Building a completely positive factorization

Immanuel M. Bomze

ISOR and VCOR, University of Vienna, Austria

Abstract

Using a bordering approach, and building upon an already known factorization of a principal block, we establish sufficient conditions under which we can extend this factorization to the full matrix. Simulations show that the approach is promising also in higher dimensions.

Key words. Copositive optimization, rounding problem.

1 Introduction

A symmetric matrix is called *completely positive,* if it admits a symmetric rectangular matrix factorization with no negative entries; see [8] which focuses on linear-algebraic and graph theoretic properties of this matrix class. The concept – the notion was probably coined by [33], see also [23] – had its origins from applications in combinatorics (block designs, [33]). Further fields of application include physics, biology and statistics (Markovian models of DNA evolution, [38]), project management (stochastic and robust optimization, [44]), and economic modeling (see [30]), and in recent years optimization applications became increasingly important.

We need some notation. By $\mathbb{R}^n$ we denote n-dimensional Euclidean space, by $\mathbb{R}^n_+$ its positive orthant, (column) vectors $\mathbf{v} \in \mathbb{R}^n$ are always in boldface while $\mathbf{v}^\top$ denotes their transpose (rows). The zero matrix of appropriate size is always denoted by O, and $A \leq O$ for a matrix A of the same size means that no entry A_{ij} is positive, while $A \leq B$ means $A - B \leq O$. I_d denotes the $d \times d$ identity matrix, with its ith column $\mathbf{e}_i \in \mathbb{R}^d$. For a scalar t, we denote by $t_+ = \max\{0, t\}$ while for a matrix A we put $A_+ = [(A_{ij})_+]_{i,j}$. Further, we designate by $A^{\bullet 2} = [(A_{ij})^2]_{i,j}$ the Hadamard square of A. The cone of all symmetric positive-semidefinite matrices of some fixed order is denoted by $\mathcal{P}$, and the cone of all symmetric matrices with no negative entries by $\mathcal{N}$. Finally, let $\sqrt{X} \in \mathcal{P}$ denote the symmetric square-root of a matrix $X \in \mathcal{P}$. Note that even if $X \geq O$, we may have negative entries in $\sqrt{X}$. If however $\sqrt{X} \geq O$, then X belongs to the the cone $\mathcal{C}$ of all (symmetric) completely positive matrices

$$\mathcal{C} = \left\{ X \in \mathcal{P} : X = F^\top F \text{ for some possibly rectangular matrix } F \geq O \right\}.$$

We call $X = F^\top F$ a *completely positive (cp) factorization.* The minimal number of rows in F yielding a cp factorization of X, the so-called *cp-rank* of X, can exceed the order n of X but is bounded by $\binom{n+1}{2} - 4 \sim \frac{n^2}{2}$, and this bound is asymptotically tight [15, 49], refuting a conjecture suggesting $n^2/4$ published 20 years ago [27].

In this paper we aim at obtaining a cp factorization of a symmetric $(n+1) \times (n+1)$ matrix $Y = H^\top H$ by a bordering approach: we assume that we know a cp factorization for a principal $n \times n$ submatrix $X = F^\top F$ of Y, and derive sufficient conditions under which we can specify a suitable factor H. This is the content of Sections 3 and 7. These sufficient conditions generalize and complement previous investigations of the same kind in [46], leading to a structurally different cp factorization (essentially, the role of a lower block-triangular factor in [46] is now played by an upper block-triangular one). In Section 4, we take an optimization-inspired approach, leading to LP- or QP-based relaxations of the main property, in order to systematically find out whether or not this new sufficient condition is satisfied. This approach also may enable us to efficiently search for constellations (i.e., selecting the bordering row) where the condition is met. A small empirical study provided in Section 5 shows that our approach is more promising. In Section 6 we show that our approach indeed suffices to obtain a cp factorization for all completely positive 3×3 matrices, establishing in an elementary way the well-known fact that the cp-rank of these does not exceed

three. This has been known before, but our approach seems less involved than the previous arguments. Inspired by this, we move on in Section 7 to discuss extensions in higher dimensions.

2 Motivation and preprocessing

Since the explicit introduction of *copositive optimization* (or copositive programming) by [11, 45], we observe a rapid evolution of this field. One reason for the success is culminating in the important paper [18] where it is shown that every mixed-binary (fractional) quadratic optimization problem can be written as a copositive optimization problem, which is a linear optimization problem over the cone $\mathcal{C}$ subject to linear constraints, see [1, 2, 13, 18], and recently many similar copositive representation results followed. For some surveys, we refer to [10, 12, 19, 28].

The terminology copositive optimization has its justification as the dual cone of $\mathcal{C}$ coincides with the cone of copositive matrices of the same order. Recall that a symmetric $n \times n$ matrix is said to be *copositive* if it generates a quadratic form taking no negative values over the positive orthant $\mathbb{R}^n_+$. The usual conic approximation algorithms for solving a copositive (or completely positive) optimization problem use (subsets of) the outer approximation $\mathcal{P} \cap \mathcal{N}$ of $\mathcal{C}$. However, often copositive optimization problems are used to reformulate hard (mixed-binary) quadratic optimization problems which in turn may encode combinatorial optimization problems (see [11, 18, 44, 45] and references therein). The optimal solution of the latter is encoded by an $r \times n$ matrix F with no negative entries, in a cp factorization $X = F^\top F \in \mathcal{C}$.

Once we arrive at a solution $X \in \mathcal{P} \cap \mathcal{N}$ of the relaxation, we should try to find this H, not only to show that in this instance, the relaxation is exact, but also to retrieve the solution of the original problem (the so-called rounding procedure). Very few recent papers deal with approximating $\mathcal{C}$ from within: a theoretical characterization of interior points of $\mathcal{C}$ is presented in [24, 29], while algorithmic aspects of factorization are the focus of [35].

Suppose that we want to find an explicit factorization of an $(n+1) \times (n+1)$ matrix Y which we scale such that

$$Y = \begin{bmatrix} 1 & \mathbf{x}^\top \\ \mathbf{x} & X \end{bmatrix}, \tag{1}$$

building upon an already known cp factorization of the principal block X. Note that as $\mathcal{C} \subseteq \mathcal{P}$, we can immediately spare our efforts if one diagonal entry of Y is negative. Similarly, positive-semidefiniteness of Y implies that a zero diagonal entry forces the whole row and column to be zero, in which case we can remove it and continue with a smaller principal submatrix with strictly positive diagonal elements. In the end, we just have to enlarge the factor H by suitably adding zero columns, to obtain the original Y. Hence we may and do assume $Y_{ii} > 0$ for all i. Next, we observe that with any positive-definite diagonal matrix Δ and any factorization $Y = H^\top H$, we get another one of $\Delta Y \Delta = (H\Delta)^\top (H\Delta)$

of the same size. As also Δ^{-1} is positive-definite, we may use $\Delta_{ii} = Y_{ii}^{-1/2} > 0$ and concentrate on the case where the diagonal of Y contains only unity entries. This is the starting point of (1).

Next we proceed as in [25] which offers an algorithmic procedure to obtain a minimal cp factorization, applied to Y with a special sparsity pattern. As a preliminary step, we check the necessary condition $Y \in \mathcal{P} \cap \mathcal{N}$; since $X \in \mathcal{C} \subseteq \mathcal{P} \cap \mathcal{N}$, we merely must check $\mathbf{x} \in \mathbb{R}_+^n$ to ensure $Y \in \mathcal{N}$; and $X - \mathbf{x}\mathbf{x}^\top \in \mathcal{P}$ to ensure $Y \in \mathcal{P}$, by Schur complementation. Note that positive-semidefiniteness of $Y \geq O$ and Y_{ii} implies $Y_{ij} \in [0, 1]$ for all i, j and also any possible cp factorization matrix H (i.e. satisfying $H^\top H = Y$) must have entries between zero and one.

Suppose that $X\mathbf{v} = \mathbf{o}$ but $\mathbf{v}^\top \mathbf{x} \neq 0$; then we would arrive at the contradiction $0 \leq \mathbf{v}^\top (X - \mathbf{x}\mathbf{x}^\top)\mathbf{v} = -(\mathbf{v}^\top \mathbf{x})^2 < 0$. So we may and do assume in the sequel that $\mathbf{x} \in (\ker X)^\perp$ or equivalently, that

$$XX^+\mathbf{x} = \mathbf{x}, \tag{2}$$

where X^+ denotes the *Moore/Penrose generalized inverse (MPGI)* of any matrix.

3 Various factorization strategies

The cp factorization problem has received considerable attention as a special (symmetric) variant of the nowadays heavily researched *nonnegative matrix factorization (NMF)* problem. For a recent survey on NMF see, e.g. [52]. In this context, the cp factorization problem is also addressed as *Symmetric NMF*, e.g. in [34] where a simple parallelizable iterative procedure is proposed which is shown to converge to a stationary point (not the global solution) of the (non-convex) least squares approximation problem $\min_H \|Y - H^\top H\|$, with application to probabilistic clustering in large instances. This article fits into the tradition of convergence analysis in Matheuristics, as performed masterly in [31]; see also [32]. In contrast to these approaches, we focus on finite, not on iterative methods, although possibly employing iterative solutions to (easy) subproblems. For many other approaches, we refer to [3, 6, 9, 7, 25, 47, 48, 51] and the recent review [5], to cite just a few. Another recent paper [53] deals with algorithmic strategies for factorization based upon conic optimization, for random instances of relatively moderate order $n \geq 10$.

The original cp factorization problem can also be seen directly: to obtain $Y = H^\top H$ with an $s \times (n+1)$ matrix $H \geq O$, solve a system of $\binom{n+2}{2}$ quadratic equations in $s(n+1)$ nonnegative variables. As detailed above, we can bound only $s \leq \binom{n+2}{2} - 4$ and hence need in the worst case $\frac{1}{2}(n^3 + 4n^2 - 3n + 6)$ nonnegative variables, which from an algorithmic perspective practically prohibitive even for small instances. As already announced, we here assume that we know the cp factorization

$$X = F^\top F \qquad \text{where} \quad F \quad \text{is an } r \times n\text{-matrix with no negative entries.} \tag{3}$$

Since $X_{ij} \in [0,1]$, also $F_{ij} \in [0,1]$ for all entries. As mentioned earlier, cp-rank(X) is defined as the smallest r such that (3) holds. Hence given any F satisfying (3), we always have cp-rank$(X) \le r$, and since the cp-rank may also be of order n^2, we can have $r > n$ if $n > 4$. See [14] for recent bounds in the range $n \le 11$ and [15] for all larger n.

Recall that for the $n \times r$ matrix $F^\top$, the MPGI is given by

$$(F^\top)^+ = F(F^\top F)^+ = F X^+ \,.$$

This concept enables us to study the linear equation system $F^\top \mathbf{y} = \mathbf{x}$ in $\mathbf{y}$. It always has a solution since $F^\top (F^\top)^+ \mathbf{x} = F^\top F X^+ \mathbf{x} = X X^+ \mathbf{x} = \mathbf{x}$ by (2), and the general solution is of the form

$$\mathbf{y} = (F^\top)^+ \mathbf{x} + (I_m - (F^\top)^+ F^\top)\mathbf{u} = \mathbf{p} + (I_m - F X^+ F^\top)\mathbf{u}, \quad \mathbf{u} \in \mathbb{R}^m, \quad (4)$$

where $\mathbf{p} = F X^+ \mathbf{x}$. The solution of minimum norm is $\mathbf{y} = \mathbf{p}$ (where $\mathbf{u} = \mathbf{o}$). Now $\mathbf{p}^\top \mathbf{p} = \mathbf{x}^\top X^+ F^\top F X^+ \mathbf{x} = \mathbf{x}^\top X^+ \mathbf{x} \le 1$, where the latter inequality follows from $X - \mathbf{x}\mathbf{x}^\top \in \mathcal{P}$. Hence there is always a solution $\mathbf{y}$ to $F^\top \mathbf{y} = \mathbf{x}$ with $\mathbf{y}^\top \mathbf{y} \le 1$. This was proved, e.g., in [46, Lem.1.1,Cor.1.3]. Sometimes $\mathbf{y} = \mathbf{p}$ is the only choice, but for $r > n$ there could be better choices, see below. If $\mathbf{p}$ happens to have no negative entries, Y is said to have the *property of positivity of least squares solution (PLSS)* in [8, pp.98ff]. PLSS ensures that the lower triangular block factorization reviewed in Subsection 3.1 below works, but this property is quite restrictive as will be documented by our empirical study. Anyhow, if X is diagonal as in a related article on cp factorization [36], PLSS holds: obviously $F = \sqrt{X} \ge O$ provides a cp factorization with $\mathbf{p} = F X^+ \mathbf{x} \in \mathbb{R}^n_+$.

3.1 Lower triangular blocks

Complete positivity of Y as in (1) is characterized in [46, Prop.1.4] as follows (with a slight change of notation): there is an $r \times n$ matrix F_0 with no negative entries, and a vector $\mathbf{y}_0 \in \mathbb{R}^r_+$ with $\mathbf{y}_0^\top \mathbf{y}_0 = 1$ such that $F_0^\top \mathbf{y}_0 = \mathbf{x}$ and $X = F_0^\top F_0$.

Since completely positive factorizations are by no means unique, knowledge of F in (3) does not imply that the above F_0 and $\mathbf{y}_0$ are known. In particular, it is not guaranteed that $F = F_0$ or $\mathbf{y} = \mathbf{y}_0$. They can have even different sizes.

If we would like to search for $(F_0, \mathbf{y}_0)$ directly, this amounts to solving a system of, again, $\binom{n+1}{2} + n + 1 = \binom{n+2}{2}$ quadratic equations in now $(r+1)n \le \frac{1}{2}(n^3 + n^2 - 6n)$ nonnegative variables, slightly less than the original problem but still prohibitively demanding.

Anyhow, assume now that there is a *nonnegative* solution $\mathbf{y} \in \mathbb{R}^r_+$ to $F^\top \mathbf{y} = \mathbf{x}$ with $\mathbf{y}^\top \mathbf{y} \le 1$. Then we can use factors H with lower block triangular structure as follows:

$$Y = H^\top H \quad \text{with} \quad H = \begin{bmatrix} \beta & \mathbf{o}^\top \\ \mathbf{y} & F \end{bmatrix} \quad \text{where} \quad \beta = \sqrt{1 - \mathbf{y}^\top \mathbf{y}}\,. \quad (5)$$

4

This can be checked by straightforward calculation. From an algorithmic perspective it could pay to first try (5), e.g. by solving the linear optimization problem

$$\min\left\{\sum_i y_i : F^\top \mathbf{y} = \mathbf{x}, \, \mathbf{y} \in \mathbb{R}_+^m\right\},\tag{6}$$

or even the convex QP variant with an objective $\mathbf{y}^\top \mathbf{y}$. Since H has one more row than F, the cp-rank increment from X to Y cannot exceed one, if F provides a minimal cp factorization. Moreover, if in this situation $\mathbf{y}^\top \mathbf{y} = 1$, then the cp-rank of Y is even equal to that of X, as observed in [8, Exerc.3.7,p.146].

3.2 Upper triangular blocks

Unfortunately, unless $F = F_0$ by chance, problem (6) can be infeasible or its feasible set can have empty intersection with the unit ball. So we will propose an alternative approach where we can allow for factorizations $X = F^\top F$ and vectors $\mathbf{y} \in \mathbb{R}^r$ such that $F^\top \mathbf{y} = \mathbf{x}$, where some entries of $\mathbf{y}$ may be negative. In this case, there are always solutions inside the unit ball, as detailed above after (4).

Theorem 3.1 *For $0 \le \alpha \le 1$ denote by $\varphi(\alpha) = 1 + \sqrt{1 - \alpha}$. Suppose $F^\top \mathbf{y} = \mathbf{x}$ with $\mathbf{y}^\top \mathbf{y} \le 1$ and such that*

$$\mathbf{y}\mathbf{x}^\top \le \varphi(\mathbf{y}^\top \mathbf{y})F,\tag{7}$$

where $\le$ is understood entrywise. Then

(a) $G := \sqrt{I_r - \mathbf{y}\mathbf{y}^\top}\,F$ *is an $r \times n$ matrix with no negative entries, and*

(b)

$$Y = H^\top H \qquad \text{with} \qquad H = \begin{bmatrix} 1 & \mathbf{x}^\top \\ \mathbf{o} & G \end{bmatrix}\tag{8}$$

gives an explicit completely positive factorization of Y.

Proof. It is easy to verify that $I_r - \mathbf{y}\mathbf{y}^\top \in \mathcal{P}$ and that

$$\sqrt{I_m - \mathbf{y}\mathbf{y}^\top} = I_m - \frac{1}{\varphi(\mathbf{y}^\top \mathbf{y})}\,\mathbf{y}\mathbf{y}^\top.\tag{9}$$

So $\varphi(\mathbf{y}^\top \mathbf{y})G = \varphi(\mathbf{y}^\top \mathbf{y})F - \mathbf{y}(F^\top \mathbf{y})^\top = \varphi(\mathbf{y}^\top \mathbf{y})F - \mathbf{y}\mathbf{x}^\top$, and (a) follows. The matrix product in (b) equals

$$\begin{bmatrix} 1 & \mathbf{x}^\top \\ \mathbf{x} & \mathbf{x}\mathbf{x}^\top + G^\top G \end{bmatrix},$$

and the lower right block is $\mathbf{x}\mathbf{x}^\top + F^\top(I_m - \mathbf{y}\mathbf{y}^\top)F = F^\top F = X$, so the product is indeed Y. □

For $\mathbf{x} = \mathbf{o}$, condition (7) is trivially satisfied, and (8) and (5) coincide. But only (7) also still holds for small $\max_j x_j$ and fixed positive F, whereas (6) could be immediately rendered infeasible even for small positive departures of $\mathbf{x}$ from $\mathbf{o}$, because feasibility of (6) is a homogeneous property; notice that from the 1600 cases generated in Section 5, only 9 satisfied $\mathbf{y} \geq \mathbf{o}$. This would not hurt for other purposes, e.g. [43] aiming at a more general factorization, but for a lower triangular factorization (5) it is essential. So only Theorem 3.1 may be viewed as a quantitative perturbation result, dealing with departures from the trivial block-diagonal case $\mathbf{x} = \mathbf{o}$.

The next section deals with condition (7). We will specify algorithmic approaches to satisfying this condition, and also show how to obtain an explicit factorization for all completely 3×3 matrices in this way.

4   En route to satisfying (7)

The function φ as defined in Theorem 3.1 is decreasing, concave, and satisfies $\varphi(0) = 2$ as well as $\varphi(1) = 1$. Therefore we have the following estimates

$$1 \leq 2 - \alpha \leq \varphi(\alpha) \quad \text{for all } \alpha \in [0,1]. \tag{10}$$

The approximation of lowest order uses the constant underestimation in (10), which results in linear constraints: so, any $\mathbf{y} \in \mathbb{R}^m$ with

$$\left.\begin{array}{rcl} F^\top \mathbf{y} & = & \mathbf{x} \\ \mathbf{y}^\top \mathbf{y} & \leq & 1 \\ \mathbf{y}\mathbf{x}^\top & \leq & F \end{array}\right\} \tag{11}$$

satisfies (7). The first-order approximation uses the linear underestimation in (10). This yields (inhomogeneous) convex quadratic constraints:

$$\left.\begin{array}{rcl} F^\top \mathbf{y} & = & \mathbf{x} \\ \mathbf{y}^\top \mathbf{y} & \leq & 1 \\ \mathbf{y}\mathbf{x}^\top & \leq & (2 - \mathbf{y}^\top \mathbf{y})F \end{array}\right\}. \tag{12}$$

Likewise, (12) implies (7). Finally, we can rewrite (7) without square roots: it is evident by elementary calculations that

$$\left.\begin{array}{rcl} F^\top \mathbf{y} & = & \mathbf{x} \\ \mathbf{y}^\top \mathbf{y} & \leq & 1 \\ \left(\left[\mathbf{y}\mathbf{x}^\top - F\right]_+\right)^{\bullet 2} & \leq & (1 - \mathbf{y}^\top \mathbf{y})F^{\bullet 2} \end{array}\right\} \tag{13}$$

is equivalent to (7). As mentioned above, only the last conditions in (11), (12) or (13) can be violated by $\mathbf{y} = \mathbf{p}$. To increase our chances of satisfying (7), we therefore employ an optimization approach in that we allow for $\mathbf{y} = \mathbf{p} + P\mathbf{u}$ as in (4) with

$$P = I_r - FX^+F^\top \tag{14}$$

the orthoprojector onto $\ker(F^\top) = (\mathrm{im}F)^\perp$. Now $\mathbf{y}^\top\mathbf{y} \le 1$ is no longer guaranteed, but we know by $\mathbf{p} = FX^+\mathbf{x} \perp P\mathbf{u}$ and $P^\top P = P$ that

$$\mathbf{y}^\top\mathbf{y} = \mathbf{p}^\top\mathbf{p} + \mathbf{u}^\top P\mathbf{u}, \tag{15}$$

and [46, Lem.1.1,Cor.1.3] guarantees that at least $\mathbf{p}^\top\mathbf{p} \le 1$. Hence we consider the optimization problems

$$\min\left\{\mathbf{y}^\top\mathbf{y} : \mathbf{y} \in \mathbb{R}^r, \, \mathbf{y} \text{ satisfies } (\triangle)\right\},$$

where $(\triangle)$ stands for (11), (12) or (13).

Proposition 4.1 *Suppose that $\mathbf{x} \ne \mathbf{o}$ in a given factorization (3) for (1).*

(a) Consider the convex, linearly constrained QP

$$\pi_0 = \min\left\{\mathbf{y}^\top\mathbf{y} : \mathbf{y} \in \mathbb{R}^r, \, F^\top\mathbf{y} = \mathbf{x}, \, \mathbf{y}\mathbf{x}^\top \le F\right\}. \tag{16}$$

Then

(a1) any feasible $\mathbf{y}$ with $\mathbf{y} \ge \mathbf{o}$ satisfies $\mathbf{y}^\top\mathbf{y} \le 1$;

(a2) any feasible $\mathbf{y}$ with $\mathbf{y}^\top\mathbf{y} \le 1$ gives rise to a cp factorization (8);

(a3) else $\pi_0 > 1$ (this includes infeasibility of (16) by putting $\pi_0 = +\infty$);
 then (11) has no solution.

(b) Consider the convex quadratically constrained QP

$$\pi_1 = \min\left\{\mathbf{y}^\top\mathbf{y} : \mathbf{y} \in \mathbb{R}^r, \, F^\top\mathbf{y} = \mathbf{x}, \, \mathbf{y}\mathbf{x}^\top + (\mathbf{y}^\top\mathbf{y})F \le 2F\right\}. \tag{17}$$

Then

(b1) any feasible $\mathbf{y}$ with $\mathbf{y} \ge \mathbf{o}$ satisfies $\mathbf{y}^\top\mathbf{y} \le 1$;

(b2) any feasible $\mathbf{y}$ with $\mathbf{y}^\top\mathbf{y} \le 1$ gives rise to a cp factorization (8);

(b3) else $\pi_1 > 1$ (this includes infeasibility of (17) by putting $\pi_1 = +\infty$);
 then (12) has no solution.

(c) consider the convex nonsmoothly constrained QP

$$\pi_e = \min\left\{\mathbf{y}^\top\mathbf{y} : \mathbf{y} \in \mathbb{R}^r, \, F^\top\mathbf{y} = \mathbf{x}, \, \left([\mathbf{y}\mathbf{x}^\top - F]_+\right)^{\bullet 2} + (\mathbf{y}^\top\mathbf{y})F^{\bullet 2} \le F^{\bullet 2}\right\}. \tag{18}$$

Then

(c1) any feasible $\mathbf{y}$ gives rise to a cp factorization (8);

(c2) else (18) is infeasible, and (13) has no solution.

Proof. Statements (a1) and (b1) follow by multiplying the inequalities in (16) and (17), respectively, by $\mathbf{y}^\top \ge \mathbf{o}$ from the left, observing $\mathbf{y}^\top F = \mathbf{x}^\top$. So only (c2) remains to be shown. Since $\mathbf{x} \ne \mathbf{o}$ by assumption, also $F^\top \ne O$. If $F_{ij} > 0$, then (18)-feasibility of $\mathbf{y}$ implies $(\mathbf{y}^\top\mathbf{y})F_{ij}^2 \le F_{ij}^2$ and thus $\mathbf{y}^\top\mathbf{y} \le 1$. □

Problem (18) can be rewritten as a convex quadratically constrained QP, introducing more variables and more (linear) constraints:

$$\pi_e = \min_{(\mathbf{y},Z)\in\mathbb{R}^r\times\mathbb{R}_+^{r\times n}}\left\{\mathbf{y}^\top\mathbf{y} : F^\top\mathbf{y} = \mathbf{x}, \mathbf{y}\mathbf{x}^\top - F \le Z, Z^{\bullet 2} + (\mathbf{y}^\top\mathbf{y})F^{\bullet 2} \le F^{\bullet 2}\right\}.$$

5 Some empirical evidence

If F is square (implying cp-rank$(X) \leq n$) and nonsingular, $\mathbf{y} = F^{-\top}\mathbf{x}$ is the unique solution to $F^{\top}\mathbf{y} = \mathbf{x}$ (here $F^{-\top} = (F^{\top})^{-1}$). Then the problems (16), (17) and (18) are of no use, and one can check directly whether or not $\mathbf{y} = \mathbf{p}$ satisfies (11),(12), or (13). In a small simulation study, 1600 such square F matrices with positive entries were generated randomly and $X = F^{\top}F$. For a parameter $\sigma \in \{1, 1.1, 2, 100\}$, a vector $\mathbf{x} \in \mathbb{R}_+^n$ was drawn at random and rescaled such that $\mathbf{x}^{\top}X^{-1}\mathbf{x} = \frac{1}{\sigma}$. Obviously $\lambda_{\min}(X - \mathbf{x}\mathbf{x}^{\top})$ increases with σ while $\sigma = 1$ means that the generated matrix $Y \in \mathcal{P} \cap \mathcal{N}$ is singular. It turns out that large values of σ favour the proposed approach even for relatively large instances, but even in singular cases a success is often encountered, particularly for moderate dimensions ($n + 1 \in \{5, 20, 100\}$ was chosen). Overall we observe successes in more than 63% of the cases, which increases to 77% if $\sigma \geq 2$. The details are given in Table 1. The four numbers in every cell count how often the conditions $\mathbf{y} \geq \mathbf{o}$, (11),(12), and (13) are met for $\mathbf{y} = \mathbf{p}$. The last column gives the success rate across all n, cumulated over all cases with $\mathbf{x}^{\top}X^{-1}\mathbf{x} \geq \frac{1}{\sigma}$.

Table 1: Success counts for conditions $[\mathbf{y} \geq \mathbf{o}]|(11)|(12)|(13)$ with $\mathbf{y}^{\top}\mathbf{y} = \frac{1}{\sigma}$. For every cell 100 random $(n + 1) \times (n + 1)$-instances were generated.

σ	$n = 3$	$n = 4$	$n = 19$	$n = 99$	succ.for $\geq \sigma$
1	15|61|71|73	1|63|72|75	0|18|26|29	0|7|8|8	63,44 %
1.1	9|53|64|67	3|54|66|70	0|20|27|30	0|7|11|13	67,92 %
2	8|78|85|86	2|72|78|79	0|33|39|45	0|10|12|12	77,63 %
100	7|93|93|93	3|98|99|99	0|99|99|99	0|86|92|93	97,75 %
success	90,75 %	80,75 %	50,75 %	31,50 %	

The figures reported in Table 1 are quite encouraging. Remember that we generated matrices $Y \in \mathcal{P} \cap \mathcal{N}$ which have a completely positive block X with cp-rank not exceeding n, but we are not sure that Y is completely positive. So the decrease of success rates with increasing dimension have to be discounted by the probability that matrices Y of this kind are completely positive with cp-rank not exceeding $n + 1$.

In the generation process, we did not use the construction as in [46] with nonnegative $s \times n$ matrix F_0 and $\mathbf{y}_0 \in \mathbb{R}_+^s$ with $\mathbf{y}_0^{\top}\mathbf{y}_0 = 1$, generating $X = F_0^{\top}F_0$ and $x = F_0^{\top}\mathbf{y}_0$ to ensure complete positivity of Y, for the following reason: then we either would have a trivial success for the simulation (if we choose F such that $F^{\top}(\mathbb{R}^r) = F_0^{\top}(\mathbb{R}^s)$ like $F = F_0$ or $F^{\top} = [F_0^{\top}, O]$ or $F^{\top} = [\frac{1}{\sqrt{2}}F_0^{\top}, \frac{1}{\sqrt{2}}F_0^{\top}]$), or else we would have to pick an essentially different F such that $F^{\top}F = X$, which is not obvious at all.

Now we turn to cp-ranks possibly exceeding dimension. Here, we must use the problem (6) and one of the problems (16), (17), or (18). For simplicity, and because the differences in Table 1 were not that pronounced with the different approaches, we chose the convex, linearly constrained quadratic problem (16). The `MatLab` solvers `linprog` and `quadprog` were used. We basically follow the

same experimental scheme, but for numerical reasons we restrict attention to the non-singular cases $\sigma \geq 1.1$. For dimension $n + 1 = 5$, we allow for $m + 1 = \binom{n+1}{2} = 10$ even larger than the maximal cp-rank, for dimension $n + 1 = 20$ we use $(m+1) = (n+1)^2/4 = 100$ suggested by [27], and for dimension $n+1 = 100$ we simply double $m + 1 = 2(n + 1)$. For smaller n, we also observed that increasing m in the range between $2n$ and $n^2/2$ increases success probability, so that the generated instances indeed are a priori not too easy for the proposed method. Nevertheless, the previously observed patterns persist and possibly get even more pronounced. Details can be found in Table 2. Finally it is worth mentioning that Kaykobad's sufficient condition [37] of diagonal dominance was almost never met in both experiments; this diagonal dominance criterion follows from the results in [46], in a straightforward way, but appears to be too stringent a sufficient condition on average even for small dimensions. A similar statement holds for the PLSS property discussed shortly before Subsection 3.1.

Table 2: Success counts for the problems (6)|(16) with $\mathbf{x}^\top X^+ \mathbf{x} = \frac{1}{\sigma}$. For every cell 100 random $(n + 1) \times (n + 1)$-instances were generated.

σ	$n = 4,\ m = 9$	$n = 19,\ m = 99$	$n = 99,\ m = 199$	succ.for $\geq \sigma$
1.1	18|52	0|5	0|0	60,00 %
2	21|70	0|74	0|0	77,50 %
100	12|88	0|100	0|100	100,00 %
success	87,00 %	59,67 %	33,33 %	

The case of $n = 4$ is particularly interesting, as we have $\mathcal{C} = \mathcal{P} \cap \mathcal{N}$ in dimension four, but strict inclusion in dimension 5. The papers [9], [20], [26] and [42] discuss the 5×5 case.

For illustration, we now specify five instances generated by above construction with $\sigma = 2$. Recall in these cases always necessarily $\mathbf{y} = \mathbf{p}$, as F is square nonsingular.

$$[F_1|\mathbf{x}_1|\mathbf{p}_1] = \begin{bmatrix} 0.0245 & 0.4916 & 0.2664 & 0.1435 & 0.0334 & 0.4044 \\ 0.0860 & 0.9231 & 0.8934 & 0.6927 & 0.0396 & -0.3237 \\ 0.4181 & 0.1568 & 0.8756 & 0.4079 & 0.0598 & -0.1253 \\ 0.2231 & 0.3427 & 0.7553 & 0.5943 & 0.0589 & 0.4648 \end{bmatrix}$$

satisfy (11) and therefore also (12) and (13), while

$$[F_2|\mathbf{x}_2|\mathbf{p}_2] = \begin{bmatrix} 0.6368 & 0.8460 & 0.8173 & 0.3300 & 0.6742 & 0.5800 \\ 0.7691 & 0.1724 & 0.2346 & 0.2055 & 0.6050 & 0.3694 \\ 0.0540 & 0.0370 & 0.5874 & 0.3645 & 0.7360 & 0.0535 \\ 0.1148 & 0.3126 & 0.9242 & 0.9757 & 0.4387 & 0.1556 \end{bmatrix}$$

violate (11), but satisfy (12) and therefore (13). On the other hand,

$$[F_3|\mathbf{x}_3|\mathbf{p}_3] = \begin{bmatrix} 0.0743 & 0.7709 & 0.5029 & 0.1131 & 0.0536 & 0.6108 \\ 0.1932 & 0.3139 & 0.9477 & 0.8121 & 0.3349 & 0.1541 \\ 0.3796 & 0.6382 & 0.8280 & 0.9083 & 0.3169 & 0.1500 \\ 0.2764 & 0.9866 & 0.9176 & 0.1564 & 0.2861 & -0.2840 \end{bmatrix}$$

violate (12) and therefore (11), but satisfy (13). Finally

$$[F_4|\mathbf{x}_4|\mathbf{p}_4] = \left[\begin{array}{ccc|c|c} 0.4154 & 0.0150 & 0.9901 & 0.4921 & 0.6135 \\ 0.3050 & 0.7680 & 0.7889 & 0.2113 & -0.1430 \\ 0.8744 & 0.9708 & 0.4387 & 0.6355 & 0.3213 \end{array}\right]$$

violate (13) and therefore our methods presented up to now do not apply to $Y_4 = \left[\begin{array}{cc} 1 & \mathbf{x}_4^\top \\ \mathbf{x}_4 & F_4^\top F_4 \end{array}\right]$. Since $Y_4 \in \mathcal{P} \cap \mathcal{N}$ is a 4×4-matrix, it must be completely positive, and in Section 7 we will factorize it by a refinement of these. The last example

$$[F_5|\mathbf{x}_5|\mathbf{p}_5] = \left[\begin{array}{ccc|c|c} 0.5369 & 0.4175 & 0.7538 & 0.3648 & 0.3416 \\ 0.0665 & 0.2923 & 0.0968 & 0.3870 & 0.5449 \\ 0.4939 & 0.2897 & 0.0769 & 0.3329 & 0.2938 \end{array}\right]$$

illustrates that the condition $\mathbf{p} \geq \mathbf{o}$ from [46] can be satisfied even if our condition (13) is violated, so that it really pays to combine all methods (although in higher dimensions, importance of the lower-triangular factorization decreases).

6 Explicit factorization of low order matrices

We start by discussing the (doubly nonnegative) 2×2 case. If a diagonal entry is zero, there is only one positive factorization (of rank zero or one). Else, we may and do rescale as before so that that $X = \left[\begin{array}{cc} 1 & a \\ a & 1 \end{array}\right] \in \mathcal{P} \cap \mathcal{N}$. We may assume $0 \leq a < 1$, because otherwise ($a = 1$ and) $X = \mathbf{e}\mathbf{e}^\top$ with $\mathbf{e}^\top = [1, 1]$. In this case, there is a factorization

$$X = F^\top F \qquad \text{with} \qquad F = \left[\begin{array}{cc} 1 & a \\ 0 & \sqrt{1 - a^2} \end{array}\right]. \qquad (19)$$

Now let us proceed to the three-dimensional case. We basically show that either of the factorizations (8) or (5) apply with the same F. So we again consider

$$Y = \left[\begin{array}{cc} 1 & \mathbf{x}^\top \\ \mathbf{x} & X \end{array}\right] \in \mathcal{P} \cap \mathcal{N}, \qquad (20)$$

which is equivalent to stipulate $X \in \mathcal{P} \cap \mathcal{N}$; $\mathbf{x} \in \mathbb{R}_+^2$; and $X - \mathbf{x}\mathbf{x}^\top \in \mathcal{P}$. Again we may and do assume diag $X = \mathbf{e} = [1, 1]^\top$, and $X = F^\top F$ with F as in (19). If X is singular, then $X = \mathbf{e}\mathbf{e}^\top$; the condition $\mathbf{e}\mathbf{e}^\top - \mathbf{x}\mathbf{x}^\top \in \mathcal{P}$ implies $\mathbf{e} = \alpha \mathbf{x}$ for some $\alpha \geq 1$ and

$$Y = \left[\begin{array}{cc} 1 & 0 \\ \mathbf{x} & \sqrt{\alpha^2 - 1}\,\mathbf{x} \end{array}\right] \left[\begin{array}{cc} 1 & \mathbf{x}^\top \\ 0 & \sqrt{\alpha^2 - 1}\,\mathbf{x}^\top \end{array}\right]$$

is the cp factorization of the form (8). However, if X is nonsingular, so is F and $F^{-\top} = \frac{1}{\sqrt{1-a^2}} \left[\begin{array}{cc} \sqrt{1 - a^2} & 0 \\ -a & 1 \end{array}\right]$. So we get $y_1 = p_1 = x_1 \geq 0$ whereas $y_2 = p_2 = \frac{x_2 - ax_1}{\sqrt{1-a^2}}$ can have either sign.

10

Next we distinguish cases according to the sign of y_2. If $y_2 \geq 0$, then $\mathbf{y} \in \mathbb{R}_+^2$ and we arrive at the factorization of the form (5):

$$Y = \left[\begin{array}{cc} \mathbf{y}^\top & \sqrt{1 - \mathbf{y}^\top \mathbf{y}} \\ F^\top & \mathbf{o} \end{array} \right] \left[\begin{array}{cc} \mathbf{y} & F \\ \sqrt{1 - \mathbf{y}^\top \mathbf{y}} & \mathbf{o}^\top \end{array} \right] \qquad (21)$$

If, however, $y_2 < 0$, then the matrix $I_2 - \mathbf{y}\mathbf{y}^\top \in \mathcal{P} \cap \mathcal{N}$ (and this is true only if $\mathbf{y}$ has at most two nonzero entries of opposite sign, which is guaranteed only for $n = 2$!), so we may factorize by taking square roots. Indeed, from (9) we have

$$Q = \sqrt{I_2 - \mathbf{y}\mathbf{y}^\top} = I_n - \beta \mathbf{y}\mathbf{y}^\top \qquad \text{with} \qquad \beta = \frac{1}{\varphi(\mathbf{y}^\top \mathbf{y})} \leq 1$$

which again has no negative entry in this particular case of order two when $y_1 y_2 \leq 0$. By consequence, and this time straighforwardly as both factors of G are nonnegative,

$$X - \mathbf{x}\mathbf{x}^\top = F^\top F - F^\top \mathbf{y}\mathbf{y}^\top F = F^\top \bar{X} F = G^\top G \quad \text{with} \quad G = QF \text{ nonnegative.} \qquad (22)$$

So we obtain the desired factorization of the form (8). This establishes also the well-known fact that the cp-rank of any completely positive 3×3 matrix is at most three. The elementary argument here differs from the more involved one in [8, Cor.2.13,p.126] which however establishes factorizations of triangular type whereas the above F is usually more dense, e.g. for the data of [8, Ex.2.19,p.126]. For alternative, still different factorization results for all matrices of *rank* 3 (which can result in $r \geq 4$ if $n \geq 4$, cf. [8, Ex.3.1,p.140]) see [4, 17].

7   Towards larger cp-rank increments

Now, if we start with a cp-rank not exceeding n, as for $n = 2$ or $n = 3$, and apply the construction of Theorem 3.1 in a recursive way to building completely positive factorization, we can reach only matrices with the same property (cp-rank less or equal order). Indeed, H has $r + 1$ rows if G has r, like F. The same is intrinsically true if we aim for the factorization (5).

Nevertheless, this still has some justification as the solutions of copositive optimization problems arising from most applications are expected to have a low cp-rank. An admittedly heuristic argument goes as follows: imitating the proof of [49, Thm.3.4], one can show that for every matrix $Y \in \mathcal{C}$ one can construct a matrix $\widehat{Y} \in \partial \mathcal{C}$ with the same or larger cp-rank. But for a boundary point $\widehat{Y} = \widehat{H}^\top \widehat{H}$ we always can find a copositive matrix $S \in \mathcal{C}^* \setminus \{O\}$ which is orthogonal to it, i.e., $trace(\widehat{Y} S) = 0$. It follows that all columns of $\widehat{H}^\top$ must be global minimizers (with optimal value zero) of the quadratic form $\mathbf{x}^\top S \mathbf{x}$ over the standard simplex $\{\mathbf{x} \in \mathbb{R}_+^n : \sum_i x_i = 1\}$. A recent study on this problem [16], corroborated by asymptotic probabilistic results [21, 22, 39, 40, 41] shows that very few S have many global minimizers (although in the worst case there can even coexist an exponential number of them).

Anyhow, we can extend the strategy of (22) towards larger cp-rank increments as follows if $X - \mathbf{x}\mathbf{x}^\top$ is completely positive. Note that below arguments

do not rely on knowledge of the cp-rank. Indeed, instead of a minimal cp factorization of X we can start with any one. We will aim at enlarging the number of rows of G which will play the same role in H as in (8). As an aside, we note that by this construction, we get two alternative factorizations for X: the starting one, $X = F^\top F$ and the resulting one, $X = \mathbf{x}\mathbf{x}^\top + G^\top G = \tilde{F}^\top \tilde{F}$ where $\tilde{F}^\top = [\mathbf{x} \,|\, G^\top]$ has one more column than $G^\top$, so the latter won't be the minimal one if we succeed with our strategy. The same was already true for Theorem 3.1 where $\tilde{F}$ had one more row than F.

First we explain why we can assume without loss of generality that for a solution $\mathbf{y} \in \mathbb{R}^n$ (with some negative entries) to $F^\top \mathbf{y} = \mathbf{x}$, we have equality $\|\mathbf{y}\| = 1$:

Proposition 7.1 *Consider a solution $\mathbf{y}$ to $F^\top \mathbf{y} = \mathbf{x}$ with $\nu := \|\mathbf{y}\| < 1$ and define $\bar{\mathbf{y}} := \frac{1}{\nu}\mathbf{y}$ as well as $\bar{\mathbf{x}} := \frac{1}{\sqrt{\nu}}\mathbf{x}$, $\bar{X} := \nu X$. Then*

(a) *If $X = F^\top F$, then $\bar{X} = \bar{F}^\top \bar{F}$ where $\bar{F} := \sqrt{\nu}F \geq O$ satisfies $\bar{F}^\top \bar{\mathbf{y}} = \bar{\mathbf{x}}$ with $\|\bar{\mathbf{y}}\| = 1$;*

(b) *from any cp factorization of $\bar{Y} := \begin{bmatrix} 1 & \bar{\mathbf{x}}^\top \\ \bar{\mathbf{x}} & \bar{X} \end{bmatrix} = \bar{H}^\top \bar{H}$ we get a cp factorization of Y as follows:*

$$Y = H^\top H \qquad with \qquad H = \begin{bmatrix} \sqrt{1 - \sqrt{\nu}}\, \mathbf{e}_1^\top \\ \sqrt{\frac{1}{\nu} - \sqrt{\nu}}\,[\mathbf{o} \ \bar{F}] \\ \nu^{1/4}\bar{H} \end{bmatrix}.$$

Proof. (a) is straightforward. Likewise, to establish (b), we calculate

$$H^\top H = (1 - \sqrt{\nu})\mathbf{e}_1\mathbf{e}_1^\top + (\frac{1}{\nu} - \sqrt{\nu})\begin{bmatrix} 0 & \mathbf{o}^\top \\ \mathbf{o} & \bar{X} \end{bmatrix} + \sqrt{\nu}\,\bar{Y} = \begin{bmatrix} 1 & \sqrt{\nu}\bar{\mathbf{x}}^\top \\ \sqrt{\nu}\bar{\mathbf{x}} & \frac{1}{\nu}\bar{X} \end{bmatrix} = Y,$$

observing in addition the alternative representation

$$\sqrt{\frac{1}{\nu} - \sqrt{\nu}}\,\bar{F} = \sqrt{\nu - \nu^{3/2}}\,F$$

which can be used to build H when starting from the original factorization $X = F^\top F$, but $\bar{F}$ seems more natural as it occurs anyhow when building $\bar{H}$. $\square$

So let us assume in the sequel that $\|\mathbf{y}\| = 1$. The next step will imitate the case $y_1 y_2 < 0$ in the previous section: decompose $\mathbf{y}^\top = [\mathbf{u}^\top \,|\, -\mathbf{v}^\top]$ such that $\mathbf{u} \in \mathbb{R}_+^k \setminus \{\mathbf{o}\}$ and $\mathbf{v} \in \mathbb{R}_+^m \setminus \{\mathbf{o}\}$ with $k + m = r$ if $\mathbf{y} \in \mathbb{R}^r$, i.e., if F is an $r \times n$ matrix. As $\mathbf{u}^\top \mathbf{u} + \mathbf{v}^\top \mathbf{v} = \mathbf{y}^\top \mathbf{y} = 1$ we thus have $0 < \|\mathbf{u}\| < 1$ and $0 < \|\mathbf{v}\| < 1$, hence both $I_k - \mathbf{u}\mathbf{u}^\top$ and $I_m - \mathbf{v}\mathbf{v}^\top$ are positive-definite (but their square roots will have negative entries unless $k = 1$ or $m = 1$). We rescale $\mathbf{u}$ and $\mathbf{v}$ and imitate the high-cp-rank construction by [50, Prop.2.1]: consider

$$Q := \left[\frac{\mathbf{v}}{\|\mathbf{u}\|} \otimes \sqrt{I_k - \mathbf{u}\mathbf{u}^\top} \,\Big|\, \sqrt{I_m - \mathbf{v}\mathbf{v}^\top} \otimes \frac{\mathbf{u}}{\|\mathbf{v}\|} \right], \qquad (23)$$

a matrix with $km \leq r^2/4$ rows and $k+m = r$ columns. So, if the signs of entries of $\mathbf{y}$ are well balanced, this leaves a chance for a larger increment in cp-rank.

Theorem 7.1 *With above assumptions and notations, suppose that $F^\top F = X$ and $\mathbf{y} \in \mathbb{R}^r$ solves $F^\top \mathbf{y} = \mathbf{x}$ with $\|\mathbf{y}\| = 1$. To obtain Q as in (23), decompose*

$$\mathbf{y} = \left[\begin{array}{c} \mathbf{u} \\ -\mathbf{v} \end{array} \right] \quad with \quad (\mathbf{u}, \mathbf{v}) \in \mathbb{R}^k_+ \times \mathbb{R}^m_+ \quad and \quad F = \left[\begin{array}{c} S \\ T \end{array} \right] \quad (24)$$

where S is a $k \times n$ matrix and T is an $m \times n$ matrix. Denote by $C := I_k - \frac{\mathbf{u}\mathbf{u}^\top}{1+\|\mathbf{v}\|}$ and $D := I_m - \frac{\mathbf{v}\mathbf{v}^\top}{1+\|\mathbf{u}\|}$. Then $C = \sqrt{I_k - \mathbf{u}\mathbf{u}^\top}$ and $D = \sqrt{I_m - \mathbf{v}\mathbf{v}^\top}$.

Furthermore, let $\gamma = \frac{\|\mathbf{v}\|}{\|\mathbf{u}\|}$ and $\Delta = \left[\begin{array}{cc} \gamma I_k & O^\top \\ O & \frac{1}{\gamma}I_m \end{array} \right]$ and define $G := Q\Delta^{-1}F$. Then

(a) *the condition $G \geq O$ can be rephrased as*

$$\frac{\mathbf{v}}{\|\mathbf{v}\|} \otimes (CS) + (DT) \otimes \frac{\mathbf{u}}{\|\mathbf{u}\|} \geq O , \quad (25)$$

while the condition $F^\top \mathbf{y} = \mathbf{x}$ amounts to $S^\top \mathbf{u} - T^\top \mathbf{v} = \mathbf{x}$.

(b) *If $G \geq O$, then for the Schur complement we have $X - \mathbf{x}\mathbf{x}^\top = G^\top G$, so it is completely positive, too;*

(c) *the matrix*

$$H := \left[\begin{array}{cc} 1 & \mathbf{x}^\top \\ \mathbf{o} & G \end{array} \right]$$

provides a cp factorization, $H^\top H = Y$;

Proof. First recall (9) and that $\varphi(\mathbf{u}^\top \mathbf{u}) = 1 + \|\mathbf{v}\|$ due to $\mathbf{u}^\top \mathbf{u} + \mathbf{v}^\top \mathbf{v} = \|\mathbf{y}\|^2 = 1$, which yields $C^2 = I_k - \mathbf{u}\mathbf{u}^\top$ and likewise $D^2 = I_m - \mathbf{v}\mathbf{v}^\top$. Next abbreviate by $\mathbf{f} = \frac{\mathbf{v}}{\|\mathbf{u}\|}$ and by $\mathbf{g} = \frac{\mathbf{u}}{\|\mathbf{v}\|}$. Then $Q = [\mathbf{f} \otimes C \,|\, D \otimes \mathbf{g}]$ as defined in (23), and

$$C\mathbf{u} = (1 - \frac{\|\mathbf{u}\|^2}{1+\|\mathbf{v}\|})\mathbf{u} = \frac{1+\|\mathbf{v}\|-\|\mathbf{u}\|^2}{1+\|\mathbf{v}\|}\mathbf{u} = \frac{(1+\|\mathbf{v}\|)\|\mathbf{v}\|}{1+\|\mathbf{v}\|}\mathbf{u} = \|\mathbf{v}\|\,\mathbf{u} ,$$

so that we have $\mathbf{g} = C^{-1}\mathbf{u}$ and likewise $\mathbf{f} = D^{-1}\mathbf{v}$. To establish (a), tedious but straightforward calculations yield $Q\Delta^{-1}F = \frac{1}{\gamma}\mathbf{f} \otimes (CS) + \gamma(DT) \otimes \mathbf{g}$, and therefore we arrive at (25). For (b), first note that by definition $\mathbf{f}^\top \mathbf{f} = \gamma^2$ and $\mathbf{g}^\top \mathbf{g} = \frac{1}{\gamma^2}$. Next recall that $C^\top \mathbf{g} = C\mathbf{g} = \mathbf{u}$ and $D^\top \mathbf{f} = D\mathbf{f} = \mathbf{v}$ by

construction, so that we obtain

$$Q^\top Q = \begin{bmatrix} (\mathbf{f}^\top \otimes C^\top)(\mathbf{f} \otimes C) & (\mathbf{f}^\top \otimes C^\top)(D \otimes \mathbf{g}) \\ (D^\top \otimes \mathbf{g}^\top)(\mathbf{f} \otimes C) & (D^\top \otimes \mathbf{g}^\top)(D \otimes \mathbf{g}) \end{bmatrix}$$

$$= \begin{bmatrix} (\mathbf{f}^\top \mathbf{f}) \otimes (C^\top C) & (\mathbf{f}^\top D) \otimes (C^\top \mathbf{g}) \\ (D^\top \mathbf{f}) \otimes (\mathbf{g}^\top C) & (D^\top D) \otimes (\mathbf{g}^\top \mathbf{g}) \end{bmatrix}$$

$$= \begin{bmatrix} \gamma^2 C^2 & \mathbf{uv}^\top \\ \mathbf{vu}^\top & \frac{1}{\gamma^2} D^2 \end{bmatrix}$$

$$= \begin{bmatrix} \gamma^2 (I_k - \mathbf{uu}^\top) & \mathbf{uv}^\top \\ \mathbf{vu}^\top & \frac{1}{\gamma^2}(I_m - \mathbf{vv}^\top) \end{bmatrix}$$

$$= \Delta(I_r - \mathbf{yy}^\top)\Delta \quad \text{since} \quad \mathbf{yy}^\top = \begin{bmatrix} \mathbf{uu}^\top & -\mathbf{uv}^\top \\ -\mathbf{vu}^\top & \mathbf{vv}^\top \end{bmatrix}.$$

Hence $G^\top G = F^\top(I_r - \mathbf{yy}^\top)F = X - \mathbf{xx}^\top$. (c) follows as before. $\qquad\square$

Similar strategies as in Section 4 can be employed to satisfy (25) which can be rephrased (with $\bar{\mathbf{u}} = \frac{\mathbf{u}}{\|\mathbf{u}\|}, \bar{\mathbf{v}} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$) as

$$\bar{\mathbf{v}} \otimes S + T \otimes \bar{\mathbf{u}} \geq (\bar{\mathbf{v}} \otimes \bar{\mathbf{u}}) \left[\frac{1 + \|\mathbf{u}\| + \|\mathbf{v}\|}{(1 + \|\mathbf{u}\|)(1 + \|\mathbf{v}\|)} (T^\top \mathbf{v}) + \frac{\|\mathbf{u}\|}{1 + \|\mathbf{v}\|} \mathbf{x} \right]^\top .$$

In case of square nonsingular F, the only choice for $\mathbf{y} = \mathbf{p}$, so the decomposition (24) is predetermined and above sufficient condition can be checked easily. Take for example the data $[F_4, \mathbf{x}_4]$ for Y_4 at the end of Section 5. After rescaling by $\nu = \|\mathbf{y}_4\| = 0.7070$ to $[\bar{F}_4, \bar{\mathbf{x}}_4]$ as in Proposition 7.1, we have

$$\bar{F}_4 = \begin{bmatrix} 0.3493 & 0.0126 & 0.8325 \\ 0.2565 & 0.6458 & 0.6633 \\ 0.7352 & 0.8163 & 0.3689 \end{bmatrix}, \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} 0.8676 \\ 0.4544 \\ 0.2021 \end{bmatrix}, G^\top = \begin{bmatrix} 0.1120 & 0.6110 \\ 0.2973 & 0.9654 \\ 0.7658 & 0.3339 \end{bmatrix},$$

which gives a cp factorization $\bar{H} = \begin{bmatrix} 1 & \bar{\mathbf{x}}_4^\top \\ \mathbf{o} & G \end{bmatrix}$ of $\bar{Y}_4$, and from Proposition 7.1 we get

$$H^\top = \begin{bmatrix} 0.3990 & 0 & 0 & 0 & 0.9170 & 0 & 0 \\ 0 & 0.2645 & 0.1942 & 0.5568 & 0.5367 & 0.1027 & 0.5602 \\ 0 & 0.0096 & 0.4891 & 0.6182 & 0.2304 & 0.2726 & 0.8852 \\ 0 & 0.6305 & 0.5024 & 0.2794 & 0.6930 & 0.7022 & 0.3062 \end{bmatrix}$$

which yields indeed $H^\top H = Y_4$.

Finally we briefly sketch a possible extension if Schur complements are no longer completely positive. Then we have to abandon the upper-triangular strategy and use a general

$$H = \begin{bmatrix} \alpha & \mathbf{p}^\top \\ \mathbf{q} & R \end{bmatrix} \quad \text{with} \quad 0 \leq \alpha < 1,$$

where

$$X - \mathbf{xx}^\top = R^\top(I_n - \mathbf{qq}^\top)R + (1 - \alpha^2)\mathbf{pp}^\top - \alpha\left[\mathbf{p}(R^\top \mathbf{q})^\top + (R^\top \mathbf{q})\mathbf{p}^\top\right],$$

14

which not necessarily is completely positive. Observe that the choice $[\mathbf{p}|R^\top] = F^\top$ would lead back to solving $F^\top \mathbf{y} = \mathbf{x}$ with $\mathbf{y} = [\alpha, \mathbf{q}^\top]^\top \in \mathbb{R}_+^r$. The difference now is that, on one hand now $\|\mathbf{y}\|^2 = \alpha^2 + \mathbf{q}^\top \mathbf{q} = 1$ automatically by construction (but the previous $\|\mathbf{y}\| \leq 1$ posed no restriction as we have seen in Proposition 7.1). On the other hand, now the north-west corner of H can be less than one, which opens more possibilities to proceed similar as above. Given above empirical evidence, this may be a promising avenue of future research.

References

[1] Paula A. Amaral and Immanuel M. Bomze. Copositivity-based approximations for mixed-integer fractional quadratic optimization. *Pacific J. Optimiz.*, 11(2):225–238, 2015.

[2] Paula A. Amaral, Immanuel M. Bomze, and Joaquim J. Júdice. Copositivity and constrained fractional quadratic problems. *Math. Program.*, 146(1–2):325–350, 2014.

[3] Kurt M. Anstreicher and Samuel Burer. Computable representations for convex hulls of low-dimensional quadratic forms. *Math. Program.*, 124(1-2, Ser. B):33–43, 2010.

[4] Francesco Barioli and Abraham Berman. The maximal cp-rank of rank k completely positive matrices. *Linear Algebra Appl.*, 363:17–33, 2003.

[5] Abraham Berman, Mirjam Dür, and Naomi Shaked-Monderer. Open problems in the theory of completely positive and copositive matrices. *Electron. J. Linear Algebra*, 29:46–58, 2015.

[6] Abraham Berman and Daniel Hershkowitz. Combinatorial results on completely positive matrices. *Linear Algebra Appl.*, 95:111–125, 1987.

[7] Abraham Berman and Uriel G. Rothblum. A note on the computation of the CP-rank. *Linear Algebra Appl.*, 419(1):1–7, 2006.

[8] Abraham Berman and Naomi Shaked-Monderer. *Completely positive matrices.* World Scientific Publishing Co. Inc., River Edge, NJ, 2003.

[9] Abraham Berman and Changqing Xu. 5×5 completely positive matrices. *Linear Algebra Appl.*, 393:55–71, 2004.

[10] Immanuel M. Bomze. Copositive optimization – recent developments and applications. *European J. Oper. Res.*, 216:509–520, 2012.

[11] Immanuel M. Bomze, Mirjam Dür, Etienne de Klerk, Cornelis Roos, Arie J. Quist, and Tamás Terlaky. On copositive programming and standard quadratic optimization problems. *J. Global Optim.*, 18(4):301–320, 2000.

[12] Immanuel M. Bomze, Mirjam Dür, and Chung-Piaw Teo. Copositive optimization. *Optima – MOS Newsletter*, 89:2–10, 2012.

[13] Immanuel M. Bomze and Florian Jarre. A note on Burer's copositive representation of mixed-binary QPs. *Optim. Lett.*, 4(3):465–472, 2010.

[14] Immanuel M. Bomze, Werner Schachinger, and Reinhard Ullrich. From seven to eleven: completely positive matrices with high cp-rank. *Linear Algebra Appl.*, 459:208–221, 2014.

[15] Immanuel M. Bomze, Werner Schachinger, and Reinhard Ullrich. New lower bounds and asymptotics for the cp-rank. *SIAM J. Matrix Anal. Appl.*, 36(1):20–37, 2015.

[16] Immanuel M. Bomze, Werner Schachinger, and Reinhard Ullrich. The complexity of simple models – a study of worst and typical hard cases for the standard quadratic optimization problem. Preprint, submitted, `http://www.optimization-online.org/DB\_HTML/2016/05/5452.html`, 2016.

[17] Jan Brandts and Michal Křížek. Factorization of cp-rank-3 completely positive matrices. *Czechoslovak Math. J.*, 66(3):955–970, 1962.

[18] Samuel Burer. On the copositive representation of binary and continuous nonconvex quadratic programs. *Math. Program.*, 120(2, Ser. A):479–495, 2009.

[19] Samuel Burer. Copositive programming. In Miguel F. Anjos and Jean Bernard Lasserre, editors, *Handbook of Semidefinite, Cone and Polynomial Optimization: Theory, Algorithms, Software and Applications*, International Series in Operations Research and Management Science, pages 201–218. Springer, New York, 2012.

[20] Samuel Burer, Kurt M. Anstreicher, and Mirjam Dür. The difference between 5×5 doubly nonnegative and completely positive matrices. *Linear Algebra Appl.*, 431(9):1539–1552, 2009.

[21] Xin Chen and Jiming Peng. New analysis on sparse solutions to random Standard Quadratic Optimization Problems and extensions. *Math. Oper. Res.*, 40(3):725–738, 2015.

[22] Xin Chen, Jiming Peng, and Shuzhong Zhang. Sparse solutions to random Standard Quadratic Optimization Problems. *Math. Programming*, 141(1):273–293, 2013.

[23] Palahenedi Hewage Diananda. On non-negative forms in real variables some or all of which are non-negative. *Proc. Cambridge Philos. Soc.*, 58:17–25, 1962.

[24] Peter J. C. Dickinson. An improved characterisation of the interior of the completely positive cone. *Electron. J. Linear Algebra*, 20:723–729, 2010.

[25] Peter J. C. Dickinson and Mirjam Dür. Linear-time complete positivity detection and decomposition of sparse matrices. *SIAM J. Matrix Anal. Appl.*, 33(3):701–720, 2012.

[26] Hongbo Dong and Kurt Anstreicher. A note on "5×5 completely positive matrices". *Linear Algebra and Appl.*, 433(5):1001–1004, 2010.

[27] John H. Drew, Charles R. Johnson, and Raphael Loewy. Completely positive matrices associated with M-matrices. *Linear Multilinear Algebra*, 37(4):303–310, 1994.

[28] Mirjam Dür. Copositive programming — a survey. In Moritz Diehl, Francois Glineur, Elias Jarlebring, and Wim Michiels, editors, *Recent Advances in Optimization and its Applications in Engineering*, pages 3–20. Springer, Berlin Heidelberg New York, 2010.

[29] Mirjam Dür and Georg Still. Interior points of the completely positive cone. *Electron. J. Linear Algebra*, 17:48–53, 2008.

[30] L. J. Gray and D. G. Wilson. Nonnegative factorization of positive semidefinite nonnegative matrices. *Linear Algebra and Appl.*, 31:119–127, 1980.

[31] Walter J. Gutjahr. Convergence analysis of metaheuristics. In Vittorio Maniezzo, Thomas Stuetzle, and Stefan Voß, editors, *Matheuristics – Hybridizing Metaheuristics and Mathematical Programming*, volume 10 of *Annals of Information Systems*, pages 159–187. Springer, New York, 1995.

[32] Walter J. Gutjahr. Stochastic search in metaheuristics. In Michel Gendreau and Jean-Yves Potvin, editors, *Handbook of Metaheuristics*, volume 146 of *International Series in Operations Research & Management Science*, pages 573–597. Springer, New York, 2010.

[33] Marshall Hall, Jr. and Morris Newman. Copositive and completely positive quadratic forms. *Proc. Cambridge Philos. Soc.*, 59:329–339, 1963.

[34] Zhaoshui He, Shengli Xie, Rafal Zdunek, and Andrzej Cichocki. Symmetric nonnegative matrix factorization: algorithms and applications to probabilistic clustering. *IEEE Trans. Neural Netw.*, 22(12):2117–2131, 2011.

[35] Florian Jarre and Katrin Schmallowsky. On the computation of C^* certificates. *J. Global Optim.*, 45(2):281–296, 2009.

[36] Vassilis Kalofolias and Efstratios Gallopoulos. Computing symmetric nonnegative rank factorizations. *Linear Algebra Appl.*, 436(2):421–435, 2012.

[37] Mohammad Kaykobad. On nonnegative factorization of matrices. *Linear Algebra Appl.*, 96:27–33, 1987.

[38] Colleen Kelly. A test of the Markovian model of DNA evolution. *Biometrics*, 50(3):653–664, 1994.

[39] Spyros Kontogiannis and Paul Spirakis. Counting stable strategies in random evolutionary games. In X. Deng and D. Du, editors, *Proceedings of the 16th Annual International Symposium on Algorithms and Computation (ISAAC 2005)*, volume 3827 of *LNCS*, pages 839–848. Springer, Heidelberg, 2006.

[40] Spyros C. Kontogiannis and Paul G. Spirakis. On the support size of stable strategies in random games. *Theoretical Computer Science*, 410(8-10):933–942, 2009.

[41] Spyros C. Kontogiannis and Paul G. Spirakis. Well supported approximate equilibria in bimatrix games. *Algorithmica*, 57(2):653–667, 2010.

[42] Raphael Loewy and Bit-Shun Tam. CP rank of completely positive matrices of order 5. *Linear Algebra Appl.*, 363:161–176, 2003.

[43] Kathik Natarajan and Chung Piaw Teo. On reduced semidefinite programs for second order moment bounds with applications. *Math. Program.*, 161:487–518, 2017.

[44] Kathik Natarajan, Chung Piaw Teo, and Zhichao Zheng. Mixed zero-one linear programs under objective uncertainty: a completely positive representation. *Operations Research*, 59(3):713–728, 2011.

[45] Arie J. Quist, Etienne de Klerk, Cornelis Roos, and Tamas Terlaky. Copositive relaxation for general quadratic programming. *Optim. Methods Softw.*, 9(1-3):185–208, 1998.

[46] Luigi Salce and Paolo Zanardo. Completely positive matrices and positivity of least squares solutions. *Linear Algebra Appl.*, 178:201–216, 1993.

[47] Naomi Shaked-Monderer. A note on upper bounds on the cp-rank. *Linear Algebra Appl.*, 431:2407–2413, 2009.

[48] Naomi Shaked-Monderer. Matrices attaining the minimum semidefinite rank of a chordal graph. *Linear Algebra Appl.*, 438(10):3804–3816, 2013.

[49] Naomi Shaked-Monderer, Abraham Berman, Immanuel M. Bomze, Florian Jarre, and Werner Schachinger. New results on the cp rank and related properties of co(mpletely)positive matrices. *Linear Multilinear Algebra*, 63(2):384–396, 2015.

[50] Naomi Shaked-Monderer, Immanuel M. Bomze, Florian Jarre, and Werner Schachinger. On the cp-rank and minimal cp factorizations of a completely positive matrix. *SIAM J. Matrix Anal. Appl.*, 34(2):355–368, 2013.

[51] Julia Sponsel and Mirjam Dür. Factorization and cutting planes for completely positive matrices by copositive projection. *Math. Program.*, 143(1–2):211–229, 2014.

[52] Yu-Xiong Wang and Yu-Jin Zhang. Nonnegative matrix factorization: a comprehensive review. *IEEE Trans. Knowledge Data Eng.*, 25(6):1336–1353, 2013.

[53] Anwa Zhou and Jinyan Fan. Interiors of completely positive cones. *J. Global Optim.*, 63(4):653–675, 2015.