}^q x_j B_j$ and $Y \mapsto \langle A_i, Y \rangle$ (for $i = 1, \ldots, p$) have the desired property. The converse is also easy, by considering the image of the coordinate vectors $e_1, \ldots, e_q$ under the first map, and self-duality of the cone $\mathcal{S}_+^k$.

The psd rank is related to another notion of rank for nonnegative matrices, namely the *nonnegative rank* [CR93].

Definition 2.4. Given a nonnegative matrix $M \in \mathbb{R}_+^{p \times q}$, a *nonnegative factorization* of M of size k is a collection of nonnegative vectors $a_1, \ldots, a_p \in \mathbb{R}_+^k$ and $b_1, \ldots, b_q \in \mathbb{R}_+^k$ such that $M_{ij} = a_i^T b_j$ for all $i = 1, \ldots, p$ and $j = 1, \ldots, q$.
The *nonnegative rank* of M, denoted $\text{rank}_+(M)$, is the smallest integer k for which M admits a nonnegative factorization of size k.

The first proposition below establishes simple inequalities between the different notions of rank, namely the usual rank, the psd rank and the nonnegative rank.

Proposition 2.5. *If $M \in \mathbb{R}_+^{p \times q}$ is a nonnegative matrix, then*

$$(1) \qquad \frac{1}{2}\sqrt{1 + 8\,\text{rank}(M)} - \frac{1}{2} \leq \text{rank}_{\text{psd}}(M) \leq \text{rank}_+(M) \leq \min(p, q).$$

Proof. The last inequality is trivially true since $M = MI_q = I_p M$ where I_k is the $k \times k$ identity matrix.

Suppose $a_1, \ldots, a_p \in \mathbb{R}_+^k$ and $b_1, \ldots, b_q \in \mathbb{R}_+^k$ give a nonnegative factorization of $M \in \mathbb{R}_+^{p \times q}$. Then the diagonal matrices $A_i := \text{diag}(a_i)$ and $B_j := \text{diag}(b_j)$ give a $\mathcal{S}_+^k$-factorization of M, and we obtain the second inequality.

Now suppose $A_1, \ldots A_p, B_1, \ldots, B_q$ give a $\mathcal{S}_+^k$-factorization of M. Consider the vectors

$$a_i = \text{vec}(A_i) \quad \text{and} \quad b_j = \text{vec}(B_j)$$

where for $X \in \mathcal{S}^k$ we define $\mathrm{vec}(X) \in \mathbb{R}^{\binom{k+1}{2}}$ by:

$$\mathrm{vec}(X) = (X_{11}, \ldots, X_{kk}, \sqrt{2}X_{12}, \ldots, \sqrt{2}X_{1k}, \sqrt{2}X_{23}, \ldots, \sqrt{2}X_{(k-1)k}).$$

Then $\langle a_i, b_j \rangle = \langle A_i, B_j \rangle = M_{ij}$ so M has rank at most $\binom{k+1}{2}$. By solving for k we get the desired inequality. $\qquad \square$

Example 2.6. To illustrate the notion of a psd factorization, consider the following matrix M known as the 3×3 *derangement* matrix:

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

This matrix M satisfies $\mathrm{rank}\,(M) = \mathrm{rank}_+(M) = 3$. One can show that M admits a psd factorization of size 2. Indeed, define:

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad A_3 = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

$$B_1 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad B_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad B_3 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

One can easily check that the matrices A_i and B_j are positive semidefinite, and that $M_{ij} = \langle A_i, B_j \rangle$ for all $i = 1, \ldots, 3$ and $j = 1, \ldots, 3$. This factorization shows that $\mathrm{rank}_{\mathrm{psd}}\,(M) \leq 2$. In fact one has $\mathrm{rank}_{\mathrm{psd}}\,(M) = 2$ since the first inequality in (1) gives $\mathrm{rank}_{\mathrm{psd}}\,(M) \geq \frac{1}{2}\sqrt{1 + 8 \cdot 3} - \frac{1}{2} = 2$.

Example 2.7. Consider more generally the following 3×3 *circulant* matrix, where a, b, c are nonnegative real numbers:

$$(2) \qquad\qquad M(a, b, c) = \begin{bmatrix} a & b & c \\ c & a & b \\ b & c & a \end{bmatrix}.$$

One can check that the usual rank of $M(a, b, c)$ is 3 unless $a = b = c$ in which case the rank is one. When $\mathrm{rank}\,(M(a, b, c)) = 3$, the bounds in (1) say that $2 \leq \mathrm{rank}_{\mathrm{psd}}\,(M(a, b, c)) \leq 3$. Using the geometric interpretation of the psd rank (cf. Section 3) one can show that

$$\mathrm{rank}_{\mathrm{psd}}\,(M(a, b, c)) \leq 2 \quad \Longleftrightarrow \quad a^2 + b^2 + c^2 \leq 2(ab + bc + ac).$$

Figure 1 shows the region described by the inequality above when $a = 1$.

If M is a matrix such that $\mathrm{rank}\,(M) = 1$ or $\mathrm{rank}\,(M) = 2$, then the psd rank is equal to the rank, as stated in the next proposition:

Proposition 2.8. *For a nonnegative matrix $M \in \mathbb{R}_+^{p \times q}$ the following is true:*

$$(3) \qquad\qquad \mathrm{rank}\,(M) = 1 \Leftrightarrow \mathrm{rank}_+(M) = 1 \Leftrightarrow \mathrm{rank}_{\mathrm{psd}}\,(M) = 1.$$

Furthermore, we have the following implication:

$$(4) \qquad\qquad \mathrm{rank}\,(M) = 2 \Rightarrow \mathrm{rank}_+(M) = \mathrm{rank}_{\mathrm{psd}}\,(M) = 2.$$

Proof. The proof of (3) is clear from the inequalities (1). For (4), we can use a result from [CR93] which states that if $\mathrm{rank}\,(M) = 2$ then $\mathrm{rank}_+(M) = 2$, from which it easily follows that $\mathrm{rank}_{\mathrm{psd}}\,(M) = 2$. In Section 3 (Remark 3.7), we give a geometric argument for (4). $\qquad \square$
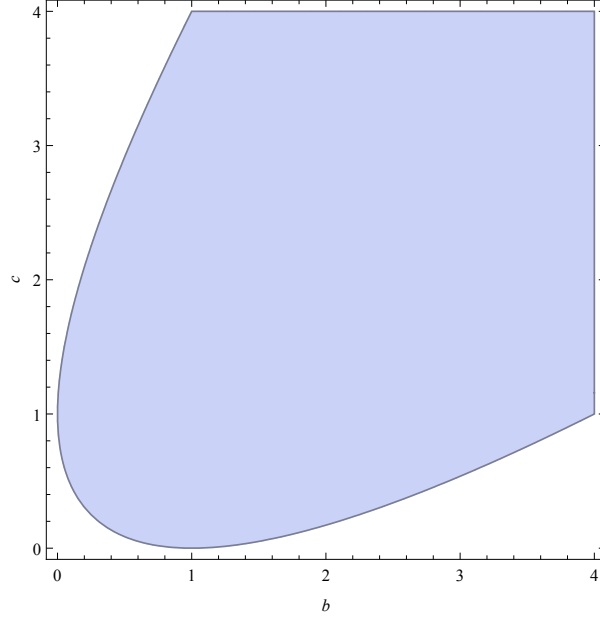
Figure 1. The blue region shows the values of (b, c) for which $\mathrm{rank}_{\mathrm{psd}}\left(M(1, b, c)\right) \leq 2$, where $M(a, b, c)$ is the 3×3 circulant matrix defined in (2). This region is defined by the inequality $1 + b^2 + c^2 \leq 2(b + c + bc)$.

2.2. **Dependence on the field.** In the definition of psd rank, Definition 2.2, we required the factors $(A_i)_{i=1,\ldots,p}$ and $(B_j)_{j=1,\ldots,q}$ to have real entries. When the matrix M has rational entries, it is natural to define a notion of psd rank where the factors $(A_i)_{i=1,\ldots,p}$ and $(B_j)_{j=1,\ldots,q}$ are required to have rational entries. If we denote this by $\mathrm{rank}^{\mathbb{Q}}_{\mathrm{psd}}(M)$, then we clearly have:

$$(5) \qquad \mathrm{rank}_{\mathrm{psd}}(M) \leq \mathrm{rank}^{\mathbb{Q}}_{\mathrm{psd}}(M).$$

In [GFR14] it was shown on a 8×6 matrix M that the inequality (5) can be strict.

It is also natural to consider a related notion of psd rank where the factors $(A_i)_{i=1,\ldots,p}$ and $(B_j)_{j=1,\ldots,q}$ in the psd factorizations are positive semidefinite *Hermitian* matrices. Denote by $\mathrm{rank}^{\mathbb{C}}_{\mathrm{psd}}(M)$ the associated psd rank. It is not difficult to see that the following inequalities hold:

$$\mathrm{rank}^{\mathbb{C}}_{\mathrm{psd}}(M) \leq \mathrm{rank}_{\mathrm{psd}}(M) \leq 2\,\mathrm{rank}^{\mathbb{C}}_{\mathrm{psd}}(M).$$

The second inequality comes from the fact that if A is a Hermitian positive semidefinite matrix, then the $2n \times 2n$ real symmetric matrix

$$(6) \qquad \frac{1}{\sqrt{2}} \begin{bmatrix} \mathrm{Re}\,A & \mathrm{Im}\,A \\ -\,\mathrm{Im}\,A & \mathrm{Re}\,A \end{bmatrix}$$

is positive semidefinite. Furthermore the function which maps any $n \times n$ Hermitian matrix A to the $2n \times 2n$ block matrix (6) preserves inner products.

One can show that the Hermitian psd rank can be strictly smaller than the real psd rank. Consider the 4×4 derangement matrix:

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

Using the inequalities (1) one can show that $\mathrm{rank}_{\mathrm{psd}}(M) \geq 3$. However one can find a psd factorization of M with Hermitian matrices of size 2, as given below:

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad A_3 = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad A_4 = \begin{bmatrix} 1 & e^{2i\pi/3} \\ e^{-2i\pi/3} & 1 \end{bmatrix}$$

$$B_1 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad B_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad B_3 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad B_4 = \begin{bmatrix} 1 & -e^{2i\pi/3} \\ -e^{-2i\pi/3} & 1 \end{bmatrix}.$$

Actually in [LWdW14], the authors exhibit a sequence of matrices (M^n) of increasing size such that $\mathrm{rank}_{\mathrm{psd}}^{\mathbb{C}}(M^n) < \mathrm{rank}_{\mathrm{psd}}(M^n)$ for all n and where the gap $\mathrm{rank}_{\mathrm{psd}}(M^n) - \mathrm{rank}_{\mathrm{psd}}^{\mathbb{C}}(M^n)$ grows with n (the ratio $\mathrm{rank}_{\mathrm{psd}}(M^n)/\mathrm{rank}_{\mathrm{psd}}^{\mathbb{C}}(M^n)$ goes asymptotically to $\sqrt{2}$). In this survey we will focus on the *real* psd rank, given in Definition 2.2.

2.3. First properties.

The next theorem establishes some structural properties satisfied by the psd rank

Theorem 2.9. *Given a nonnegative matrix $M \in \mathbb{R}_+^{p \times q}$, we have:*

(i) $\mathrm{rank}_{\mathrm{psd}}(M) = \mathrm{rank}_{\mathrm{psd}}(M^T)$.

(ii) *If $D_1 \in \mathbb{R}_+^{p \times p}, D_2 \in \mathbb{R}_+^{q \times q}$ are diagonal matrices with strictly positive elements on the diagonal, then $\mathrm{rank}_{\mathrm{psd}}(D_1 M D_2) = \mathrm{rank}_{\mathrm{psd}}(M)$.*

(iii) *If $N \in \mathbb{R}_+^{p \times q}$, then $\mathrm{rank}_{\mathrm{psd}}(M + N) \leq \mathrm{rank}_{\mathrm{psd}}(M) + \mathrm{rank}_{\mathrm{psd}}(N)$.*

(iv) *If $N \in \mathbb{R}_+^{q \times r}$ then $\mathrm{rank}_{\mathrm{psd}}(MN) \leq \min(\mathrm{rank}_{\mathrm{psd}}(M), \mathrm{rank}_{\mathrm{psd}}(N))$.*

(v) $\mathrm{rank}_{\mathrm{psd}}(M \odot M) \leq \mathrm{rank}(M)$, *where $\odot$ denotes Hadamard (entrywise) product.*

Proof.

(i) Property (i) is clear.

(ii) If $M_{ij} = \langle A_i, B_j \rangle$ is a psd factorization of M, then

$$(D_1 M D_2)_{ij} = \langle (D_1)_{ii} A_i, (D_2)_{jj} B_j \rangle$$

is a psd factorization of $D_1 M D_2$ of the same size. Thus since the diagonal elements of D_1 and D_2 are strictly positive we easily get that $\mathrm{rank}_{\mathrm{psd}}(M) = \mathrm{rank}_{\mathrm{psd}}(D_1 M D_2)$.

(iii) Let $M_{ij} = \langle A_i, B_j \rangle$ and $N_{ij} = \langle A_i', B_j' \rangle$ be psd factorizations of M and N of size respectively $\mathrm{rank}_{\mathrm{psd}}(M)$ and $\mathrm{rank}_{\mathrm{psd}}(N)$. Define

$$C_i = \begin{bmatrix} A_i & 0 \\ 0 & A_i' \end{bmatrix} \quad \text{and} \quad D_j = \begin{bmatrix} B_j & 0 \\ 0 & B_j' \end{bmatrix}.$$

Note that C_i and D_j are psd matrices of size $\mathrm{rank}_{\mathrm{psd}}(M) + \mathrm{rank}_{\mathrm{psd}}(N)$. Furthermore we clearly have $M_{ij} + N_{ij} = \langle C_i, D_j \rangle$. Thus $\mathrm{rank}_{\mathrm{psd}}(M + N) \leq \mathrm{rank}_{\mathrm{psd}}(M) + \mathrm{rank}_{\mathrm{psd}}(N)$.

(iv) Let $k = \mathrm{rank}_{\mathrm{psd}}(M)$ and let $M_{ij} = \langle A_i, B_j \rangle$ be a psd factorization of M of size k. For $j \in [r]$, define $C_j = \sum_{t=1}^{q} N_{tj} B_t$. Note that $C_j \in \mathcal{S}_+^k$ since $N_{tj} \geq 0$ and $B_t \in \mathcal{S}_+^k$. Then we verify that $\langle A_i, C_j \rangle = \sum_{t=1}^{q} N_{tj} \langle A_i, B_t \rangle = \sum_{t=1}^{q} N_{tj} M_{it} = (MN)_{ij}$ and so we get a psd factorization of MN of size k. This shows that $\mathrm{rank}_{\mathrm{psd}}(MN) \leq \mathrm{rank}_{\mathrm{psd}}(M)$. A similar argument shows that $\mathrm{rank}_{\mathrm{psd}}(MN) \leq \mathrm{rank}_{\mathrm{psd}}(N)$.

(v) Let $M_{ij} = \langle a_i, b_j \rangle$ be a factorization of M where $a_i, b_j \in \mathbb{R}^r$ where $r = \mathrm{rank}(M)$. Define $A_i = a_i a_i^T \in \mathcal{S}_+^r$ and $B_j = b_j b_j^T$ for $i = 1, \ldots, p$ and $j = 1, \ldots, q$. Then A_i, B_j give a psd factorization of $M \odot M$ of size r. Indeed:

$$\langle A_i, B_j \rangle = \langle a_i a_i^T, b_j b_j^T \rangle = (\langle a_i, b_j \rangle)^2 = M_{ij}^2 = (M \odot M)_{ij}.$$

Hence $\mathrm{rank}_{\mathrm{psd}}(M \odot M) \leq \mathrm{rank}(M)$.

$\square$

The next theorem analyzes the psd rank of block-triangular matrices (the result below was also found independently by Gábor Braun and Sebastian Pokutta as well as in [LWdW14]):

Theorem 2.10. *Let $P \in \mathbb{R}_+^{p_1 \times q_1}, Q \in \mathbb{R}_+^{p_2 \times q_1}, R \in \mathbb{R}_+^{p_2 \times q_2}$ be nonnegative matrices and let M be the block matrix of size $(p_1 + p_2) \times (q_1 + q_2)$:*

$$M = \begin{bmatrix} P & 0 \\ Q & R \end{bmatrix}.$$

Then

$$(7) \qquad \operatorname{rank}_{\mathrm{psd}}(M) \geq \operatorname{rank}_{\mathrm{psd}}(P) + \operatorname{rank}_{\mathrm{psd}}(R).$$

Furthermore, when $Q = 0$ we have equality.

Proof. We first show the inequality (7). Assume the matrix M has a psd factorization of size k where the $p_1 + p_2$ row factors are called $A_1, \ldots, A_{p_1}, \widehat{A_1}, \ldots, \widehat{A_{p_2}} \in \mathcal{S}_+^k$ and the $q_1 + q_2$ column factors are $B_1, \ldots, B_{q_1}, \widehat{B_1}, \ldots, \widehat{B_{q_2}} \in \mathcal{S}_+^k$. Since the upper-right block of M is zero, we have for all $(i, j) \in [p_1] \times [q_2]$, $\langle A_i, \widehat{B_j} \rangle = 0$. Hence, by Proposition 2.1, $A_i \widehat{B_j} = 0$. Thus if we let $F = \sum_{j=1}^{q_2} \operatorname{range}(\widehat{B_j}) \subseteq \mathbb{R}^k$, we have that $F \subseteq \ker(A_i)$ for all $i = 1, \ldots, p_1$. Since A_i is symmetric this is equivalent to $\operatorname{range}(A_i) \subseteq F^\perp$. Let U be an orthonormal matrix whose columns consist of an orthonormal basis for $F^\perp$ concatenated with an orthonormal basis of F. Since $\operatorname{range}(A_i) \subseteq F^\perp$ we know that A_i has the form:

$$(8) \qquad A_i = U \begin{bmatrix} A_i' & 0 \\ 0 & 0 \end{bmatrix} U^T$$

where A_i' is of size d where $d = \dim F^\perp$. Furthermore, since $\operatorname{range}(\widehat{B_j}) \subseteq F$ we have:

$$(9) \qquad \widehat{B_j} = U \begin{bmatrix} 0 & 0 \\ 0 & \widehat{B_j'} \end{bmatrix} U^T$$

where $\widehat{B_j'}$ is of size $k - d = \dim F$. Note that if we conjugate all the factors of the psd factorization of M by U (i.e., replace A_i by $U^T A_i U$, etc.) we get another valid psd factorization of M of the same size. Thus we can assume without loss of generality that $U = I$ and that A_i and $\widehat{B_j}$ are block-diagonal.

If we now let B_j' be the upper-left $d \times d$ block of B_j, then we have $P_{ij} = \langle A_i, B_j \rangle = \langle A_i', B_j' \rangle$, since A_i has the block-diagonal structure (8) (with $U = I$). Thus this shows that $\operatorname{rank}_{\mathrm{psd}}(P) \leq d$. Similarly, if we let $\widehat{A_i'}$ be the bottom-right $(k-d) \times (k-d)$ block of $\widehat{A_i}$, then we get $R_{ij} = \langle \widehat{A_i}, \widehat{B_j} \rangle = \langle \widehat{A_i'}, \widehat{B_j'} \rangle$ and thus $\operatorname{rank}_{\mathrm{psd}}(R) \leq k - d$. Thus we finally get that $\operatorname{rank}_{\mathrm{psd}}(P) + \operatorname{rank}_{\mathrm{psd}}(R) \leq d + (k - d) = k$ which what we want.

We now show that when $Q = 0$ we have $\operatorname{rank}_{\mathrm{psd}}(M) = \operatorname{rank}_{\mathrm{psd}}(P) + \operatorname{rank}_{\mathrm{psd}}(R)$. Indeed let $P_{ij} = \langle C_i, D_j \rangle$ and $R_{ij} = \langle E_i, F_j \rangle$ be psd factorizations of P and R respectively of size $\operatorname{rank}_{\mathrm{psd}}(P)$ and $\operatorname{rank}_{\mathrm{psd}}(R)$. Define:

$$A_i = \begin{bmatrix} C_i & 0 \\ 0 & 0 \end{bmatrix} \quad \forall i \in [p_1], \quad \widehat{A_i} = \begin{bmatrix} 0 & 0 \\ 0 & E_i \end{bmatrix} \quad \forall i \in [p_2]$$

and

$$B_j = \begin{bmatrix} D_j & 0 \\ 0 & 0 \end{bmatrix} \quad \forall j \in [q_1], \quad \widehat{B_j} = \begin{bmatrix} 0 & 0 \\ 0 & F_j \end{bmatrix} \quad \forall j \in [q_2].$$

It is easy to see that the factors $A_1, \ldots, A_{p_1}, \widehat{A_1}, \ldots, \widehat{A_{p_2}}$ and $B_1, \ldots, B_{q_1}, \widehat{B_1}, \ldots, \widehat{B_{q_2}}$ give a psd factorization of the block-diagonal matrix $\begin{bmatrix} P & 0 \\ 0 & R \end{bmatrix}$ of size $\operatorname{rank}_{\mathrm{psd}}(P) + \operatorname{rank}_{\mathrm{psd}}(R)$. Thus this shows, together with the inequality proved above, that

$$\operatorname{rank}_{\mathrm{psd}} \begin{bmatrix} P & 0 \\ 0 & R \end{bmatrix} = \operatorname{rank}_{\mathrm{psd}}(P) + \operatorname{rank}_{\mathrm{psd}}(R).$$

$\square$

Example 2.11. A consequence of Theorem 2.10 is that the psd rank of the identity matrix I_n is equal to n, since $I_n = \mathrm{diag}(1, \ldots, 1)$. In fact more generally the psd rank of a nonnegative diagonal matrix is equal to the number of nonzero diagonal elements.

Kronecker product. The Kronecker (tensor) product of two matrices $M \in \mathbb{R}^{p \times q}$ and $N \in \mathbb{R}^{p' \times q'}$ is the $pp' \times qq'$ matrix $M \otimes N$ defined by:

$$M \otimes N = \begin{bmatrix} M_{11}N & \ldots & M_{1q}N \\ \vdots & & \vdots \\ M_{p1}N & \ldots & M_{pq}N \end{bmatrix}.$$

It is well-known that the rank of the Kronecker product $M \otimes N$ is equal to the product of the ranks of M and N: $\mathrm{rank}\,(M \otimes N) = \mathrm{rank}\,(M)\mathrm{rank}\,(N)$. A natural question is to know whether the same is true for the psd rank. In [LWdW14] the authors give a counterexample to this, where they show that the psd rank of $M \otimes N$ can be strictly smaller than $\mathrm{rank}_{\mathrm{psd}}\,(M)\mathrm{rank}_{\mathrm{psd}}\,(N)$ (note that the inequality $\mathrm{rank}_{\mathrm{psd}}\,(M \otimes N) \leq \mathrm{rank}_{\mathrm{psd}}\,(M)\mathrm{rank}_{\mathrm{psd}}\,(N)$ is always true).

2.4. **Lower semicontinuity of psd rank.** In this subsection we show that for any $k \in \mathbb{N}$, the set of matrices of psd rank less or equal to k is closed (under the standard topology in $\mathbb{R}^{p \times q}$). We prove the following:

Theorem 2.12. *Let $(M^n)_{n \in \mathbb{N}}$ be a sequence of nonnegative matrices converging to $M \in \mathbb{R}_+^{p \times q}$ such that $\mathrm{rank}_{\mathrm{psd}}\,(M^n) \leq k$ for all $n \in \mathbb{N}$. Then $\mathrm{rank}_{\mathrm{psd}}\,(M) \leq k$.*

Proof. The main ingredient to prove this result is to show that the factors A_i, B_j in a psd factorization can always be chosen to be bounded. We have the following lemma:

Lemma 2.13. *Let $M \in \mathbb{R}_+^{p \times q}$ and assume that M has a psd factorization of size k. Then M admits a psd factorization $M_{ij} = \langle A_i, B_j \rangle$ of size k where the factors satisfy $\mathrm{trace}(A_i) \leq k$ and $\mathrm{trace}(B_j) = \sum_{i=1}^{p} M_{ij}$.*

Proof. We defer the proof of this Lemma to Section 5 where we discuss in more detail the issue of scaling the factors in a psd factorization. $\square$

Let (M^n) be a sequence of nonnegative matrices converging to M. Since (M^n) is a convergent sequence the entries of M^n are all bounded from above by some positive constant (independent of n). The previous lemma shows that for each n, one can find a psd factorization of M^n of the form:

$$(M^n)_{ij} = \langle A_i^n, B_j^n \rangle$$

where $A_i^n, B_j^n \in \mathcal{S}_+^k$ and such that the sequences $(A_i^n)_{n \in \mathbb{N}}$ and $(B_j^n)_{n \in \mathbb{N}}$ are all bounded in $\mathcal{S}^k$. Thus one can extract convergent subsequences $(A_i^{\phi(n)}), (B_j^{\phi(n)})$ where $\phi : \mathbb{N} \to \mathbb{N}$ is increasing and $A_i^{\phi(n)} \to A_i$ and $B_j^{\phi(n)} \to B_j$ when $n \to +\infty$. Since $\mathcal{S}_+^k$ is closed we have $A_i, B_j \in \mathcal{S}_+^k$ and we get:

$$M_{ij} = \lim_{n \to +\infty} \langle A_i^{\phi(n)}, B_j^{\phi(n)} \rangle = \langle A_i, B_j \rangle$$

which is a valid psd factorization of M of size k. Thus $\mathrm{rank}_{\mathrm{psd}}\,(M) \leq k$. $\square$

Remark 2.14. The result above shows that the function $\mathrm{rank}_{\mathrm{psd}} : \mathbb{R}_+^{p \times q} \to \mathbb{N}$ is *lower semicontinuous*, i.e., for any convergent sequence $M^n \to M$ it holds:

$$\mathrm{rank}_{\mathrm{psd}}\,(M) \leq \liminf_{n \to +\infty} \mathrm{rank}_{\mathrm{psd}}\,(M^n).$$

It is well-known that the usual rank is also lower semicontinuous, as well as the nonnegative rank (cf. [BCR11] for the lower semicontinuity property of the nonnegative rank). However, some notions of

rank can fail to have this property. A well-known example is the rank of tensors of order greater or equal to 3 which is not lower semicontinuous, giving rise to the notion of border rank.

3. MOTIVATION AND EXAMPLES

3.1. **Geometric interpretation.** In this section we discuss the geometric interpretation of the psd rank. This interpretation was the original motivation that led to the definition of the psd rank in [GPT13].

Semidefinite programming is the problem of optimizing a linear function over an affine slice of the psd cone:

$$\text{minimize } L(X) \text{ subject to } X \in \mathcal{S}_+^k \cap \mathcal{L}$$

where $L : \mathcal{S}^k \to \mathbb{R}$ is a linear function and $\mathcal{L}$ is an affine subspace of $\mathcal{S}^k$. The feasible set $\mathcal{S}_+^k \cap \mathcal{L}$ of a semidefinite program is known as a *spectrahedron* and can also be written as the solution set of a linear matrix inequality $\{x \in \mathbb{R}^d : A_0 + x_1 A_1 + \cdots + x_d A_d \succeq 0\}$ where the A_i are symmetric matrices that span the subspace $\mathcal{L}$. Semidefinite programs can be solved to arbitrary precision in polynomial-time, and have many applications in different areas of science and engineering [VB96].

Let $P \subset \mathbb{R}^n$ be a polytope and assume we want to minimize a linear function ℓ over P, i.e., we want to compute $\min\{\ell(x) : x \in P\}$. Observe that if P admits a representation of the form

$$(10) \qquad\qquad P = \pi(\mathcal{S}_+^k \cap \mathcal{L})$$

where $\mathcal{L} \subset \mathcal{S}^k$ is an affine subspace and π is a linear map, then one can write the linear optimization problem over P as a semidefinite program of size k, since:

$$\min_{x \in P} \ell(x) = \min_{y \in \mathcal{S}_+^k \cap \mathcal{L}} (\ell \circ \pi)(y)$$

and $\ell \circ \pi$ is linear. A representation of the polytope P of the form (10) is called a *psd lift* of size k. Such a representation is interesting in practice when the size d of the lift is much smaller than the number of facets of P, which is the size of the trivial representation of P using linear inequalities. A natural question to ask is thus: what is the smallest k such that P admits a psd lift of size k?

It turns out that the answer to this question is tightly related to the psd rank considered in this paper. For this we need to introduce the notion of a slack matrix:

Definition 3.1. Let $P \subset \mathbb{R}^n$ be a polytope (i.e., a bounded polyhedron) and $Q \subset \mathbb{R}^n$ be a polyhedron with $P \subseteq Q$. Let $x_1, \ldots, x_v$ be such that $P = \text{conv}(x_1, \ldots, x_v)$ and let $a_j \in \mathbb{R}^n, b_j \in \mathbb{R}$, $(j = 1, \ldots, f)$ be such that $Q = \{x \in \mathbb{R}^n : a_j^T x \leq b_j \ \forall j\}$. Then the slack matrix of the pair P, Q, denoted $S_{P,Q}$ is the nonnegative $v \times f$ matrix whose (i, j)-th entry is $b_j - a_j^T x_i$. When $P = Q$ we write $S_P := S_{P,P}$ and we call it the slack matrix of P.

Remark 3.2. Note that the entries of slack matrix of $S_{P,Q}$ can depend on the inequality description of Q and the vertex description of P (e.g., different scalings, redundant inequalities), however it is not hard to see that these do not affect the various ranks of the matrix, namely the usual rank, nonnegative rank and psd rank. Thus we will often refer to a slack matrix of a pair P, Q as "the" slack matrix of P, Q.

The next theorem gives an answer to the question of psd lifts formulated above, using the psd rank: it shows that the size of the smallest psd lift of a polytope P is equal to the psd rank of the slack matrix S_P of P (this is the case $P = Q$ of the statement below).

Theorem 3.3 (cf. Proposition 3.6 in [GRT13b]). *Let $P \subset \mathbb{R}^n$ be a polytope with nonempty interior and $Q \subset \mathbb{R}^n$ be a polytope such that $P \subseteq Q$, and let $S_{P,Q}$ be the slack matrix of the pair P, Q (cf. Definition 3.1). Then $\text{rank}_{\text{psd}} S_{P,Q}$ is the smallest integer k for which there exists an affine subspace $\mathcal{L}$ of $\mathcal{S}^k$ and a linear map π such that $P \subseteq \pi(\mathcal{S}_+^k \cap \mathcal{L}) \subseteq Q$.*

Sketch. Let $k = \text{rank}_{\text{psd}}\, S_{P,Q}$. We first show how to construct a spectrahedron $\mathcal{S}_+^k \cap \mathcal{L}$ of size k such that $P \subseteq \pi(\mathcal{S}_+^k \cap \mathcal{L}) \subseteq Q$ for some linear map π. Let $x_1, \ldots, x_v$ be the vertices of P and let $Q = \{x \in \mathbb{R}^n : a_j^T x \leq b_j \,\forall j = 1, \ldots, f\}$ be a facet description of Q. Let $A_1, \ldots, A_v, B_1, \ldots, B_f \in \mathcal{S}_+^k$ be a psd factorization of $S_{P,Q}$ of size k:

$$b_j - a_j^T x_i = \langle A_i, B_j \rangle \quad \forall i = 1, \ldots, v, \; j = 1, \ldots, f.$$

Consider the convex set C:

(11) $$C = \{x \in \mathbb{R}^n : \exists A \in \mathcal{S}_+^k \text{ such that } b_j - a_j^T x = \langle A, B_j \rangle \,\forall j = 1, \ldots, f\}.$$

It is easy to verify that C is contained between P and Q: indeed $C \subseteq Q$ because any $x \in C$ satisfies $b_j - a_j^T x \geq 0$ for all $j = 1, \ldots, f$; also $P \subseteq C$ because the vertices x_i of P satisfy (11) with $A = A_i$. Also it is not too difficult to show that, since Q is bounded, C can be expressed in the desired form $C = \pi(\mathcal{S}_+^k \cap \mathcal{L})$ where $\mathcal{L}$ is an affine subspace of $\mathcal{S}^k$ and π is a linear projection map (we refer to [GRT13b, Proposition 3.6] for the details). Thus this proves the first direction.

Assume now that we can write $P \subseteq \pi(\mathcal{S}_+^k \cap \mathcal{L}) \subseteq Q$ where $\mathcal{L}$ is an affine subspace of $\mathcal{S}^k$ and π is a linear map. We show how to construct a psd factorization of $S_{P,Q}$ of size k. Note that we can assume $\mathcal{S}_+^k \cap \mathcal{L}$ to contain a positive definite matrix, since otherwise we can reduce k. Let $C = \pi(\mathcal{S}_+^k \cap \mathcal{L})$. Using a suitable choice of basis for $\mathcal{L}$, the convex set C can be written as:

$$C = \{x \in \mathbb{R}^n \; : \; \exists y \in \mathbb{R}^m \text{ such that } T(x, y) \in \mathcal{S}_+^k\}$$

where $T : \mathbb{R}^n \times \mathbb{R}^m \to \mathcal{S}^k$ is an affine linear map, i.e., T has the form $T(x, y) = U_0 + x_1 U_1 + \cdots + x_n U_n + y_1 V_1 + \cdots + y_m V_m$ for some $U_0, U_1, \ldots, U_n, V_1, \ldots, V_m \in \mathcal{S}^k$. Observe that since $C \subseteq Q$ we have for any $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$:

$$T(x, y) \in \mathcal{S}_+^k \;\Rightarrow\; b_j - a_j^T x \geq 0 \,\forall j = 1, \ldots, f.$$

Using strong duality for semidefinite programming (we assumed that $\mathcal{S}_+^k \cap \mathcal{L}$ contains a positive definite matrix), it follows that for any $j = 1, \ldots, f$, there exists $B_j \in \mathcal{S}_+^k$ and $c_j \geq 0$ such that:

(12) $$b_j - a_j^T x = c_j + \langle T(x, y), B_j \rangle \quad \forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^m.$$

Since C has nonempty interior and is bounded (because P has nonempty interior and Q is bounded) one can show that c_j can be taken to be 0 in (12) (see [Pas12, Lemma 6.1] for a proof in the case of linear programming lifts; the proof for semidefinite programming lifts is very similar). Since $P \subseteq C$ we know that for any x_i vertex of P there exists y_i such that $T(x_i, y_i) \in \mathcal{S}_+^k$. Thus if we let $A_i = T(x_i, y_i)$ then we get the following psd factorization of size k of the slack matrix $S_{P,Q}$:

$$b_j - a_j^T x_i = \langle A_i, B_j \rangle \quad \forall i = 1, \ldots, v, j = 1, \ldots, f.$$

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Note that the proof of Theorem 3.3 is constructive: it shows how to construct the spectrahedron $\mathcal{S}_+^k \cap \mathcal{L}$ and the linear map π from a psd factorization of $S_{P,Q}$, and vice-versa. We now illustrate Theorem 3.3 using two simple examples.

Example 3.4. Let $P = Q = [-1, 1]^2$ be the square in the plane. The polytope P has 4 facets and 4 vertices and the slack matrix of P can be shown to be equal to the following 4×4 matrix:

(13) $$M = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}.$$

One can construct the following psd factorization of M of size 3, showing that $\text{rank}_{\text{psd}} M \leq 3$: $M_{ij} = \langle A_i, B_j \rangle$ where $A_i = u_i u_i^T \in \mathcal{S}_+^3$, $B_j = v_j v_j^T \in \mathcal{S}_+^3$, $i, j = 1, \ldots, 4$ with:

$$u_1 = (1,0,0), \quad u_2 = (0,1,0), \quad u_3 = (0,0,1), \quad u_4 = (1,1,1)$$
$$v_1 = (1,0,0), \quad v_2 = (1,-1,0), \quad v_3 = (0,1,-1), \quad v_4 = (0,0,1).$$

Thus by Theorem 3.3, this means that one can represent the polytope $P = [-1,1]^2$ as the linear image of a spectrahedron of size 3. One can in fact show that P is the projection onto the (x,y) coordinates of the following spectrahedron T of size 3:

$$T = \left\{ (x,y,z) \in \mathbb{R}^3 \ : \ \begin{bmatrix} 1 & x & y \\ x & 1 & z \\ y & z & 1 \end{bmatrix} \succeq 0 \right\}.$$

The spectrahedron T (also known as the *elliptope*) is depicted in Figure 2. Note that no smaller representation of the square $[-1,1]^2$ is possible: it was proved in [GRT13a] that the psd rank of any n-dimensional polytope is at least $n+1$ which in this case means that $\text{rank}_{\text{psd}} M \geq 2 + 1 = 3$.
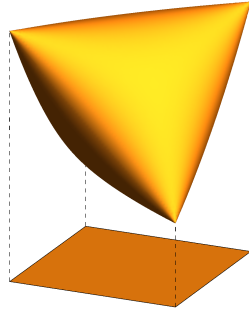

Figure 2. A psd lift of the square $[-1,1]^2$ of size 3: the elliptope $\{X \in \mathcal{S}_+^3 \ : \ \text{diag}(X) = \mathbb{1}\}$ linearly projects onto the square $[-1,1]^2$.

Example 3.5. We now give another illustration of Theorem 3.3 where the polytopes P and Q are different. Let $Q = [-1,1]^2$ and let now $P = [-a,a] \times [-b,b]$ be the rectangle centered at the origin with side lengths $2a$ and $2b$ with $0 \leq a,b \leq 1$ (cf. Figure 3). The slack matrix of the pair P, Q can be easily computed and is given by:

$$(14) \qquad\qquad M = \begin{bmatrix} 1+a & 1+b & 1-a & 1-b \\ 1-a & 1+b & 1+a & 1-b \\ 1-a & 1-b & 1+a & 1+b \\ 1+a & 1-b & 1-a & 1+b \end{bmatrix}.$$

Theorem 3.3 says that $\text{rank}_{\text{psd}} M$ is equal to the smallest size of a spectrahedron which has a linear projection that is contained between P and Q. In the previous example we saw a spectrahedron of size 3 which projects onto $Q = [-1,1]^2$ and thus this shows that $\text{rank}_{\text{psd}} M \leq 3$ for all $a,b \leq 1$. It is natural to ask whether the psd rank of M can be equal to 2 for some values of a, b? One can actually show that the psd rank of M is equal to 2 if, and only if, there is an ellipse E such that $P \subseteq E \subseteq Q$ (cf. e.g., [GRT13b, Section 4]). It is not hard to see that such an ellipse exists if and only if $a^2 + b^2 \leq 1$. Thus we have the following:

$$\text{rank}_{\text{psd}} M = \begin{cases} 3 & \text{if } a^2 + b^2 > 1 \\ 2 & \text{if } 0 < a^2 + b^2 \leq 1 \\ 1 & \text{if } a = b = 0 \end{cases}.$$
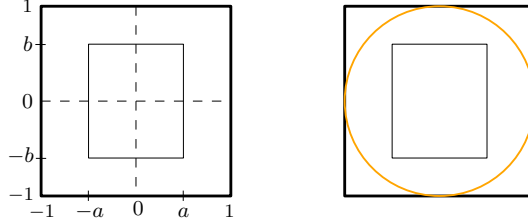
Figure 3. The geometric problem associated to the slack matrix M of Equation (14). The inner polytope is $P = [-a, a] \times [-b, b]$ and the outer polytope is $Q = [-1, 1]^2$. The right figure shows an instance where there exists an ellipse E such that $P \subset E \subset Q$.

The geometric interpretation of the psd rank given in Theorem 3.3 can be used to study the psd rank of any *arbitrary* nonnegative matrix M, since one can show that any nonnegative matrix M is the slack matrix of some pair of polytopes P, Q. In fact one can show the following result which appears in [GG12, Theorem 1]:

Lemma 3.6. *Let $M \in \mathbb{R}_+^{p \times q}$ be a nonnegative matrix and assume that $M\mathbb{1} = \mathbb{1}$. Let $r = \operatorname{rank} M$. Then there exist polytopes P, Q in $\mathbb{R}^{r-1}$ (where P and Q are bounded) such that $P \subset Q$ and such that M is the slack matrix of the pair P, Q.*

Proof. The proof is in [GG12] and we reproduce it here for completeness. In [GG12] it is shown that one can always find a factorization of M of the form $M = AB$ where $A \in \mathbb{R}^{p \times r}, B \in \mathbb{R}^{r \times q}$ and $A\mathbb{1} = \mathbb{1}$ and $B\mathbb{1} = \mathbb{1}$. Write A and B as:

$$A = \begin{bmatrix} a_1^T & t_1 \\ \vdots & \\ a_p^T & t_p \end{bmatrix} \quad B = \begin{bmatrix} b_1^T \\ \vdots \\ b_r^T \end{bmatrix},$$

where $a_i \in \mathbb{R}^{r-1}$, $t_i \in \mathbb{R}$ for $i = 1, \ldots, p$ and $b_j \in \mathbb{R}^q$ for all $j = 1, \ldots, r$. Note that since $A\mathbb{1} = \mathbb{1}$ we have $t_i = 1 - \mathbb{1}^T a_i$. Define the polytopes P and Q as follows:

$$P = \operatorname{conv}(a_1, \ldots, a_p) \subset \mathbb{R}^{r-1}$$

and

$$Q = \left\{ x \in \mathbb{R}^{r-1} \ : \ \sum_{i=1}^{r-1} x_i b_i + \left(1 - \sum_{i=1}^{r-1} x_i\right) b_r \geq 0 \right\}.$$

Note that Q is defined using q linear inequalities. It is not difficult to verify that M is the slack matrix of the pair P, Q. It remains to show that Q is bounded. Assume for contradiction that $x_0 + \alpha z \in Q$ for all $\alpha \geq 0$ where $x_0 \in Q$. Then one can show that this implies that

$$w := \sum_{i=1}^{r-1} z_i b_i - \left(\sum_{i=1}^{r-1} z_i\right) b_r \geq 0.$$

Note that we have $\mathbb{1}^T w = 0$ since $\mathbb{1}^T b_i = 1$ for all $i = 1, \ldots, r$. Thus since $w \geq 0$ and $\mathbb{1}^T w = 0$, this means that $w = 0$, i.e., $\sum_{i=1}^{r-1} z_i b_i - \left(\sum_{i=1}^{r-1} z_i\right) b_r = 0$. Since B is full row rank this necessarily means that $z = 0$. We have thus shown that Q is bounded. $\qquad \square$

Remark 3.7. As an application of Lemma 3.6, note for example that if M is a nonnegative matrix with rank two, then it is the slack matrix of two nested intervals in the real line. It thus follows easily from this geometric interpretation and from Theorem 3.3 that the psd rank of any rank-two matrix is equal to 2 (this was already shown in Proposition 2.8 using a result from [CR93]).

Rank-three matrices of psd rank two. Using the geometric interpretation of psd rank one can decide if a given matrix M of rank 3 has $\operatorname{rank}_{\mathrm{psd}} M \leq 2$ using semidefinite programming. Indeed assume that $\operatorname{rank} M = 3$ and let $P \subset Q \subset \mathbb{R}^2$ be two polytopes in the plane such that M is the slack matrix of P with respect to Q. From [GRT13b, Proposition 4.1], we know that $\operatorname{rank}_{\mathrm{psd}} M = 2$ if, and only if, there exists an ellipse E such that $P \subset E \subset Q$. Using a vertex description of P, and a facet description of Q, this can be decided using semidefinite programming: Indeed, let $x_1, \ldots, x_v$ be the vertices of P, and let $Q = \{x \in \mathbb{R}^2 \ : \ Gx \leq h\}$ be a facet description of Q where G has f rows. One can show that there exists an ellipse sandwiched between P and Q if, and only if, there exist $A \in \mathcal{S}^2, b \in \mathbb{R}^2$ and $c \in \mathbb{R}$ such that:

$$(i) \quad A \succeq 0, \ \operatorname{trace}(A) = 1;$$

$$(ii) \quad \begin{bmatrix} x_j \\ 1 \end{bmatrix}^T \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \begin{bmatrix} x_j \\ 1 \end{bmatrix} \leq 0 \quad \forall j = 1, \ldots, v;$$

$$(iii) \quad \exists \lambda_i \geq 0 \ : \ \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \succeq \lambda_i \begin{bmatrix} 0 & g_i^T/2 \\ g_i/2 & -h_i \end{bmatrix} \quad \forall i = 1, \ldots, f.$$

The ellipse E that satisfies $P \subset E \subset Q$ is then defined by:

$$E = \left\{ x \in \mathbb{R}^2 \ : \ \begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} \leq 0 \right\}.$$

Note that the constraint (ii) above corresponds to the condition $P \subseteq E$ and the constraint (iii) corresponds to $E \subseteq Q$. The latter uses the following result commonly known as the *S-lemma* [BV04, Appendix B]:

Lemma 3.8. *Let $A_i \in \mathcal{S}^n, b_i \in \mathbb{R}^n, c_i \in \mathbb{R}$ for $i = 1, 2$ and assume that the following implication holds for all $x \in \mathbb{R}^n$:*

$$x^T A_1 x + 2b_1^T x + c_1 \leq 0 \implies x^T A_2 x + 2b_2^T x + c_2 \leq 0.$$

Then there exists a $\lambda \geq 0$ such that:

$$\begin{bmatrix} A_2 & b_2 \\ b_2^T & c_2 \end{bmatrix} \preceq \lambda \begin{bmatrix} A_1 & b_1 \\ b_1^T & c_1 \end{bmatrix}.$$

3.2. **Information theoretic interpretation.** We now describe a different application of the psd rank in the area of quantum information theory. Let M be a $p \times q$ nonnegative matrix and assume that the entries of M all sum up to 1. Then M can be seen as a joint probability distribution of a pair of random variables (X, Y), where $M_{ij} = P(X = i, Y = j)$. It is known that a nonnegative factorization of M can be interpreted as a representation of (X, Y) as a mixture of independent random variables, see e.g., [CR93, Section 6]. As such the *nonnegative rank* of M defines a certain measure of correlation between random variables X and Y. In this section we show that a similar interpretation of the psd rank holds, and that $\operatorname{rank}_{\mathrm{psd}} M$ gives a measure of correlation between X and Y in terms of *quantum information theory*. This quantum interpretation of the psd rank was first pointed out in the paper [JSWZ13].

Remark 3.9. We remark that this is not the only known interpretation of the psd rank in quantum information theory: in [FMP+12] the authors show that the psd rank of a matrix M is equal to *the one-way quantum communication complexity of computing the matrix M in expectation*. Also the psd rank is tightly related to the problem of determining the smallest dimension of a Hilbert space that explains certain measured correlations, see e.g., [BPA+08, WCD08] for more details. For space reasons, however, we focus only on the interpretation of [JSWZ13] in terms of correlation of two random variables X, Y.

3.2.1. *Correlation generation.* Given a pair of random variables (X, Y), consider the following *correlation generation* game: Two parties, Alice and Bob (for short, A and B), want to generate samples from the pair of variables (X, Y). Alice outputs samples from X and Bob outputs samples from Y and they want to do it in such a way that the samples follow the *joint* distribution of (X, Y). Note that if X and Y were independent each party could independently sample from the marginals and they would successfully achieve their objective. However if X and Y are correlated then the two parties A and B must either communicate together or share some common information in order to achieve their task. We will show here that the minimum amount of *quantum* information that A and B need to have in common is precisely $\log \text{rank}_{\text{psd}} M$ where M is the matrix giving the joint distribution of (X, Y). Thus this shows that $\log \text{rank}_{\text{psd}} M$ gives a measure of the correlation between X and Y in terms of quantum information theory.
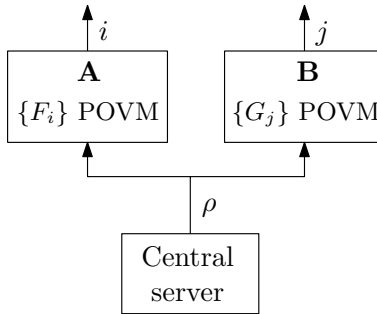
Figure 4. Quantum correlation generation problem: A and B generate samples from the joint distribution (X, Y) using shared information provided by a central server. The psd rank characterizes the minimum amount of quantum information that has to be shared between the two parties.

We first recall some basic terminology from quantum information theory. The state of a finite-dimensional quantum system is represented by a Hermitian positive semidefinite matrix ρ with trace 1, called a *density operator*. For convenience, we will work here with real symmetric matrices (instead of complex Hermitian) since our definition of psd rank involves real symmetric matrices. The state of a *bipartite* system is described by a density operator ρ of dimension $n_1 n_2$ where n_1 is the dimension of the first part (or subsystem) and n_2 is the dimension of the second part. Measurements in quantum mechanics are formalized using the concept of *POVM* (short for Positive Operator-Valued Measure). A POVM is a finite collection of psd matrices $F_1, \ldots, F_p \in \mathcal{S}_+^k$ that satisfy $\sum_{i=1}^p F_i = I_k$ where I_k is the identity matrix. The outcome of measuring a state ρ using a POVM $\{F_1, \ldots, F_p\}$ is $i \in \{1, \ldots, p\}$ with probability $\text{trace}(F_i \rho)$. Note that $\text{trace}(F_i \rho) \geq 0$ and $\sum_{i=1}^p \text{trace}(F_i \rho) = \text{trace}(\rho) = 1$.

Let M be a $p \times q$ nonnegative matrix such that $M_{ij} = P(X = i, Y = j)$ where (X, Y) is a pair of random variables. Assume we have a decomposition of M of the form:

$$(15) \qquad M_{ij} = \text{trace}((F_i \otimes G_j)\rho) \quad \forall i = 1, \ldots, p, \ j = 1, \ldots, q$$

where $\rho \in \mathcal{S}^{k^2}$ is a bipartite quantum state (where each subsystem has dimension k) and where $\{F_i\}$ and $\{G_j\}$ are POVMs, i.e., $F_i, G_j \in \mathcal{S}_+^k$ and $\sum_{i=1}^p F_i = \sum_{j=1}^q G_j = I_k$. The notation $\otimes$ here indicates Kronecker product. If there is such a decomposition of M, then one can produce samples from the pair (X, Y) using the help of a central server as follows (cf. Figure 4): the central server sends the first part of the bipartite state ρ to Alice and the second part to Bob (each part has dimension k). Alice and Bob perform measurements using POVMs $\{F_i\}$ and $\{G_j\}$ respectively and output the outcomes i and j of their measurements. The laws of quantum mechanics say that the outcome (i, j) occurs with probability $\text{trace}((F_i \otimes G_j)\rho)$. Identity (15) thus guarantees that the outputs of Alice and Bob follow the joint distribution of (X, Y).

The cost of the protocol described above is the number of quantum bits communicated by the central server to Alice and Bob, which in this case is $\log k$ (a quantum system of dimension n is represented using $\log n$ qubits). We are thus interested in the smallest k for which a decomposition of M of the form (15) exist. It turns out that this is equal to the psd rank of M as we show in the next proposition.

Proposition 3.10. [JSWZ13] *Let M be a $p \times q$ nonnegative matrix where all the entries sum up to one. Let $k \geq 1$. Then the following are equivalent:*
(i) M admits a psd factorization of size k, i.e., there exist $A_i, B_j \in \mathcal{S}_+^k$ such that $M_{ij} = \langle A_i, B_j \rangle$ for all $i = 1, \ldots, p$ and $j = 1, \ldots, q$.
(ii) There is a quantum protocol for the correlation generation problem using $\log k$ qubits, i.e., M admits a factorization of the form (15) of size k.

Proof. (ii) $\Rightarrow$ (i): Assume we have a decomposition of M of the form

$$M_{ij} = \text{trace}((F_i \otimes G_j)\rho) \quad \forall i = 1, \ldots, p, \ j = 1, \ldots, q$$

where $F_i \in \mathcal{S}_+^k$ and $G_j \in \mathcal{S}_+^k$ are psd matrices such that $\sum_{i=1}^p F_i = \sum_{j=1}^q G_j = I_k$ and $\rho \in \mathcal{S}^{k^2}$ is psd such that $\text{trace}(\rho) = 1$. Assume for simplicity that ρ is rank-one, i.e., $\rho = \psi\psi^T$ where $\psi \in \mathbb{R}^{k^2}$ (the general case is very similar). Since $\psi \in \mathbb{R}^{k^2} \cong \mathbb{R}^{k \times k}$ we know that $\text{rank}\,\psi \leq k$ and so we can write:

$$\psi = \sum_{s=1}^k v_s \otimes w_s,$$

where $v_k, w_k \in \mathbb{R}^k$. Let V and W be the matrices with the v_s and w_s in columns, i.e., $V = [v_1 | \ldots | v_k]$ and $W = [w_1 | \ldots | w_k]$. Define $A_i = V^T F_i V$ and $B_j = W^T G_j W$ for $i = 1, \ldots, p$ and $j = 1, \ldots, q$. Clearly A_i and B_j are psd and have size k. We claim that A_i, B_j give a psd factorization of M of size k. Indeed we have:

$$\begin{aligned}
\text{trace}((F_i \otimes G_j)\psi\psi^T) &= \psi^T(F_i \otimes G_j)\psi \\
&= \sum_{1 \leq s,t \leq k} (v_s \otimes w_s)^T(F_i \otimes G_j)(v_t \otimes w_t) \\
&\overset{(*)}{=} \sum_{1 \leq s,t \leq k} (v_s^T F_i v_t)(w_s^T G_j w_t) \\
&= \langle V^T F_i V, W^T G_j W \rangle = \langle A_i, B_j \rangle.
\end{aligned}$$

where in $(*)$ we used the mixed-product property of the Kronecker product $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$.

(i) $\Rightarrow$ (ii): We now prove the other direction. Assume we have a psd factorization of M of the form $M_{ij} = \langle A_i, B_j \rangle$ where $A_i, B_j \in \mathcal{S}_+^k$. We show how to construct a factorization of the form (15). Let Σ_A, Σ_B be defined respectively by $\Sigma_A = \sum_{i=1}^p A_i$ and $\Sigma_B = \sum_{j=1}^q B_j$. Note that Σ_A and Σ_B can be assumed to be invertible (otherwise we can reduce the size of the psd factorization). Consider the matrices F_i and G_j defined by:

$$F_i = \Sigma_A^{-1/2} A_i \Sigma_A^{-1/2} \ (i = 1, \ldots, p) \quad \text{and} \quad G_j = \Sigma_B^{-1/2} B_j \Sigma_B^{-1/2} \ (j = 1, \ldots, q).$$

Then $F_i, G_j \succeq 0$ and $\sum_{i=1}^p F_i = \sum_{j=1}^q G_j = I_k$. We now construct the state $\rho \in \mathcal{S}^{k^2}$ of the protocol. To do so recall that we have the following simple fact: If A and B are symmetric matrices of size k, then

$$\text{trace}(AB) = e^T(A \otimes B)e$$

where $e = \mathrm{vec}(I_k) \in \mathbb{R}^{k^2}$ is the vector obtained by stacking all the columns of I_k into a single column of dimension k^2. Let $\rho \in \mathcal{S}^{k^2}$ be defined by $\rho = \psi\psi^T$ where

$$\psi = (\Sigma_A^{1/2} \otimes \Sigma_B^{1/2})e.$$

First note that ρ is a valid state and $\mathrm{trace}(\rho) = 1$ since

$$\mathrm{trace}(\rho) = \psi^T\psi = e^T(\Sigma_A \otimes \Sigma_B)e = \mathrm{trace}(\Sigma_A\Sigma_B) = \sum_{\substack{1 \le i \le p \\ 1 \le j \le q}} \mathrm{trace}(A_iB_j) = \sum_{\substack{1 \le i \le p \\ 1 \le j \le q}} M_{ij} = 1.$$

We now claim that the choice of $\{F_i\}, \{G_j\}$ and ρ gives a valid decomposition of M as in (15). Indeed we have:

$$\begin{aligned}
\mathrm{trace}((F_i \otimes G_j)\rho) &= \psi^T(F_i \otimes G_j)\psi \\
&= e^T(\Sigma_A^{1/2} \otimes \Sigma_B^{1/2})(\Sigma_A^{-1/2}A_i\Sigma_A^{-1/2} \otimes \Sigma_B^{-1/2}B_j\Sigma_B^{-1/2})(\Sigma_A^{1/2} \otimes \Sigma_B^{1/2})e \\
&= e^T(A_i \otimes B_j)e = \mathrm{trace}(A_iB_j) = M_{ij}.
\end{aligned}$$

$\square$

4. Relationships between ranks

Recall from Proposition 2.5 that for a nonnegative matrix $M \in \mathbb{R}_+^{p \times q}$,

$$(16) \qquad \frac{1}{2}\sqrt{1 + 8\,\mathrm{rank}\,(M)} - \frac{1}{2} \le \mathrm{rank}_{\mathrm{psd}}\,(M) \le \mathrm{rank}_+\,(M).$$

The first inequality is equivalent to saying that for all nonnegative matrices M,

$$(17) \qquad \mathrm{rank}\,(M) \le \binom{\mathrm{rank}_{\mathrm{psd}}\,(M) + 1}{2}.$$

This says that while rank may be larger than psd rank, it cannot be much larger, since it is bounded above by a quadratic function of the psd rank. In this section, we examine the relationships between the three ranks present in inequality (16) and a fourth type of rank called *square root rank*. We begin by showing that all inequalities in (16) can be tight. An easy example for the second inequality is the $n \times n$ identity matrix for which $\mathrm{rank}_{\mathrm{psd}}\,(I_n) = \mathrm{rank}_+\,(I_n) = n$.

Example 4.1 (**Derangement matrices**)**.** The $n \times n$ *derangement matrix* D_n is the matrix with zeros on the diagonal and ones elsewhere. It verifies $\mathrm{rank}\,(D_n) = n$ for all n. Fix a positive integer k and let $n := \binom{k+1}{2}$. We will exhibit a factorization of D_n through $\mathcal{S}_+^k$ which will show that $\mathrm{rank}_{\mathrm{psd}}\,(D_n) \le k$, making the first inequality tight.

To construct a psd factorization of D_n through $\mathcal{S}_+^k$ choose factors as follows: For $i = 1, \ldots, k$, let $A_i = e_ie_i^T$ where e_i is the ith standard basis vector in $\mathbb{R}^k$. Since $\binom{k+1}{2} = \binom{k}{2} + k$, we need to define $\binom{k}{2}$ further A_i matrices. Let $F = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$. For each $i, j \in \{1, \ldots, k\}$ with $i < j$, define $F_{i,j}$ to be equal to the $k \times k$ matrix that has its i, j principal submatrix equal to F and all other entries equal to 0. Let $A_{k+1} = F_{1,2}$, $A_{k+2} = F_{1,3}$, $\ldots$, $A_{2k} = F_{2,3}$, $A_{2k+1} = F_{2,4}$, and so on. Now we define matrices $B_1, \ldots, B_n$ for the columns. First, let E be the $(k-1) \times (k-1)$ matrix with ones on the diagonal and $\frac{1}{2}$ everywhere else. For $i = 1, \ldots, k$, let B_i be the matrix whose ith row and column are identically zero and whose remaining entries form the matrix E. For $i > k$, we obtain the matrix B_i from A_i by the following: First change all nonzero entries and all diagonal entries of A_i to ones. Then change all remaining zero entries to $\frac{1}{2}$. Call the resulting matrix B_i. The matrices A_i, B_j form a psd factorization of D_n.

We present the case $k = 3$ below:

$$D_6 = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

Then $A_1, \ldots, A_6$ are:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix},$$

and $B_1, \ldots, B_6$ are:

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & \frac{1}{2} & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ \frac{1}{2} & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & \frac{1}{2} \\ 1 & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix}, \begin{bmatrix} 1 & \frac{1}{2} & 1 \\ \frac{1}{2} & 1 & \frac{1}{2} \\ 1 & \frac{1}{2} & 1 \end{bmatrix}, \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 & 1 \\ \frac{1}{2} & 1 & 1 \end{bmatrix}.$$

We showed that $\mathrm{rank}_{\mathrm{psd}}(D_n) = k$ whenever $n = \binom{k+1}{2}$. Now suppose that n is strictly between $\binom{k}{2}$ and $\binom{k+1}{2}$. Then the rank lower bound (first inequality in (16)) implies that $\mathrm{rank}_{\mathrm{psd}}(D_n) > k-1$. Since D_n is a submatrix of $D_{\binom{k+1}{2}}$ for which we know a size k psd factorization, $\mathrm{rank}_{\mathrm{psd}}(D_n) \leq k$. Thus, $\mathrm{rank}_{\mathrm{psd}}(D_n) = k$ for these intermediate values of n, or equivalently,

$$\mathrm{rank}_{\mathrm{psd}}(D_n) = \min\left\{ k \: : \: n \leq \binom{k+1}{2} \right\} \text{ for all } n.$$

4.1. Square root rank: an upper bound for psd rank.

Definition 4.2. The *square root rank* of a nonnegative matrix M, denoted as $\mathrm{rank}_{\sqrt{}}(M)$, is the minimum rank of a Hadamard square root of M.

For a quick example of square root rank, note that the following matrix M (of rank 3) has $\mathrm{rank}_{\sqrt{}}(M) = 2$ as evidenced by the shown square root S.

$$M = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 4 \\ 1 & 1 & 1 \end{pmatrix} \longrightarrow S = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 1 & 1 & -1 \end{pmatrix}$$

Recall from the proof of Theorem 2.9 (v) that if a Hadamard square root of M has rank r then there is a psd factorization of M by matrices of rank one lying in the psd cone $\mathcal{S}_+^r$. This implies the following corollary.

Corollary 4.3. *For a nonnegative matrix $M \in \mathbb{R}^{p \times q}$, $\mathrm{rank}_{\mathrm{psd}}(M) \leq \mathrm{rank}_{\sqrt{}}(M)$. In particular, if M is a 0/1 matrix, then $\mathrm{rank}_{\mathrm{psd}}(M) \leq \mathrm{rank}(M)$.*

The second statement of Corollary 4.3 says that if a matrix has only the two distinct entries 0 and 1, then its psd rank is bounded above by rank. This was extended by Barvinok [Bar12] to an upper bound on the psd rank of a matrix in terms of its rank and number of distinct entries.

Lemma 4.4. [Bar12, Lemma 4.4]
Let $A = (a_{ij})$ be a real matrix and $f : \mathbb{R} \to \mathbb{R}$ be a polynomial of degree k. If $B = (b_{ij})$ is such that $b_{ij} = f(a_{ij})$ for all i, j, then

$$\mathrm{rank}(B) \leq \binom{k + \mathrm{rank}(A)}{k}.$$

Corollary 4.5. [Bar12, Lemma 4.6] *If the number of distinct entries in a nonnegative matrix M does not exceed k, then* $\mathrm{rank}_{\mathrm{psd}}(M) \leq \binom{k-1+\mathrm{rank}(M)}{k-1}$.

Proof. Let $\mathcal{M}$ be the set of distinct entries in M and $\phi : \mathcal{M} \to \mathbb{R}_+$ be the square root function taking nonnegative values. Since $|\mathcal{M}| \leq k$, there exists a polynomial $f(t)$ of degree $k-1$ such that $\phi(t) = f(t)$ on $\mathcal{M}$. If we let S be the nonnegative entrywise square root of M, then by Lemma 4.4 and Corollary 4.3 we have that

$$\mathrm{rank}_{\mathrm{psd}}(M) \leq \mathrm{rank}(S) \leq \binom{k-1+\mathrm{rank}(M)}{k-1}.$$

$\square$

While we strongly suspect that it is NP-hard to compute psd rank, there is no proof of this fact at the moment. The situation is clearer for square root rank.

Theorem 4.6. *The square root rank of a nonnegative matrix is NP-hard to compute.*

Proof. Recall that given a list of n positive integers $a_1, \ldots, a_n$, the partition problem asks whether there exist sign choices $s_1, \ldots, s_n \in \{-1, 1\}$ such that $\sum_{i=1}^{n} s_i a_i = 0$. This problem is known to be NP-complete [GJ79].

Given the integers $a_1, \ldots, a_n$, define an $(n+1) \times (n+1)$ matrix A of the form:

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & a_1^2 \\ 0 & 1 & 0 & \cdots & 0 & a_2^2 \\ 0 & 0 & 1 & \cdots & 0 & a_3^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & a_n^2 \\ 1 & 1 & 1 & \cdots & 1 & 0 \end{pmatrix}.$$

Since A contains the $n \times n$ identity matrix as a submatrix, the square root rank of A must be either n or $n+1$. If S is a Hadamard square root of A, then we may scale rows and columns of S by -1 and not affect the rank. Thus, we may assume that the first n columns of S are composed of zeros and positive ones. With this assumption, we see immediately that there exists a Hadamard square root of rank n if and only if the partition problem for $a_1, \ldots, a_n$ is satisfiable. $\square$

Remark 4.7. Although the partition problem is NP-complete, it is only weakly NP-complete and admits a pseudo-polynomial time algorithm. Thus, the above theorem does not rule out the existence of an algorithm for deciding $\mathrm{rank}_{\sqrt{}}$ that runs in time polynomial in the problem dimension and the *magnitude* (not encoding length) of the matrix entries. Furthermore, this embedding of the partition problem cannot hope to show that psd rank is NP-hard to compute. To see this, consider the matrix A corresponding to the partition problem with integers 5, 12, and 13:

$$\begin{pmatrix} 1 & 0 & 0 & 25 \\ 0 & 1 & 0 & 144 \\ 0 & 0 & 1 & 169 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

This instance of the partition problem is not satisfiable, yet the matrix A has a 3×3 psd factorization. Such a factorization is obtained by placing the matrices

$$\begin{pmatrix} 1 & 0 & -\frac{5}{13} \\ 0 & 1 & -\frac{12}{13} \\ -\frac{5}{13} & -\frac{12}{13} & 1 \end{pmatrix}, \begin{pmatrix} 25 & 60 & 65 \\ 60 & 144 & 156 \\ 65 & 156 & 169 \end{pmatrix}$$

on the fourth row and the fourth column, respectively, of A and by placing the standard basis factorization of the identity in the first three rows and columns.

4.2. **Lower bounds for psd rank.** Lower bounding the psd rank has shown to be a difficult task. In this section, we discuss the known lower bounding techniques and their limitations.

We say that two matrices of the same dimensions have the same *support* if they share the same zero/nonzero pattern in their entries. Lower bounds based solely on the support of the matrix have been shown to be quite powerful in the case of nonnegative rank (see [FKPT13] for an overview). In the case of psd rank, their power is much more limited. Given a matrix M, the entry-wise square $M \odot M$ has the same support as M and has psd rank bounded above by $\text{rank}(M)$ (Theorem 2.9, part (v)). Thus, a purely support-based bound cannot produce a lower bound that is higher than the rank of M. This observation was extended by Lee and Theis in [LT12] as follows:

Theorem 4.8. [LT12, Theorem 1.1] *Fix a support Z and let $\mathcal{M}_Z$ be the set of all matrices sharing this support. Then*

$$\min_{A \in \mathcal{M}_Z} \text{rank}(A) = \min_{A \in \mathcal{M}_Z, A \geq 0} \text{rank}_{\text{psd}}(A).$$

If a nonnegative matrix has the property that it achieves the minimum rank possible among all matrices sharing its support, then the rank is a lower bound to the psd rank. In particular, slack matrices of polytopes have this property. In [GRT13a], the authors showed the following corollary and characterized those polytopes that achieve this lower bound in $\mathbb{R}^2$ and $\mathbb{R}^3$.

Corollary 4.9. [GRT13a, Proposition 3.2] *If P is an n-dimensional polytope with slack matrix S_P, then $\text{rank}(S_P) = n + 1 \leq \text{rank}_{\text{psd}}(S_P)$.*

To obtain stronger lower bounds, we need to move past using only the support of a matrix. The only known lower bounding techniques that are not purely support based rely on the quantifier elimination theory of Renegar [Ren92] as seen in [GPT13].

We give a high level discussion of the idea behind this technique and then the result. For complete proofs, see [GPT13]. Suppose we are given a convex set $C \subset \mathbb{R}^n$ that has a psd lift into $\mathcal{S}^k_+$, i.e. there exists a linear map π and an affine subspace $\mathcal{L}$ such that $C = \pi(\mathcal{L} \cap \mathcal{S}^k_+)$. Then $\mathcal{L} \cap \mathcal{S}^k_+$ is a semialgebraic set where the bounding polynomials have degree at most k.

Theorem 4.10. [Ren06] *Let $Q = \{z \in \mathbb{R}^m : C + \sum z_i A_i \succeq 0\}$ be a spectrahedron with $E := C + \sum z'_i A_i \succ 0$ for some $z' \in Q$, and C, A_i are symmetric matrices of size $k \times k$. Then Q is a semialgebraic set described by $g^{(i)}(z) \geq 0$ for $i = 1, \ldots, k$ where $g^{(0)}(z) := \det(C + \sum z_i A_i)$ and $g^{(i)}(z)$ is the i-th Renegar derivative of $g^{(0)}(z)$ in direction E.*

The work of Renegar says that when we project this set, the degree and number of the resulting bounding polynomials of C are bounded in k and n.

Theorem 4.11. [Ren92, Theorem 1.1] *Given a formula of the form*

$$\exists \, y \in \mathbb{R}^{m-n} : g_i(x, y) \geq 0 \ \ \forall i = 1, \ldots, s$$

where $x \in \mathbb{R}^n$ and $g_i \in \mathbb{R}[x, y]$ are polynomials of degree at most d, there exists a quantifier elimination method that produces a quantifier free formula of the form

$$(18) \qquad \bigvee_{i=1}^{I} \bigwedge_{j=1}^{J_i} (h_{ij}(x) \, \Delta_{ij} \, 0)$$

where $h_{ij} \in \mathbb{R}[x]$, $\Delta_{ij} \in \{>, \geq, =, \neq, \leq, <\}$ such that

$$I \leq (sd)^{Kn(m-n)}, \ J_i \leq (sd)^{K(m-n)}$$

and the degree of h_{ij} is at most $(sd)^{K(m-n)}$, where K is a constant.

Multiplying all of these polynomials together, we obtain a single polynomial, whose degree is bounded in k and n, that vanishes on the boundary of C. Hence, if we know that every polynomial

that vanishes on the boundary of C must have very high degree, then we can say that C does not have a $\mathcal{S}_+^k$-lift for small k. Recall that the *Zariski closure* of a set $X \subseteq \mathbb{R}^n$ is the smallest algebraic variety in $\mathbb{R}^n$ that contains X. The Zariski closure of the boundary of C is a hypersurface in $\mathbb{R}^n$ since the boundary of C has codimension one. We define the *degree of C* to be the minimal degree of a (nonzero) polynomial whose zero set is the Zariski closure of the boundary of C. By construction, this polynomial vanishes on the boundary of C.

Proposition 4.12. [GPT13, Proposition 6] *If $C \subseteq \mathbb{R}^n$ is a full-dimensional convex semialgebraic set with a $\mathcal{S}_+^k$-lift, then the degree of C is at most $k^{O(k^2 n)}$.*

Note that when C is a polytope, the degree of C is equal to the number of facets of C: indeed if a polynomial p vanishes on the boundary of C then each linear polynomial defining a facet of C must divide p and so the degree of p is at least the number of facets of C; conversely the polynomial obtained as the product of all the linear polynomials defining the facets of C vanishes on the boundary of C. This allows us to lower bound the psd rank of slack matrices of polytopes.

Corollary 4.13. [GPT13, Corollary 4] *If $C \subset \mathbb{R}^n$ is a full-dimensional polytope whose slack matrix has psd rank k, then C has at most $k^{O(k^2 n)}$ facets.*

Example 4.14. Corollary 4.13 shows that as the number of facets in an n-dimensional polytope in $\mathbb{R}^n$ increases, the psd rank of the slack matrix of the polytope has to increase. However, the rank of any such slack matrix stays fixed at $n + 1$. For example, let S_d be the slack matrix of a d-gon in the plane. Then by Corollary 4.13, $\mathrm{rank}_{\mathrm{psd}}(S_d)$ grows to infinity as d increases. As we have seen before, however, $\mathrm{rank}(S_d) = 3$ for all d. This provides a first example of a family of matrices with arbitrarily large gap between rank and psd rank.

For non-slack matrices, it can still be possible to apply this lower bound by viewing the matrix as a generalized slack matrix.

Example 4.15. We now construct a matrix family that has the same zero pattern as the derangement matrices and for which rank is three and psd rank grows arbitrarily large. Let C_d be a convex semialgebraic set in the plane whose bounding polynomial has degree d. By results of Scheiderer [Sch12], C_d has a $\mathcal{S}_+^r$-lift for some finite r, and suppose k is the smallest such r. By Proposition 4.12, $d \leq k^{O(k^2)}$.

Now pick $d^2 + 1$ distinct points on the boundary of C_d and let P be the convex hull of these points. Also, let Q be the polyhedron whose facet inequalities are given by the tangent lines to C_d at the vertices of P. Then the slack matrix of the pair P, Q, which was denoted as $S_{P,Q}$ in Section 3.1, is a nonnegative matrix with the same zero pattern as the derangement matrix. Call this nonnegative matrix M_d. By construction, the set C_d is sandwiched between P and Q and its boundary passes through the vertices of P. If C' is another convex semialgebraic set that is sandwiched between P and Q, then its boundary must also contain the vertices of P because P and Q touch at these vertices. By Bezout's theorem, the degree of C' must be at least $d + 1$. By Theorem 3.3, the $\mathrm{rank}_{\mathrm{psd}} M_d$ is the smallest k such that a slice of $\mathcal{S}_+^k$ projects to a convex set nested between P and Q. Since this smallest k grows as the degree of the polynomial bounding the projected spectrahedron grows, it must be that $\mathrm{rank}_{\mathrm{psd}} M_d$ grows with d.

On the other hand, $\mathrm{rank}(M_d) = 3$ since it is the slack matrix of a pair of polygons. Therefore, by choosing a family of convex semialgebraic sets in the plane of increasing degree with the requirements specified above, one can obtain a family of nonnegative matrices of rank three and growing psd rank. For instance, take

$$C_{d=2t} := \{(x, y) \in \mathbb{R}^2 \ : \ x^{2t} + y^{2t} \leq 1\}.$$

This quantifier elimination lower bound framework has proven useful for showing that certain families of matrices must have growing psd rank. Its usefulness is limited, however, when considering

the psd rank of a single matrix. Other techniques have been developed to show matrices with high psd rank. Briët et. al. used a counting argument in [BDP13] to show that most 0/1-polytopes in $\mathbb{R}^n$ have psd rank that is at least exponential in n. Gouveia et. al. produced a lower bound for generic polytopes (polytopes whose vertices are algebraically independent) in [GRT13b], but again, these techniques are of limited usefulness when considering a single specific matrix. An answer to the following problem would likely provide a new technique that is applicable to many open questions in this field.

Problem 4.16. Produce a 10×10 nonnegative matrix M with integral entries such that $\text{rank}(M) = 3$ and $\text{rank}_{\text{psd}}(M) \geq 5$.

4.3. **Comparisons between ranks.** We can now compare all the ranks seen so far. To keep track, we summarize the relationships in Table 1.

Table 1. Relationships between various ranks

	rank	rank$_+$	rank$_{\text{psd}}$	rank$_{\sqrt{}}$
rank	$=$	$\ll$ (4.17)	$\ll$ (4.14, 4.15) , $>$ (4.1)	$\ll$ (4.18), $>$ (4.17)
rank$_+$	$\gg$	$=$	$\gg$ (4.17)	$\ll$ (4.18), $\gg$ (4.17)
rank$_{\text{psd}}$	$<, \gg$	$\ll$	$=$	$\ll$ (4.18)
rank$_{\sqrt{}}$	$<, \gg$	$\ll$, $\gg$	$\gg$	$=$

The symbol $\ll$ refers to the rank indexing the row being arbitrarily smaller than the one indexing the column (i.e. there does not exist a function of the row rank that upper bounds the column rank). The symbol $<$ indicates that the rank on the row may be smaller than the rank on the column, but the gap cannot not be arbitrarily large. For example, the entry in the $(1, 2)$-position says that rank may be arbitrarily smaller than nonnegative rank, but never larger. The $(1, 3)$-entry says that rank may be smaller or larger than psd rank. The gap in the first case may be arbitrarily large, but the gap in the second case is controlled (see (17)). The numbers in the table refer to examples exhibiting the relationship.

Example 4.17 (**Euclidean distance matrices**). Consider the $n \times n$ *Euclidean distance matrix* M_n whose (i, j)-entry is $(i - j)^2$. The rank of M_n is three for all n since $M_n = A_n^T B_n$ where column i of A_n is $(i^2, -2i, 1)$ and column j of B_n is $(1, j, j^2)$.

The square root rank and psd rank of M_n are two for all n since the matrix with (i, j)-entry equal to $i - j$ has usual rank two. So for all n, the matrix M_n has constant rank, psd rank, and square root rank.

Now we show that M_n has growing nonnegative rank. Suppose M_n has a $\mathbb{R}_+^k$-factorization. Then there exists $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{R}_+^k$ such that $\langle a_i, b_j \rangle \neq 0$ for all $i \neq j$. Notice that if $\text{supp}(b_j) \subseteq \text{supp}(b_i)$ then $\langle a_i, b_i \rangle = 0$ implies $\langle a_i, b_j \rangle = 0$, and hence, all the b_i's (and also all the a_i's) must have supports that are pairwise incomparable. Since there are at most 2^k possibilities for these supports, $n \leq 2^k$, or equivalently, $k \geq \log_2 n$. Therefore we get that $\text{rank}_+(M_n) \geq \log_2 n$. (Note that one can actually obtain a slightly better lower bound using Sperner's lemma: indeed, since we know that there can be at most $\binom{k}{\lfloor \frac{k}{2} \rfloor}$ incomparable support patterns we get that $\binom{k}{\lfloor \frac{k}{2} \rfloor} \geq n$, which is slightly better than $2^k \geq n$).

In [Hru12], Hrubeš exhibited a nonnegative factorization to show that the nonnegative rank of this family is actually $\Theta(\log_2 n)$.

Example 4.18 (**Prime matrices**). Let $n_1, n_2, n_3, \ldots$ be an increasing sequence of positive integers such that $2n_k - 1$ is prime for each k. Let $\mathcal{P}_k$ denote the set of all primes strictly less than $2n_k - 1$. Define a $k \times k$ matrix Q^k such that $Q_{ij}^k = n_i + n_j - 1$. Then Q^k has usual rank two for all k.

Consequently, by Proposition 2.8, the nonnegative rank and psd rank of Q^k are also two for all k. For example, suppose our sequence has the form $2, 3, 4, 6, \ldots$ Then $Q^1, \ldots, Q^4$ will have the form:

$$(\,3\,), \begin{pmatrix} 3 & 4 \\ 4 & 5 \end{pmatrix}, \begin{pmatrix} 3 & 4 & 5 \\ 4 & 5 & 6 \\ 5 & 6 & 7 \end{pmatrix}, \begin{pmatrix} 3 & 4 & 5 & 7 \\ 4 & 5 & 6 & 8 \\ 5 & 6 & 7 & 9 \\ 7 & 8 & 9 & 11 \end{pmatrix}.$$

Note that the top left block of each Q^k is Q^{k-1} and that the diagonal entries of Q^k are the increasing sequence of primes, $2n_1 - 1, 2n_2 - 1, \ldots$

We will prove by induction that Q^k has full square root rank for each k. The base case is clear so assume that Q^{k-1} has full square root rank, i.e. every possible Hadamard square root of Q^{k-1} has rank equal to $k-1$. Fix a Hadamard square root of Q^k and let M be the matrix equal to this square root in every entry except M_{kk}. Let M_{kk} be the variable x. The determinant of M has the form $\alpha x + \beta$ where α and β are in the extension field $\mathbb{Q}(\sqrt{\mathcal{P}_k})$. By properties of the determinant, α is equal to the determinant of the top left $(k-1) \times (k-1)$ block. Thus by the induction assumption, α is nonzero. Hence, any x making the determinant zero must also lie in $\mathbb{Q}(\sqrt{\mathcal{P}_k})$. However, our square root of Q^k must have $x = \pm\sqrt{2n_k - 1}$. Thus, the square root of Q^k must have full rank.

The remaining relationship shown in Table 1 that we have not discussed is $\text{rank} > \text{rank}_{\sqrt{}}$. In the example after Definition 4.2, we saw that rank can be larger than square root rank. The possible gap is controlled, however, since square root rank is an upper bound to psd rank and the gap between rank and psd rank is controlled.

5. PROPERTIES OF FACTORS

The matrices used as factors in a positive semidefinite factorization can sometimes be chosen to satisfy specific constraints. In this section we explore the ranks and norms of the factors used in a psd factorization.

5.1. **Rank of factors.** In [LT12] it has been shown that we can always pick the factors of a factorization to have some bounded rank, depending only on the size of the matrix:

Proposition 5.1. [LT12, Lemma 4.5] *If a $p \times q$ nonnegative matrix $M \in \mathbb{R}^{p \times q}_+$ has $\text{rank}_{\text{psd}} M \leq k$, then it has a $\mathcal{S}^k_+$ factorization $M_{ij} = \langle A_i, B_j \rangle$ where $\text{rank}\, A_i \leq \sqrt{8q+1}/2$ and $\text{rank}\, B_j \leq \sqrt{8p+1}/2$ for all $i = 1, \ldots, p$ and $j = 1, \ldots, q$.*

The reason is simple. If we fix the factors corresponding to rows, the set of valid factors for any given column is the feasible set of a semidefinite program, and standard results from convex optimization guarantee the existence of a solution with bounded rank ([Pat98],[Bar01]). Fixing these column factors and repeating the process over the rows we get the result.

We are particularly interested in knowing when the factors be chosen to have rank one. The answer turns out to be given by the $\text{rank}_{\sqrt{}}(M)$ (Definition 4.2).

Proposition 5.2. *The square root rank of M, $\text{rank}_{\sqrt{}}(M)$, is precisely the smallest size of a psd factorization of M comprised solely of rank one factors.*

Proof. As seen in the proof of Theorem 2.9 (v), if $M = N \odot N$ and N has rank k, then we can take a rank factorization of N, $N = A^T B$, and use it to create the matrices $A_i = a_i a_i^T$ and $B_j = b_j b_j^T$, where a_i and b_j range over the columns of A and B respectively. These matrices A_i and B_j have rank one and, by construction, form a $\mathcal{S}^k_+$ factorization of M.

To prove the reverse implication just note that if $v_i v_i^T$ and $w_j w_j^T$ form a $\mathcal{S}^k_+$ factorization of M then setting V and W to be the matrices whose columns are the v_i and the w_j respectively, we can obtain a matrix $V^T W$ that has rank k and is a Hadamard square root of M. $\qquad\square$

In general, we expect the gap between $\text{rank}_{\sqrt{}}(M)$ and $\text{rank}_{\text{psd}}(M)$ to be arbitrarily high, as illustrated in Example 4.18, but good knowledge of rank one factorizations can potentially provide some insight into general factorizations.

Remark 5.3. Let M be a $p \times q$ nonnegative matrix with $\text{rank}_{\text{psd}}(M) = k$. Suppose A_i, $i = 1, \ldots, p$ and B_j, $j = 1, \ldots, q$ form a $\mathcal{S}_+^k$ factorization of M. Each A_i can be written as $\sum_{l=1}^k v_{i,l} v_{i,l}^T$, and each B_j as $\sum_{l=1}^k w_{j,l} w_{j,l}^T$. Define N as the matrix indexed by $\{1, \ldots, p\} \times \{1, \ldots, k\}$ and $\{1, \ldots, q\} \times \{1, \ldots, k\}$ whose entry $N_{\{i,s\},\{j,r\}}$ is given by $(v_{i,s}^T w_{j,r})^2$.

Then, $N \in \mathbb{R}_+^{pk \times qk}$ consists of $p \times q$ blocks of size $k \times k$ and $\text{rank}_{\sqrt{}}(N) = k$. Furthermore, summing all the entries of block (i,j) of N gives us entry (i,j) of M.

This remark allows us to transfer properties from rank one factorizations to general factorizations, an example can be seen at the end of the next subsection.

5.2. Norms of factors.
Besides rank, another useful quantity to control in the factors is their size. In Section 2 we used a lemma showing that the factors of a semidefinite factorization can be rescaled to have small trace. This was instrumental in establishing the lower semicontinuity of psd rank. We restate the lemma here and provide a proof:

Lemma 5.4. *Let $M \in \mathbb{R}_+^{p \times q}$ and assume that M has a psd factorization of size k. Then M admits a psd factorization $M_{ij} = \langle A_i, B_j \rangle$ of size k where the factors satisfy $\text{trace}(A_i) \leq k$ and $\text{trace}(B_j) = \sum_{i=1}^p M_{ij}$.*

Proof. Let $M_{ij} = \langle \widehat{A}_i, \widehat{B}_j \rangle$ be an arbitrary psd factorization of M of size k. Let $S = \sum_{i=1}^p \widehat{A}_i \in \mathcal{S}_+^k$. Note that we can assume that S is invertible (otherwise the size of the psd factorization can be reduced). Define $A_i = S^{-1/2} \widehat{A}_i S^{-1/2}$ and let $B_j = S^{1/2} \widehat{B}_j S^{1/2}$. Note that $\langle A_i, B_j \rangle = \langle \widehat{A}_i, \widehat{B}_j \rangle = M_{ij}$ and so $A_1, \ldots, A_p, B_1, \ldots, B_q$ give a valid psd factorization of M of size k. Observe that by construction we have $\sum_{i=1}^p A_i = I$ (where I is the $k \times k$ identity matrix), thus necessarily each A_i satisfies $A_i \preceq I$ and thus $\text{trace}(A_i) \leq k$. Also we have for any $j = 1, \ldots, q$, $\text{trace}(B_j) = \text{trace}((\sum_{i=1}^p A_i) B_j) = \sum_{i=1}^p \text{trace}(A_i B_j) = \sum_{i=1}^p M_{ij}$ as desired. $\qquad \square$

Note that this is equivalent to saying that we can choose A_i and B_j all with trace bounded by $\sqrt{k \|M\|_{1,1}}$ where $\|M\|_{1,1}$ is the matrix norm induced by the 1-norm in $\mathbb{R}^k$. This looks very similar to another rescaling result that has proven very useful, the rescaling result in [BDP13], which was used to show that there are $0/1$-polytopes with only exponential-sized semidefinite representations. The result states that if a matrix M has psd rank k, then it has a semidefinite factorization where each factor has largest eigenvalue less than or equal to $\sqrt{k\|M\|_\infty}$, where $\|M\|_\infty$ is the maximum absolute value of an entry in M.

In the remainder of this section we present a new, simplified proof of a stronger version of this fact. As in [BDP13], the main tool we need is a version of *John's ellipsoid theorem*.

Theorem 5.5 (John's Theorem). *Let C be a full dimensional convex set in $\mathbb{R}^n$ and let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map such that the image of the unit ball B^n under T is the unique minimum volume ellipsoid E containing $C \cup -C$. Then*

$$\frac{1}{n} T T^T \in \text{conv}(\{vv^T : v \in \text{Boundary}(C) \cap \text{Boundary}(E)\}).$$

A simple consequence of John's Theorem is the following result concerning scalability of usual rank factorizations.

Corollary 5.6. *If $M \in \mathbb{R}^{p \times q}$ has rank k, then there exist $A \in \mathbb{R}^{k \times p}$, $B \in \mathbb{R}^{k \times q}$ such that $M = A^T B$ and the maximum 2-norm of a column of A or B is at most $k^{1/4} \sqrt{\|M\|_\infty}$.*

Proof. Start with any rank factorization $M = U^T V$ and let C be the convex hull of the columns of U. Let $T : \mathbb{R}^k \to \mathbb{R}^k$ such that $T(B^k)$ is the minimum volume ellipsoid containing $C \cup -C$. For all columns u of U, $\|T^{-1}(u)\|_2 \leq 1$ by construction. Furthermore by John's theorem for any column v of V

$$\|T^T(v)\|_2^2 = v^T T T^T v = v^T \left(k \sum_i \lambda_i u_i u_i^T \right) v = k \sum_i \lambda_i \langle u_i, v \rangle^2$$

with $\lambda_i \geq 0$ and $\sum \lambda_i = 1$, and where the u_i's are columns of U. But this implies $\|T^T(v)\|_2^2 \leq k\|M\|_\infty^2$ and therefore $\|T^T(v)\|_2 \leq \sqrt{k}\|M\|_\infty$. By making $L = k^{1/4}\sqrt{\|M\|_\infty}T^{-1}$, we have that $A = LU$ and $B = (L^{-1})^T V$ have the intended properties. $\square$

As mentioned in the introduction, factorizations where the factors have small norm have been studied in different contexts, particularly in Banach space theory. For example, [LMSS07, Lemma 4.2] is closely related to Corollary 5.6. The scalability of psd factorizations also follows readily using the same ideas.

Corollary 5.7. *If $M \in \mathbb{R}_+^{p \times q}$ has psd rank k, then there exist $A_1, \ldots, A_p, B_1, \ldots, B_q \in \mathcal{S}_+^k$ such that $M_{i,j} = \langle A_i, B_j \rangle$ and the largest eigenvalue of A_i is bounded above by 1 and the trace norm of B_j is bounded above by $k\|M\|_\infty$.*

Proof. Start with a factorization $A_1', \ldots, A_p', B_1', \ldots, B_q' \in \mathcal{S}_+^k$ of M, and let

$$U = \{u \in \mathbb{R}^k \ : \ A_i' - uu^T \succeq 0 \text{ for some } i\}.$$

Applying John's Theorem to $\text{conv}(U)$, we get an operator T such that $T^{-1}(U)$ is contained in the unit ball, and $TT^T = k \sum_i \lambda_i u_i u_i^T$ with $\lambda_i \geq 0$ and $\sum \lambda_i = 1$ and each u_i verifying $A_l' \succeq u_i u_i^T$ for some l. Let $A_i = T^{-1} A_i' (T^{-1})^T$ and $B_j = T^T B_j' T$ then

$$T^{-1}(U) = \{u \in \mathbb{R}^k \ : \ A_i - uu^T \succeq 0 \text{ for some } i\}.$$

If α is the largest eigenvalue of A_i and v is an associated unit norm eigenvector, then note that $\sqrt{\alpha} v \in U$ and so $\alpha = \|\sqrt{\alpha}v\|_2^2 \leq 1$. Thus this shows that the largest eigenvalue of each A_i is at most 1. For B_j note that

$$\text{trace}\,(B_j) = \text{trace}\,(TT^T B_j') = k\,\text{trace}\,\left(\sum_i \lambda_i u_i u_i^T B_j' \right) = k \sum_i \lambda_i u_i^T B_j' u_i.$$

Since for each i there is l such that $A_l' - u_i u_i^T \succeq 0$, we have $u_i^T B_j' u_i \leq \text{trace}\,(A_l' B_j') = M_{lj} \leq \|M\|_\infty$, and so

$$\text{trace}\,(B_j) \leq k \sum_i \lambda_i \|M\|_\infty = k\|M\|_\infty.$$

$\square$

Note that this does indeed imply the scaling result of [BDP13], since the trace norm is larger than the largest eigenvalue, hence by rescaling the A_i and B_j we can get them all with largest eigenvalue at most $\sqrt{k\|M\|_\infty}$. If we are only interested in getting a bound of $k^{5/4}\sqrt{\|M\|_\infty}$ (which is already enough for the application in [BDP13]) we can derive it directly from Corollary 5.6, together with Remark 5.3, illustrating that properties valid for rank one factors can sometimes be extended to general factorizations.

Remark 5.8. It is worth noting that while Lemma 5.4 and the result of [BDP13] look similar, the bounds they provide are in general not comparable. In general, if M is dense, $\|M\|_{1,1}$ is expected to be much larger than $\|M\|_\infty$ so if the psd rank is low compared to the number of rows of M, we expect the bound from Corollary 5.7 to be significantly smaller. For sparse matrices, the same is not true: when applied to the identity matrix for example, Corollary 5.7 can only guarantee factors

of largest eigenvalue at most $\sqrt{k}$, hence the trace is at most $k\sqrt{k}$, a worse guarantee than that obtained directly from Lemma 5.4.

6. Space of factorizations

In this section, we fix a nonnegative matrix M and consider the set of all valid psd factorizations of M as a topological space. In the special case where $\text{rank}(M) = \binom{\text{rank}_{\text{psd}}(M)+1}{2}$, we show in Proposition 6.3 and Corollary 6.4 that this topological space is closely related to the space of all linear images of the psd cone that nest between two polyhedral cones coming from M. An extension of this result to general M is not possible, as seen in Example 6.5. In Examples 6.7 and 6.8, we use this machinery to construct psd factorizations from the linear embedding of the psd cone. Finally in Proposition 6.9, we show that for rank three matrices with psd rank two, the space of psd factorizations is connected. This contrasts with the nonnegative rank case where it is known that the space of nonnegative factorizations can be disconnected for rank three matrices with nonnegative rank three [MSvS03].

For this section, let $M \in \mathbb{R}_+^{p \times q}$ be a nonnegative matrix with psd rank k. As before, we define a *psd factorization* to be a set of matrices $(A_1, \ldots, A_p, B_1, \ldots, B_q) \in (\mathcal{S}^k)^{p+q}$ such that each of the component matrices is psd and $M_{ij} = \langle A_i, B_j \rangle$ for each entry in M. We define the set of all such psd factorizations to be the *space of psd factorizations* associated to M and denote it by $\mathcal{SF}(M)$. Note that this definition only considers matrices whose size is equal to the psd rank of M.

As a warm-up, it is straightforward to see that $\mathcal{SF}(M)$ is closed and infinite. To see that $\mathcal{SF}(M)$ is infinite, simply note that for any psd factorization $(A_1, \ldots, B_q)$ and any matrix $L \in GL(k)$ (the group of invertible $k \times k$ matrices), the matrices $\left(L^T A_1 L, \ldots, L^{-1} B_q (L^{-1})^T\right)$ also form a psd factorization of M. We refer to the set of all such psd factorizations as the *orbit* of $(A_1, \ldots, B_q)$ in $\mathcal{SF}(M)$. In some cases the entire space of psd factorizations is equivalent to a single orbit.

Example 6.1. Let M be the 3×3 derangement matrix from Example 2.6 (i.e. $M_{ii} = 0$ and $M_{ij} = 1$ for $i \neq j$). This matrix has usual rank three and psd rank two as shown by the factorization

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

Let $\mathcal{X}$ denote an arbitrary psd factorization of M. The zero pattern of M implies that the matrices composing $\mathcal{X}$ must all be rank one. Now it is straightforward to see that there exists an invertible matrix such that conjugation by this matrix will send $\mathcal{X}$ to the explicit factorization above. Hence, $\mathcal{SF}(M)$ is composed of a single orbit.

The next proposition gives a geometric picture of the space of factorizations in the special case where $\text{rank}_{\text{psd}}(M) = k$ and $\text{rank}(M) = \binom{k+1}{2}$. Before proceeding, we present a homogenized version of the geometric interpretation of psd rank that was given in Section 3.1, which will be easier to work with in this section. First, we homogenize Definition 3.1:

Definition 6.2. Let P and Q be polyhedral cones with $P \subset Q \subset \mathbb{R}^n$. Let $x_1, \ldots, x_v$ be the extreme rays of P and let $a_j^T x \geq 0$, $(j = 1, \ldots, f)$ be the inequalities defining Q. Then the slack matrix of the pair P, Q, denoted $S_{P,Q}$, is the nonnegative $v \times f$ matrix whose (i,j)th entry is $a_j^T x_i$.

By taking a rank factorization, it is easy to see that any nonnegative matrix M can be viewed as $S_{P,Q}$ for some polyhedral cones P and Q whose dimension is equal to $\text{rank}(M)$. Furthermore, Theorem 3.3 extends straightforwardly to this conic setting: $\text{rank}_{\text{psd}}(S_{P,Q})$ is the smallest integer k for which there exists a subspace $\mathcal{L}$ and a linear map π such that $P \subset \pi(\mathcal{S}_+^k \cap \mathcal{L}) \subset Q$.

In the special case where the cones P and Q come from a matrix M with $\text{rank}_{\text{psd}}(M) = k$ and $\text{rank}(M) = \binom{k+1}{2}$, we can count dimensions to see that the map π is invertible and the subspace $\mathcal{L}$ is all of $\mathcal{S}^k$. If we define $\Delta_k(P, Q)$ to be the space of all linear maps $\pi : \mathcal{S}^k \to \mathbb{R}^{\binom{k+1}{2}}$ such that

$P \subset \pi(\mathcal{S}_+^k) \subset Q$, then $\Delta_k(P,Q)$ is nonempty with M, P, and Q as above. We can actually say much more about $\Delta_k(P,Q)$ in this special case.

Proposition 6.3. *Let* $M \in \mathbb{R}_+^{p \times q}$ *with* $\mathrm{rank}\,(M) = \binom{k+1}{2}$ *and* $\mathrm{rank}_{\mathrm{psd}}(M) = k$. *Fix a rank factorization of* $M = UV$ *where* u_i *is the ith row of* U *and* v_j *is the jth column of* V. *Let* $P = \mathrm{cone}(u_1, \ldots, u_p)$ *and* $Q = \left\{ x \mid v_j^T x \geq 0 \text{ for all } j \right\}$ *be the cones generated by this rank factorization so that* $M = S_{P,Q}$. *Then* $\mathcal{SF}(M)$ *is homeomorphic to* $\Delta_k(P,Q)$.

Proof. Suppose $(A_1, \ldots, B_q)$ is a psd factorization of M. The set $(A_1, \ldots, A_p)$ spans $\mathcal{S}^k$ (else we could find a lower dimensional rank factorization of M), so we can define a linear map π by making $\pi(A_i) = u_i$. This map is well-defined since if $\sum_i \alpha_i A_i$ and $\sum_j \beta_j A_j$ are two representations of the same matrix in $\mathcal{S}^k$, then we have that $\left(\sum_i \alpha_i u_i - \sum_j \beta_j u_j \right)^T V = 0$. Since V has full row rank, this implies that $\sum_i \alpha_i u_i = \sum_j \beta_j u_j$. By the definition of π, it is immediate that $P \subset \pi(\mathcal{S}_+^k)$. Since π has the property that $\langle \pi(L), v_j \rangle = \langle L, B_j \rangle$ for each j, we also have that $\pi(\mathcal{S}_+^k) \subset Q$. Thus, we have defined a map from $\mathcal{SF}(M)$ to $\Delta_k(P,Q)$.

Next, suppose that we have $\pi \in \Delta_k(P,Q)$. Define $A_i = \pi^{-1}(u_i)$ and $B_j = \pi^*(v_j)$ where π^* is the adjoint map. Then $A_i \in \mathcal{S}_+^k$ and $B_j \in (\mathcal{S}_+^k)^* = \mathcal{S}_+^k$ and these matrices form a psd factorization of M. This map is the inverse of the one defined above and both of the maps are continuous. Hence, the spaces are homeomorphic. $\square$

Both of the spaces in the previous proposition permit a natural action by $GL(k)$. The action on $\mathcal{SF}(M)$ was mentioned above when we discussed the orbits of $\mathcal{SF}(M)$. The action on $\Delta_k(P,Q)$ takes the form $g \cdot \pi(L) = \pi(gLg^T)$, i.e. we compose the map π with an automorphism of the psd cone. The homeomorphism in the previous proposition respects these group actions so we can descend to the quotient to see the following.

Corollary 6.4. *Under the same assumptions as the previous proposition, the spaces* $\mathcal{SF}(M)/GL(k)$ *and* $\Delta_k(P,Q)/GL(k)$ *are homeomorphic. Furthermore,* $\Delta_k(P,Q)/GL(k)$ *is homeomorphic to the space of all linear images* C *of* $\mathcal{S}_+^k$ *such that* $P \subset C \subset Q$.

Proof. The first statement is shown by descending to the quotient as discussed prior to the corollary. The second homeomorphism is just given by $[\pi] \mapsto \pi(\mathcal{S}_+^k)$. It is straightforward to check that this map is a well-defined homeomorphism. $\square$

The next example shows that the conclusion of Corollary 6.4 cannot hold for general M.

Example 6.5. Let M be the slack matrix of the regular hexagon, i.e. M is the 6×6 circulant matrix defined by the vector $(0,1,2,2,1,0)$. It was shown in [GRT13a] that M has rank three, psd rank four, and at least two distinct factorization orbits (since there exists a factorization consisting entirely of rank one matrices and another factorization using both rank one and rank two matrices). Since this matrix is a slack matrix of a polytope, however, the cones P and Q must be equal and the only image nested between them must be P itself. Hence, there cannot exist a bijection between factorization orbits and images nested between P and Q.

In the next examples, we apply our machinery to matrices with rank three and psd rank two.

Remark 6.6. In light of Corollary 6.4, we can gain new insight on Example 6.1. By taking the trivial rank factorization of the 3×3 derangement matrix, we obtain the cones

$$P = \mathrm{cone}\left((0,1,1),(1,0,1),(1,1,0)\right) \subset \mathbb{R}_+^3 = Q.$$

Dehomogenizing these cones gives us the two triangles seen in Figure 5. In this dehomogenized picture, linear images of the psd cone correspond to ellipses and it is straightforward to see that there is a unique ellipse that fits between the two triangles. Hence, $\Delta_2(P,Q)/GL(2)$ consists of a single point and by the corollary, the space of psd factorizations is composed of a single orbit.
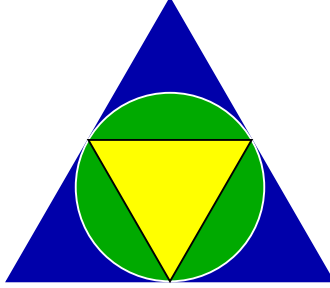
Figure 5. A space of psd factorizations consisting of a single orbit: The yellow and blue triangles correspond to the dehomogenized cones coming from a rank factorization of the 3×3 derangement matrix. The green circle is the unique ellipse that can be nested between the two triangles.

We now show how to apply Proposition 6.3 to find different psd factorizations of a matrix.

Example 6.7. In this example, we consider the following matrix M of rank three (shown along with a rank factorization):

$$\begin{pmatrix} 3 & 3 & 1 & 1 \\ 1 & 3 & 3 & 1 \\ 1 & 1 & 3 & 3 \\ 3 & 1 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 2 & 2 & 2 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \end{pmatrix}.$$

By forming the cones P and Q corresponding to this rank factorization and then dehomogenizing through the plane $\{(1, x_2, x_3) : x_2, x_3 \in \mathbb{R}\}$, we see that P corresponds to the square centered at the origin of side length two and that Q corresponds to the same square scaled by a factor of two (see Figure 6). Now any linear image of $\mathcal{S}_+^2$ corresponds to an ellipse in this dehomogenized picture. So to get a psd factorization of M, we just need to pick an ellipse, figure out a linear image of $\mathcal{S}_+^2$ that corresponds to it, and apply the homeomorphism discussed in Proposition 6.3.

For the circle centered at the origin with radius $\sqrt{2}$, we get the following (where $\alpha = \frac{1}{\sqrt{2}}$):

$$\begin{pmatrix} 1+\alpha & \alpha \\ \alpha & 1-\alpha \end{pmatrix}, \begin{pmatrix} 1-\alpha & \alpha \\ \alpha & 1+\alpha \end{pmatrix}, \begin{pmatrix} 1-\alpha & -\alpha \\ -\alpha & 1+\alpha \end{pmatrix}, \begin{pmatrix} 1+\alpha & -\alpha \\ -\alpha & 1-\alpha \end{pmatrix},$$

$$\begin{pmatrix} 1+\alpha & 0 \\ 0 & 1-\alpha \end{pmatrix}, \begin{pmatrix} 1 & \alpha \\ \alpha & 1 \end{pmatrix}, \begin{pmatrix} 1-\alpha & 0 \\ 0 & 1+\alpha \end{pmatrix}, \begin{pmatrix} 1 & -\alpha \\ -\alpha & 1 \end{pmatrix}.$$

For the ellipse centered at the origin with horizontal axis of length four and vertical axis of length three, we get the factorization:

$$\begin{pmatrix} 5/3 & 1/2 \\ 1/2 & 1/3 \end{pmatrix}, \begin{pmatrix} 1/3 & 1/2 \\ 1/2 & 5/3 \end{pmatrix}, \begin{pmatrix} 1/3 & -1/2 \\ -1/2 & 5/3 \end{pmatrix}, \begin{pmatrix} 5/3 & -1/2 \\ -1/2 & 1/3 \end{pmatrix},$$

$$\begin{pmatrix} 7/4 & 0 \\ 0 & 1/4 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1/4 & 0 \\ 0 & 7/4 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

It is interesting to note how the ranks of the factors change depending on whether the ellipse contacts the vertices of P or the facets of Q. For example, in the second factorization, the matrices corresponding to the columns are rank one exactly when the corresponding facet of the outer square is tight to the ellipse. Of course, this is not a coincidence, but due to how we construct the psd factorization once we know the linear embedding of the psd cone.

In every example of a psd factorization that has been presented so far, either the matrices corresponding to the rows or the matrices corresponding to the columns can be chosen to be rank
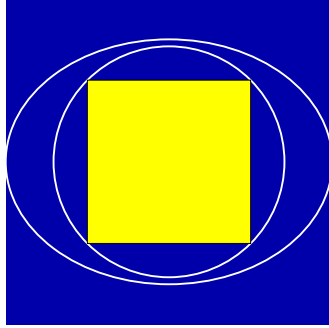
Figure 6. This shows the situation described in Example 6.7. The inner and outer squares correspond to the dehomogenized cones P and Q and any ellipse nested between the two squares corresponds to an orbit of psd factorizations. In the example, we showed factorizations corresponding to both the circle and the ellipse drawn in the figure.

one matrices. Initial attempts to construct a matrix without this property proved fruitless. With the machinery of this section, finding such an example becomes almost trivial.

Example 6.8. In this example, we present a 4×4 matrix with psd rank two such that every 2×2 psd factorization must have a rank two matrix on a row and a rank two matrix on a column. To construct this example, we start with the 3×3 derangement matrix as in Example 6.1, which corresponds to the picture shown in Figure 5. Now we add an extra vertex to the inner triangle and an extra facet to the outer triangle so that neither the new vertex nor the new facet touch the circle, as shown in Figure 7. This corresponds to a new matrix

$$M = \begin{bmatrix} 0 & 1 & 1 & 2 \\ 1 & 0 & 1 & 2 \\ 1 & 1 & 0 & 6 \\ 1 & 1 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -1 \end{bmatrix}.$$

The same circle as before is still the unique ellipse nested between the two polytopes so the space of factorizations consists of a single orbit. When we construct a psd factorization in this orbit, the matrix corresponding to the new vertex must lie in the interior of $\mathcal{S}_+^2$ and the matrix corresponding to the new facet must lie in the interior of $(\mathcal{S}_+^2)^*$. Hence, they must have rank two. Such a factorization may be obtained by augmenting our previous factorization for the derangement matrix with the matrix $\begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}$ for the new vertex and the matrix $\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ for the new facet.

For the special case where M has rank three and we are considering 2×2 psd factorizations, we can show that the space of factorization orbits is connected.

Proposition 6.9. *Let M be a nonnegative $p \times q$ matrix with psd rank two and usual rank three. Then $\mathcal{SF}(M)/GL(2)$ is connected.*

Proof. Let P and Q be the cones in $\mathbb{R}^3$ arising from a rank factorization as above. By Corollary 6.4, it is enough to show that $\Delta_2(P,Q)/GL(2)$ is connected. To do this, we will dehomogenize the cones so that we can work with polytopes and ellipses.

The cone Q must be pointed since it was formed from a full-rank matrix. Thus, we can find an affine hyperplane such that the dehomogenization of Q through this hyperplane is bounded. We dehomogenize through this hyperplane to get polytopes $\widetilde{P} \subset \widetilde{Q}$. Any element of $\Delta_2(P,Q)/GL(2)$ corresponds to an ellipse nested between $\widetilde{P}$ and $\widetilde{Q}$. Thus, to finish the proof, it is enough to show that any two ellipses that are nested between $\widetilde{P}$ and $\widetilde{Q}$ can be connected by a path of ellipses.
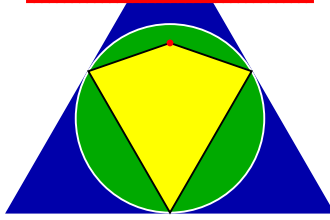
Figure 7. Here we take the picture corresponding to the 3×3 derangement matrix and add an additional facet to the outer polytope and an additional vertex to the inner polytope (shown in red). Since the new facet and vertex are not incident to the boundary of the linear embedding of $\mathcal{S}_+^2$, their corresponding matrices in the psd factorization must have full rank.

Suppose E_0 and E_1 are ellipses nested between the two polytopes. Then there exist quadratic polynomials q_0 and q_1 such that $E_i = \{x \mid q_i(x) \geq 0\}$. For $t \in [0, 1]$, define a quadratic polynomial $q_t = (1-t)q_0 + tq_1$ and the corresponding ellipse E_t. Since q_0 and q_1 are nonnegative on the points of $\widetilde{P}$, so is q_t and we have that $\widetilde{P} \subset E_t$. Similarly, since q_0 and q_1 are negative on $(E_0 \cup E_1)^c$ (where $(E_0 \cup E_1)^c$ denotes the complement of $E_0 \cup E_1$ in the ambient space), we have that $E_t \subset E_0 \cup E_1 \subset \widetilde{Q}$. Thus, E_t gives the desired path of ellipses. $\qquad\square$

We are not sure if Proposition 6.9 extends to matrices M with $\operatorname{rank}_{\mathrm{psd}}(M) = k$ and $\operatorname{rank}(M) = \binom{k+1}{2}$ for $k > 2$. The proof in the $k = 2$ case relied on the fact that bounded spectrahedra in $\mathcal{S}_+^2$ can be represented by a single polynomial inequality. Higher dimensional spectrahedra require several polynomial inequalities and it is not clear if the proof can be extended to this setting. Searching for a counterexample has also been difficult, since the next case involves linear images of $\mathcal{S}_+^3$ nested between six-dimensional cones.

7. Symmetric factorizations

In this section we consider nonnegative matrices that admit psd factorizations where the row and column factors A_i and B_j are required to be the same. Matrices that admit such a factorization are called *completely psd* [LP13], by analogy to *completely positive matrices* [BSM03]:

Definition 7.1. A symmetric matrix $M \in \mathcal{S}^n$ is called *completely psd* if there exists $k \in \mathbb{N}$ and $A_1, \ldots, A_n \in \mathcal{S}_+^k$ such that $M_{ij} = \langle A_i, A_j \rangle$ for all $i, j = 1, \ldots, n$. Such a factorization of M is called a *completely-psd factorization*.

The set of matrices that are *completely psd* forms a convex cone in $\mathcal{S}^n$; we denote this cone by $\mathrm{CP}_{\mathrm{psd}}^n$. Completely psd matrices find applications in quantum information theory for the computation of certain quantum graph parameters [LP13].

Recall that a matrix M is called *completely positive* if there exist *nonnegative* vectors $a_i \in \mathbb{R}_+^d$ such that $M_{ij} = \langle a_i, a_j \rangle$ for all $i, j = 1, \ldots, n$. The convex cone of $n \times n$ completely positive matrices is denoted by CP^n. By representing a nonnegative vector as a diagonal psd matrix it is easy to see that any completely positive matrix is also completely psd, i.e., $\mathrm{CP}^n \subseteq \mathrm{CP}_{\mathrm{psd}}^n$. It is also clear from the definition that any completely psd matrix M is necessarily nonnegative and positive semidefinite. Thus we have the inclusion

$$\mathrm{CP}^n \subseteq \mathrm{CP}_{\mathrm{psd}}^n \subseteq \mathrm{DN}^n \tag{19}$$

where DN^n is the cone of matrices that are nonnegative and positive semidefinite (also called *doubly nonnegative matrices*).

When $n \leq 4$ it is known that $\mathrm{CP}^n = \mathrm{DN}^n$ and thus the inclusions (19) are all equalities. It is known that for $n = 5$ the two inclusions (19) are strict [LP13, FW10]:

- To show that $\mathrm{CP}^5 \neq \mathrm{CP}^5_{\mathrm{psd}}$, one can consider the 5×5 matrix M defined by:

$$M_{ij} = \cos^2\left(\frac{4\pi}{5}(i-j)\right), \quad i,j = 1,\ldots,5.$$

The matrix M is completely psd since it admits the factorization $M_{ij} = \langle a_i a_i^T, a_j a_j^T \rangle$ where $a_i = (\cos(4\pi i/5), \sin(4\pi i/5)) \in \mathbb{R}^2$. On the other hand $M \notin \mathrm{CP}^5$ since $\langle H, M \rangle < 0$ where H is an element of the dual cone $(\mathrm{CP}^5)^*$ known as the *Horn form* (see e.g., [Dür10]):

$$H = \begin{bmatrix} 1 & -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 & 1 \end{bmatrix}.$$

- In [LP13, FW10] it was shown that any matrix $M \in \mathrm{DN}^5$ such that rank $M = 3$ and whose sparsity pattern is the 5-cycle is *not* completely psd. One can easily exhibit such a matrix; thus this shows that $\mathrm{CP}^5_{\mathrm{psd}} \neq \mathrm{DN}^5$ (cf. [LP13, Eq. 3.3] or [FW10, Section 3] for concrete examples).

Several fundamental questions are open concerning the cone $\mathrm{CP}^n_{\mathrm{psd}}$. One important question is to know whether the cone $\mathrm{CP}^n_{\mathrm{psd}}$ is closed. For a completely-psd matrix M one can define the *cpsd-rank* of M as the smallest integer k for which M admits a completely-psd factorization of size k. The closedness question concerning $\mathrm{CP}^n_{\mathrm{psd}}$ is related to the question of whether the cpsd-rank of a matrix $M \in \mathrm{CP}^n_{\mathrm{psd}}$ can be bounded from above by a function that depends only on n. We refer the reader to [BLP15] for recent progress concerning this question.

8. Open questions

8.1. **Psd rank of special matrices.** There are very few matrices for which one can determine the psd rank precisely, and when we can, it is usually in very special cases where our coarse bounds such as the square root rank or the trivial rank bound turn out to be sufficient. As such, any new insight on determining the psd rank of concrete matrices will provide new tools to analyze psd rank in general. In that spirit, we propose a few more or less concrete matrices whose psd ranks we would like to know, and would provide a starting point for this program.

Problem 8.1. Consider the 10 by 10 matrix A whose rows and columns are indexed by subsets of $\{1,\ldots,5\}$ of size 2 and 3 respectively, and defined by $A_{I,J} = |I \cap J|$. What is the psd rank of A?

One geometric interpretation for the matrix A, is to take the 10-vertices of a *rectified* 5-*cell* inscribed in a 3-sphere and take the generalized slack matrix with respect to the tangents at the 10-points. Since the unit ball in $\mathbb{R}^4$ has a semidefinite representation of size 4, we know that the psd rank of A is at most 4. The usual rank of A being 5, we know that its psd rank must be at least 3. If one can show that it is 4 it would prove that there is no smaller semidefinite representation of the 3-sphere. A more general version of this problem can be attained by allowing different set sizes and different numbers of elements.

Problem 8.2. Let $A(n)$ be the matrix whose rows and columns are indexed by subsets of $\{1,\ldots,n\}$ of size $\lfloor n/2 \rfloor$ and $\lceil n/2 \rceil$ respectively, and defined by $A_{I,J} = |I \cap J|$. What is the psd rank of $A(n)$?

The matrices $A(n)$ again have an interpretation in terms of an inscribed polytope in the $(n-2)$-sphere. In general, we know only that their psd rank is at most $2\lceil \sqrt{n} \rceil$ and at least $\frac{1}{2}\sqrt{1+8n} - \frac{1}{2}$. These matrices are very special, in the sense that they have a rich symmetry structure. In fact they are related to the Johnson scheme $J(n,k)$, with $k = \lfloor n/2 \rfloor$, since if $A_i, i = 0,\ldots,k$ are the generators of the associated algebra, $A(n) = \sum_{i=0}^{k} iA_i$. This justifies posing a more ambitious and less well-defined question.

Problem 8.3. Let A be a matrix in the Bose-Mesner algebra of the Johnson scheme (or any other association scheme). Can one exploit the symmetry in these matrices to provide non-trivial bounds on their psd ranks?

8.2. **Geometry of psd rank.** In Section 6, we defined $\mathcal{SF}(M)$, the space of psd factorizations of size k of a nonnegative matrix M of rank $\binom{k+1}{2}$ and psd rank k. We showed that the quotient space $\mathcal{SF}(M)/GL(k)$ is connected when $k = 2$. There is no reason to believe that such a connectivity result holds in general.

Problem 8.4. If M is a nonnegative matrix with $\mathrm{rank}_{\mathrm{psd}}(M) = k$ and $\mathrm{rank}(M) = \binom{k+1}{2}$, then is the space of factorization orbits $\mathcal{SF}(M)/GL(k)$ always connected?

The analogous question for nonnegative rank was studied in [MSvS03]. The authors showed examples where M has rank three and the space of nonnegative factorization orbits of size three is disconnected (here the orbit is generated by the permutation group S_3, acting by permuting the entries of the nonnegative factorization). A natural question to ask is whether these disconnected factorizations remain disconnected when we embed them in the space of 3×3 psd factorization orbits. Embedding these factorizations is straightforward (just make the nonnegative vectors into diagonal matrices), but doing so changes the setting of Problem 8.4 since we are now considering the possible 3×3 psd factorizations of a matrix M with rank three and psd rank two. We are able to show that the factorizations that were disconnected in the nonnegative case become connected when we embed them in the set of 3×3 psd factorizations. However, this does not settle Problem 8.4.

8.3. **Complexity and algorithms.** Several complexity questions concerning psd rank are open.

Problem 8.5. For a fixed constant k, is it NP-hard to decide if $\mathrm{rank}_{\mathrm{psd}}(M) \leq k$? For example, is it NP-hard to decide if $\mathrm{rank}_{\mathrm{psd}}(M) = 3$?

The equivalent question for the nonnegative rank was shown to be decidable in polynomial time [AGKM12]. The complexity of the algorithm proposed in [AGKM12] is polynomial in p, q (the dimensions of the matrix) but doubly-exponential in k; this was later improved to a singly-exponential algorithm in k in [Moi13]. Both algorithms express the problem of finding a nonnegative factorization of size k as a semialgebraic system which is then solved using quantifier elimination algorithms [Ren92]. Since the complexity of quantifier elimination algorithms has an exponential dependence on the number of variables, one has to make sure that the number of variables in the semialgebraic system is independent of p, q (the size of the nonnegative matrix) in order to get an algorithm that is polynomial in p, q. The semialgebraic formulations proposed in [Moi13] and [AGKM12] rely on the key fact that the solution of any linear program is a rational function of the data. This fact however is far from being true in semidefinite programming [NRS10], and this is one major obstacle in extending these algorithms to the psd rank case.

Another important question is to know the complexity of computing the psd rank (here k is not a fixed constant anymore). In Section 2 we saw that the psd rank of a matrix $M \in \mathbb{R}_+^{p \times q}$ satisfies the following inequalities:

$$(20) \qquad \mathrm{rank}(M) \leq \binom{\mathrm{rank}_{\mathrm{psd}}(M) + 1}{2} \quad \text{and} \quad \mathrm{rank}_{\mathrm{psd}}(M) \leq \min(p, q).$$

It is already not known whether deciding if any of these inequalities is tight can be achieved in polynomial-time.

Problem 8.6. Show that the psd rank is NP-hard to compute. More specifically show that the problems of deciding whether inequalities (20) are tight are NP-hard.

Note that in [GRT13b, Theorem 4.6], an algorithm is proposed to decide whether the first inequality of (20) is tight, however the complexity of the algorithm has an exponential dependence

on rank (M) (the complexity of the algorithm is polynomial when rank (M) is a fixed parameter). For the case of the nonnegative rank, it was shown [Vav09] that the problem of deciding whether rank $_+(M)$ = rank (M) is NP-hard. Another question is to find efficient algorithms to approximate the psd rank.

Problem 8.7. Is there a polynomial time approximation algorithm for psd rank (or nonnegative rank) that will find a factorization of size at most $\alpha \cdot \text{rank}_{\text{psd}}$ (or $\alpha \cdot \text{rank}_+$) for some constant α?

8.4. **Slack matrices.** The psd rank of slack matrices of polytopes has been of special interest due to its applications to semidefinite lifts of polytopes as described in Section 3. A 0/1 polytope is one whose vertices are 0/1 vectors. It was shown in [BDP13] that as n grows, most 0/1-polytopes of dimension n must have psd rank exponential in n. The proof works via a counting argument and does not identify specific families of polytopes with such exponential psd rank. However in the recent paper [LRS14] a super-polynomial lower bound for the psd rank of the cut polytope, TSP polytope and stable set polytope was proved.

Another interesting question concerning slack matrices of polytopes is whether there exists a family of polytopes that can be expressed by liftings to small psd cones but require large polyhedral lifts. In the language of nonnegative and psd rank, this leads to the following question.

Problem 8.8. Find a family of polytopes that exhibits a large (e.g. exponential) gap between its psd and nonnegative ranks.

In the paper [FSP15] it was shown that there is a family of polytopes $P_d \subset \mathbb{R}^{2d}$ such that $\text{rank}_{\text{psd}}(P_d) \leq O(d \log d)$ and $\text{rank}_+(P_d) \geq \Omega(d^2)$. It is still an open question to construct a family of polytopes where the ratio of nonnegative rank to psd rank grows faster than $d/\log d$.

There are several families of polytopes with exponentially many facets (in the dimension of the polytope) that can be expressed by small polyhedral lifts or small psd lifts. The stable set polytope of a perfect graph on n vertices has psd rank $n + 1$ [GRT13a, Theorem 4.12]. On the other hand, Yannakakis [Yan91, Theorem 5] proved that its nonnegative rank is at most $O(n^{\log n})$. Is nonnegative rank of the stable set polytope of a perfect graph polynomial in n ? Even if it is not, the gap between psd rank and nonnegative rank would not be dramatic for this family. Polytopes with a large gap between psd rank and nonnegative rank as in Problem 8.8 would show that semidefinite programming is truly more powerful than linear programming in expressing linear optimization problems over polytopes.

Goemans observed in [Goe14] that the nonnegative rank of a slack matrix is bounded below by $\log(v)$ where v is the number of vertices of the polytope. The same argument shows that $\log(\sum f_i)$ is a lower bound to nonnegative rank of a polytope where f_i is the number of i-dimensional faces of the polytope. It would be interesting to know if there are similar lower bounds for the psd rank of a polytope coming from the combinatorial structure of the polytope.

Problem 8.9. Develop good bounds on the psd rank of a polytope that use information about its combinatorial/facial structure.

We remark that Corollary 4.13 provides a lower bound to the psd rank of a polytope in terms of the number of facets of the polytope. However, the unknown constants in the bound prevent us from using it to understand specific polytopes. A result in the spirit of Problem 8.9 from [GRT13b] is that generic polytopes of dimension n with v vertices have psd rank at least $(nv)^{1/4}$.

8.5. **Completely positive semidefinite matrices.** The following question about symmetric psd factorizations was raised in Section 7.

Problem 8.10. Let CP_{psd}^n be the set of $n \times n$ matrices M that admit a factorization of the form

$$(21) \qquad\qquad\qquad\qquad M_{ij} = \langle A_i, A_j \rangle$$

where $A_1, \ldots, A_n$ are psd matrices. Does there exist a function $f(n)$ such that any matrix $M \in \mathrm{CP}^n_{\mathrm{psd}}$ admits a factorization of the form (21) where the factors $A_1, \ldots, A_n$ have size at most $f(n)$?

References

[AGKM12] S. Arora, R. Ge, R. Kannan, and A. Moitra. Computing a nonnegative matrix factorization–provably. In *Proceedings of the 44th Symposium on Theory of Computing (STOC)*, pages 145–162. ACM, 2012. 31

[Bar01] A. Barvinok. A remark on the rank of positive semidefinite matrices subject to affine constraints. *Discrete & Computational Geometry*, 25(1):23–31, 2001. 22

[Bar12] A. Barvinok. Approximations of convex bodies by polytopes and by projections of spectrahedra. arXiv preprint arXiv:1204:0471, 2012. 17, 18

[BCR11] C. Bocci, E. Carlini, and F. Rapallo. Perturbation of matrices and nonnegative rank with a view toward statistical models. *SIAM Journal on Matrix Analysis and Applications*, 32(4):1500–1512, 2011. 8

[BDP13] J. Briët, D. Dadush, and S. Pokutta. On the existence of 0/1 polytopes with high semidefinite extension complexity. In *Algorithms, ESA 2013*, volume 8125 of *Lecture Notes in Computer Science*, pages 217–228. Springer Berlin Heidelberg, 2013. 21, 23, 24, 32

[BLP15] S. Burgdorf, M. Laurent, and T. Piovesan. On the closure of the completely positive semidefinite cone and linear approximations to quantum colorings. *arXiv preprint arXiv:1502.02842*, 2015. 30

[BPA+08] N. Brunner, S. Pironio, A. Acin, N. Gisin, A.A. Méthot, and V. Scarani. Testing the dimension of Hilbert spaces. *Physical Review Letters*, 100(21):210503, 2008. 13

[BPT12] G. Blekherman, P. A. Parrilo, and R. Thomas, editors. *Semidefinite Optimization and Convex Algebraic Geometry*, volume 13 of *MOS-SIAM Series on Optimization*. SIAM, 2012. 2

[BSM03] A. Berman and N. Shaked-Monderer. *Completely Positive Matrices*. World Scientific Pub Co Inc, 2003. 29

[BV04] S.P. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. 13

[CR93] J.E. Cohen and U.G. Rothblum. Nonnegative ranks, decompositions, and factorizations of nonnegative matrices. *Linear Algebra and its Applications*, 190:149–168, 1993. 1, 3, 4, 12, 13

[Dür10] M. Dür. Copositive programming–a survey. *Recent Advances in Optimization and its Applications in Engineering*, pages 3–20, 2010. 30

[FKPT13] S. Fiorini, V. Kaibel, K. Pashkovich, and D. O. Theis. Combinatorial bounds on nonnegative rank and extended formulations. *Discrete Mathematics*, 313(1):67 – 83, 2013. 19

[FMP+12] S. Fiorini, S. Massar, S. Pokutta, H.R. Tiwary, and R. de Wolf. Linear vs. semidefinite extended formulations: Exponential separation and strong lower bounds. In *Proceedings of the Forty-fourth Annual ACM Symposium on Theory of Computing*, STOC '12, pages 95–106. ACM, 2012. 2, 13

[FSP15] H. Fawzi, J. Saunderson, and P.A. Parrilo. Sparse sum-of-squares certificates on finite abelian groups. *arXiv preprint arXiv:1503.01207*, 2015. 32

[FW10] P.E. Frenkel and M. Weiner. On vector configurations that can be realized in the cone of positive matrices. *arXiv preprint arXiv:1004.0686*, 2010. 29, 30

[GFR14] J. Gouveia, H. Fawzi, and R. Z. Robinson. Rational and real positive semidefinite rank can be different. *arXiv preprint arXiv:1404.4864*, 2014. 5

[GG12] N. Gillis and F. Glineur. On the geometric interpretation of the nonnegative rank. *Linear Algebra and its Applications*, 437(11):2685–2712, 2012. 12

[GJ79] M. R. Garey and D. S. Johnson. *Computers and Intractability: A guide to the theory of NP-completeness*. W. H. Freeman and Company, 1979. 18

[Goe14] M. Goemans. Smallest compact formulation for the permutahedron. *Mathematical Programming Series B*, pages 1–7, 2014. 32

[GPT13] J. Gouveia, P.A. Parrilo, and R.R. Thomas. Lifts of convex sets and cone factorizations. *Mathematics of Operations Research*, 38(2):248–264, 2013. 2, 3, 9, 19, 20

[GRT13a] J. Gouveia, R. Z. Robinson, and R. R. Thomas. Polytopes of minimum positive semidefinite rank. *Discrete & Computational Geometry*, 50(3):679–699, 2013. 11, 19, 26, 32

[GRT13b] J. Gouveia, R. Z. Robinson, and R. R. Thomas. Worst-case results for positive semidefinite rank. *arXiv preprint arXiv:1305.4600*, 2013. 9, 10, 11, 13, 21, 31, 32

[Hru12] P. Hrubeš. On the nonnegative rank of distance matrices. *Information Processing Letters*, 112:457–461, 2012. 21

[Jol02]   I. Jolliffe. *Principal component analysis*. Springer Series in Statistics. Springer, 2nd edition, 2002. 1

[JSWZ13]   R. Jain, Y. Shi, Z. Wei, and S. Zhang. Efficient protocols for generating bipartite classical distributions
            and quantum states. *IEEE Transactions on Information Theory*, 59(8):5171–5178, 2013. 2, 13, 15

[Kal63]   R. E. Kalman. Mathematical description of linear dynamical systems. *Journal of the Society for Industrial
            & Applied Mathematics, Series A: Control*, 1(2):152–192, 1963. 1

[LMSS07]   N. Linial, S. Mendelson, G. Schechtman, and A. Shraibman. Complexity measures of sign matrices.
            Combinatorica, 27(4):439–463, 2007. 1, 24

[LP13]   M. Laurent and T. Piovesan. Conic approach to quantum graph parameters using linear optimization
            over the completely positive semidefinite cone. *arXiv preprint arXiv:1312.6643*, 2013. 29, 30

[LRS14]   J.R. Lee, P. Raghavendra, and D. Steurer. Lower bounds on the size of semidefinite programming relax-
            ations. *arXiv preprint arXiv:1411.6317*, 2014. 32

[LS99]   D. D. Lee and H. S. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*,
            401(6755):788–791, 1999. 1

[LS09]   N. Linial and A. Shraibman. Lower bounds in communication complexity based on factorization norms.
            Random Structures & Algorithms, 34(3):368–394, 2009. 1

[LT12]   T. Lee and D. O. Theis. Support-based lower bounds for the positive semidefinite rank of a nonnegative
            matrix. arXiv preprint arXiv:1203.3961, 2012. 19, 22

[LWdW14]   T. Lee, Z. Wei, and R. de Wolf. Some upper and lower bounds on psd rank. In preparation, 2014. 6, 8

[Moi13]   A. Moitra. An almost optimal algorithm for computing nonnegative rank. In *SODA*, pages 1454–1464.
            SIAM, 2013. 31

[Moo81]   B. Moore. Principal component analysis in linear systems: Controllability, observability, and model re-
            duction. *IEEE Transactions on Automatic Control*, 26(1):17–32, 1981. 1

[MSvS03]   D. Mond, J. Smith, and D. van Straten. Stochastic factorizations, sandwiched simplices and the topology
            of the space of explanations. *Royal Soc. Proc., Math. Phys. and Engineering Sciences*, 459(2039):2821–
            2845, 2003. 1, 25, 31

[NRS10]   J. Nie, K. Ranestad, and B. Sturmfels. The algebraic degree of semidefinite programming. *Mathematical
            Programming*, 122(2):379–405, 2010. 31

[Pas12]   K. Pashkovich. *Extended formulations for combinatorial polytopes*. PhD thesis, Otto-von-Guericke-
            Universität Magdeburg, 2012. 10

[Pat98]   G. Pataki. On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal
            eigenvalues. *Mathematics of operations research*, 23(2):339–358, 1998. 22

[Ren92]   J. Renegar. On the computational complexity and geometry of the first-order theory of the reals. Part I:
            Introduction. preliminaries. The geometry of semi-algebraic sets. The decision problem for the existential
            theory of the reals. *Journal of Symbolic Computation*, 13(3):255 – 299, 1992. 19, 31

[Ren06]   J. Renegar. Hyperbolic programs, and their derivative relaxations. *Foundations of Computational Math-
            ematics*, 6(1):59–79, 2006. 19

[Sch12]   C. Scheiderer. Semidefinite representation for convex hulls of real algebraic curves. arXiv preprint
            arXiv:1208.3865, 2012. 20

[Vav09]   S. A. Vavasis. On the complexity of nonnegative matrix factorization. *SIAM Journal on Optimization*,
            20(3):1364–1377, 2009. 32

[VB96]   L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, 1996. 9

[WCD08]   S. Wehner, M. Christandl, and A.C. Doherty. Lower bound on the dimension of a quantum system given
            measured data. *Physical Review A*, 78(6):062112, 2008. 13

[Yan91]   M. Yannakakis. Expressing combinatorial optimization problems by linear programs. *J. Comput. System
            Sci.*, 43(3):441–466, 1991. 2, 32

Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology, Cambridge, MA 02139, USA

E-mail address: `hfawzi@mit.edu`

CMUC, Department of Mathematics, University of Coimbra, 3001-454 Coimbra, Portugal

E-mail address: `jgouveia@mat.uc.pt`

Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology, Cambridge, MA 02139, USA

E-mail address: `parrilo@mit.edu`

Department of Mathematics, University of Washington, Box 354350, Seattle, WA 98195, USA

E-mail address: `rzr@uw.edu`

Department of Mathematics, University of Washington, Box 354350, Seattle, WA 98195, USA

E-mail address: `rrthomas@uw.edu`