Noname manuscript No.
(will be inserted by the editor)

Quadratic Optimization with Switching Variables: The Convex Hull for $n = 2$

Kurt M. Anstreicher · **Samuel Burer**

Received: date / Accepted: date

Abstract We consider quadratic optimization in variables (x, y) where $0 \leq x \leq y$, and $y \in \{0,1\}^n$. Such binary y are commonly referred to as *indicator* or *switching* variables and occur commonly in applications. One approach to such problems is based on representing or approximating the convex hull of the set $\{(x, xx^T, yy^T) : 0 \leq x \leq y \in \{0,1\}^n\}$. A representation for the case $n = 1$ is known and has been widely used. We give an exact representation for the case $n = 2$ by starting with a disjunctive representation for the convex hull and then eliminating auxiliary variables and constraints that do not change the projection onto the original variables. An alternative derivation for this representation leads to an appealing conjecture for a simplified representation of the convex hull for $n = 2$ when the product term $y_1 y_2$ is ignored.

Keywords Quadratic optimization · switching variables · convex hull · perspective cone · semidefinite programming

Mathematics Subject Classification (2010) 90C11 · 90C20 · 90C22 · 90C25 · 90C26

1 Introduction

This paper concerns quadratic optimization in variables $x \in \mathbb{R}^n$ and $y \in \{0,1\}^n$, where $0 \leq x \leq y$. The y variables are refered to as *indicator* or *switching* variables and occur frequently in applications, including electrical power production [9, 11, 13], constrained portfolio optimization [9, 10, 19], nonlinear machine scheduling [1] and chemical pooling [7]. In the formulations of these problems the objective function is typically separable in x and y and cross-terms $y_i y_j$ with $i \neq j$ may or may not appear.

K.M. Anstreicher
Department of Business Analytics, University of Iowa, Iowa City, IA, 52242-1994, USA.
E-mail: kurt-anstreicher@uiowa.edu

S. Burer
Department of Business Analytics, University of Iowa, Iowa City, IA, 52242-1994, USA.
E-mail: samuel-burer@uiowa.edu

One approach for such problems is to consider symmetric matrix variables X and Y that replace the rank-1 matrices xx^T and yy^T, respectively. Using such variables, an objective of the form $c^Tx + x^TQx + y^TDy$ can be replaced by the linear function $c^Tx + Q \bullet X + D \bullet Y$, where (x, X, Y) should then be in the set

$$\mathscr{H} := \text{conv}\{(x, xx^T, yy^T) : 0 \le x \le y \in \{0,1\}^n\}.$$

The problem is then to represent $\mathscr{H}$ in a manner that is amenable to computation, e.g., with a polynomial-size outer formulation, disjunctive formulation, or convex relaxation. Note that, because y is binary, $\text{diag}(Y)$ captures y, and in particular, when the cross-terms y_iy_j are not of interest, we may consider the simpler convex hull

$$\mathscr{H}' := \text{conv}\{(x, xx^T, y) : 0 \le x \le y \in \{0,1\}^n\}.$$

For general n, determining computable representations of $\mathscr{H}$ and $\mathscr{H}'$ is difficult. For example, even when y is fixed to e, the resulting convex hull, called QPB in [5] for "quadratic programming over the box," is intractable. When $n = 2$, an exact representation for QPB was given in [2], but such a representation is not known for $n \ge 3$. For general n, the paper [8] studies valid inequalities for $\mathscr{H}'$. For the case $n = 1$, $\mathscr{H} = \mathscr{H}'$ since there are no cross-terms, and a computable representation was given in [10] based on prior work in [9]. This representation has subsequently been used in a variety of applications; see for example [11, 13]. Several authors have also studied the case when $n = 2$ but have focused on convexifying in the space of (x, y, γ), where γ is a scalar associated with the epigraph of a specially structured quadratic function, e.g., a convex quadratic one; see [3] and references therein.

In Section 2, we consider the case of $n = 1$ and reprove the representation of $\mathscr{H} = \mathscr{H}'$ in a new way which incorporates standard ideas from the literature on constructing strong semidefinite programming (SDP) relaxations of quadratic programs. In particular, our proof can be viewed as establishing that $\mathscr{H}$ for $n = 1$ is captured exactly by the relaxation which uses the standard positive semidefinite (PSD) condition along with the standard Reformulation–Linearization Technique (RLT) constraints [18].

Our main result in this paper is a representation of $\mathscr{H}$ for $n = 2$, which we derive in several steps. Note that in this case there is only a single cross-term y_1y_2, and we can write $\mathscr{H}$ in the form

$$\mathscr{H} = \text{conv}\{(x, xx^T, y, y_1y_2) : 0 \le x \le y \in \{0,1\}^2\}.$$

First, in Section 3, we give a disjunctive representation of $\mathscr{H}$ that involves additional variables $\alpha \in \mathbb{R}^2$, $\beta \in \mathbb{R}^2$ as well as 2 second-order cone (SOC) constraints and one 3×3 PSD condition. In Section 4 we project out β and remove the SOC constraints by replacing the single PSD constraint with four PSD constraints. The primary effort in the paper occurs in Section 5, where we show that it is in fact only necessary to impose one of these four PSD constraints in order to represent $\mathscr{H}$. This analysis is relatively complex due to the fact that we are attempting to characterize the projection of $(x, X, y, Y_{12}, \alpha)$ onto (x, X, y, Y_{12}) where the constraints on $(x, X, y, Y_{12}, \alpha)$ include PSD conditions. If all constraints on $(x, X, y, Y_{12}, \alpha)$ were linear, we could use standard polyhedral techniques such as Fourier-Motzkin elimination to perform this projection. However, since our case includes PSD conditions, we are unaware of any general methodolgy for characterizing such a projection, and therefore our proof technique is tailored to the structure of $\mathscr{H}$ for $n = 2$.

The final representation that we obtain for $\mathscr{H}$ retains the variables $\alpha \in \mathbb{R}^2$ and one 5×5 PSD constraint. From a computational standpoint this is only a modest improvement over the original disjunctive representation. However the primary goal of the paper is to better understand the exact structure of $\mathscr{H}$. To this end, in Section 6, we describe an alternative derivation for the representation of $\mathscr{H}$ obtained in Section 5. In this alternative derivation the α variables that appear in our representation arise naturally in order to complete a 5×5 PSD matrix V that is a relaxation of the rank-one matrix vv^T, where $v = (1, x_1, x_2, t_1, t_2)$ and $t = e - y$ is also binary. Moreover the PSD condition in our final representation of $\mathscr{H}$ is a strengthening of the condition $V \succeq 0$. This derivation also leads to a conjecture that the weaker condition $V \succeq 0$ is sufficient to characterize $\mathscr{H}'$, i.e., when there are no cross terms, for $n = 2$. If true, this conjecture would establish that $\mathscr{H}'$ can be represented using PSD, RLT, and simple linear conditions derived from the binary nature of y, thus generalizing the results of Section 2 for $n = 1$ as well as the representation of QPB for $n = 2$ from [2]. This conjecture is supported by extensive numerical computations but remains unproved.

Since our representation for $\mathscr{H}$ with $n = 2$ retains the variables $\alpha \in \mathbb{R}^2$, it is natural to wonder if there might be a representation of $\mathscr{H}$ that involves only convex constraints in the variables (x, X, y, Y_{12}). We have not pursued this topic for several reasons. First, as described above, the α variables arise naturally as elements of the 5×5 matrix $V \succeq 0$. Second, in situations where a convex set has a spectrahedral representation (that is, a representation involving PSD conditions and linear constraints), and an alternative representation using a set of convex constraints $f_i(\cdot) \le 0$, $i = 1, \ldots, m$ where all $f_i(\cdot)$ are convex, the representation using convex constraints may be surprisingly complex, possibly involving case conditions and/or a dissection of the domain of variables; see for example the explicit representations for quadratic optimization over a triangle or rectangle in $\mathbb{R}^2$ in [15] and [14], respectively, compared to the spectrahedral representations in [2]. Finally, it is known that a set represented as a projected spectahedron, such as our representation of $\mathscr{H}$ that includes the variables α, may fail to be a basic semialgebraic set [17, Lemma 3.14] and may therefore have no possible representation in terms of unquantified polynomial inequalities.

Although our analysis is focused on the case of $n = 2$, our results can also be applied to problems with larger n by utilizing the convex hull characterization for pairs of variables $x_i \le y_i$, $x_j \le y_j$, $i, j \in \{1, \ldots, n\}$. Since there are $O(n^2)$ such pairs, the required constraints would only be added *a priori* for small n; for larger n one could iteratively search for violated conditions and add them to strengthen the relaxation of $\mathscr{H}$. The addition of this information based on $n = 2$ is a promising approach to strengthen computational methods that have previously utilized only the convex hull representation for $n = 1$.

Notation. We use e to denote a vector of arbitrary dimension with each component equal to one, and e_i to denote an elementary vector with all components equal to zero except for a one in component i. For symmetric matrices X and Y, $X \succeq Y$ denotes that $X - Y$ is positive semidefinite (PSD) and $X \succ Y$ denotes that $X - Y$ is positive definite. The vector whose components are those of the diagonal entries of a matrix X is denoted $\text{diag}(X)$, and $\text{Diag}(x)$ puts the vector x into a square diagonal matrix. The convex hull of a set is denoted $\text{conv}\{\cdot\}$.

2 The convex hull for $n = 1$

In this section we consider the representation of $\mathscr{H}$ for $n = 1$; note that $\mathscr{H} = \mathscr{H}'$ in this case. The representation given in Theorem 1 below is known, but to our knowledge the proof given here is new. We define

$$\text{PER} := \left\{ (\alpha, \beta, \gamma) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R} : \begin{array}{l} \alpha^2 \le \beta\gamma \\ 0 \le \beta \le \alpha \le \gamma \end{array} \right\}$$

to be the so-called *perspective cone* in $\mathbb{R}^3$. In particular, the constraint $\alpha^2 \le \beta\gamma$ is called a *perspective* constraint in the literature [10].

Theorem 1 *For $n = 1$, $\mathscr{H} = \mathscr{H}' = \{(x_1, X_{11}, y_1) \in \text{PER} : y_1 \le 1\}$.*

Proof Let $t_1 = 1 - y_1$. Then the constraints $0 \le x_1 \le y_1, y_1 \in \{0,1\}$ can be written in the form $x_1 + s_1 + t_1 = 1$, $x_1 \ge 0$, $s_1 \ge 0$, $t_1 \in \{0,1\}$. By relaxing the rank-one matrix $(1, x_1, s_1, t_1)^T (1, x_1, s_1, t_1)$ we obtain a matrix

$$W = \begin{pmatrix} 1 & x_1 & s_1 & t_1 \\ x_1 & X_{11} & Z_{11} & 0 \\ s_1 & Z_{11} & S_{11} & 0 \\ t_1 & 0 & 0 & t_1 \end{pmatrix}, \tag{1}$$

where we are using the fact that, for binary t_1, it holds that $t_1^2 = t_1$ and $x_1 t_1 = s_1 t_1 = 0$. Multiplying $x_1 + s_1 + t_1 = 1$ in turn by the variables x_1 and s_1, we next obtain the RLT constraints $X_{11} + Z_{11} = x_1$ and $S_{11} + Z_{11} = s_1$. Let

$$\begin{aligned} \mathscr{C} &= \text{conv}\{(1, x_1, s_1, t_1)^T (1, x_1, s_1, t_1) : x_1 + s_1 + t_1 = 1, x_1 \ge 0, s_1 \ge 0, t_1 \in \{0,1\}\}, \\ \mathscr{D} &= \{W \in \text{DNN} : x_1 + s_1 + t_1 = 1, X_{11} + Z_{11} = x_1, S_{11} + Z_{11} = s_1\}, \end{aligned}$$

where the matrix W in the definition of $\mathscr{D}$ has the form (1), and DNN denotes the cone of doubly nonnegative matrices, that is, matrices that are both componentwise nonnegative and PSD. We claim that $\mathscr{C} = \mathscr{D}$. The inclusion $\mathscr{C} \subset \mathscr{D}$ is obvious by standard SDP-relaxation techniques. However, from [4, Corollary 2.5] we know that

$$\mathscr{C} = \{W \in \text{CP} : x_1 + s_1 + t_1 = 1,\ X_{11} + S_{11} + t_1 + 2Z_{11} = 1\},$$

where CP denotes the cone of completely positive matrices, that is, matrices that can be represented as a sum of nonnegative rank-one matrices. Note that $X_{11} + S_{11} + t_1 + 2Z_{11} = 1$ is the "squared" constraint obtained by substituting appropriate variables into the expression $(x_1 + s_1 + t_1)^2 = 1$. Then $\mathscr{C} = \mathscr{D}$ follows from the facts that since W is 4×4, $W \in \text{CP} \iff W \in \text{DNN}$ [16], and the constraints $x_1 + s_1 + t_1 = 1$, $X_{11} + Z_{11} = x_1$ and $S_{11} + Z_{11} = s_1$ together imply $X_{11} + S_{11} + t_1 + 2Z_{11} = 1$.

From $\mathscr{C} = \mathscr{D}$ we conclude that $\operatorname{conv}\{(x_1, x_1^2, y_1) : 0 \le x_1 \le y_1, y_1 \in \{0,1\}\} = \{(x_1, X_{11}, 1-t_1) : x_1 + s_1 + t_1 = 1, X_{11} + Z_{11} = x_1, S_{11} + Z_{11} = s_1, W \in \text{DNN}\}$. To complete the proof we will simplify the condition that $W \succeq 0$. Note that

$$W = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & x_1 & t_1 \\ x_1 & X_{11} & 0 \\ t_1 & 0 & t_1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}.$$

Then $W \succeq 0$ if and only if

$$\begin{pmatrix} 1 & x_1 & t_1 \\ x_1 & X_{11} & 0 \\ t_1 & 0 & t_1 \end{pmatrix} \succeq 0 \quad \Leftrightarrow \quad \begin{pmatrix} 1-t_1 & x_1 \\ x_1 & X_{11} \end{pmatrix} \succeq 0,$$

which using $y_1 = 1 - t_1$ is equivalent to $y_1 \ge 0$, $X_{11} \ge 0$, $y_1 X_{11} \ge x_1^2$. The conditions of the theorem thus ensure that $W \in \text{DNN}$, where $t_1 = 1 - y_1 \ge 0$, $s_1 = 1 - t_1 - x_1 = y_1 - x_1 \ge 0$, $Z_{11} = x_1 - X_{11} \ge 0$ and $S_{11} = 1 + X_{11} - 2x_1 - t_1 = y_1 + X_{11} - 2x_1 \ge 0$. □

Note that the characterization in Theorem 1 is sometimes written in terms of the lower convex envelope rather than the convex hull, in which case the condition $X_{11} \le x_1$ is omitted.

3 The disjunctive convex hull for $n = 2$

In this section, we develop an explicit disjunctive formulation for the convex hull $\mathscr{H}$ when $n = 2$. As described Section 1, we will use that fact that $\operatorname{diag}(Y) = y$ and that there is only one cross-term $y_1 y_2$ to write (x, X, y, Y_{12}) for points in $\mathscr{H}$.

The representation for $\mathscr{H}$ obtained in this section is based on the four values of $y \in \{0,1\}^2 = \{0, e_1, e_2, e\}$. Specifically, note that $\mathscr{H} = \operatorname{conv}(\mathscr{H}_0 \cup \mathscr{H}_{e_1} \cup \mathscr{H}_{e_2} \cup \mathscr{H}_e)$, where for each fixed y,

$$\mathscr{H}_y := \operatorname{conv}\left\{(x, xx^T, y, y_1 y_2) : 0 \le x \le y\right\}.$$

Each such $\mathscr{H}_y$ has a known representation. $\mathscr{H}_0$ is just a singleton, and for $y = e_1$ and $y = e_2$ representations based on PER are provided by Theorem 1. For $y = e$, a representation is given in [2] as follows. Define

$$\text{RLT}_x := \left\{ \begin{pmatrix} \lambda & x^T \\ x & X \end{pmatrix} : \begin{array}{l} \lambda \ge 0,\ 0 \le \operatorname{diag}(X) \le x \\ \max\{0, x_1 + x_2 - \lambda\} \le X_{12} \le \min\{x_1, x_2\} \end{array} \right\},$$

which is the homogenization of those points (x, X) satisfying the standard RLT constraints associated with $0 \le x \le e$. It is proven in [2] that

$$\mathscr{H}_e = \left\{ (x, X, y, Y_{12}) : \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \in \text{PSD} \cap \text{RLT}_x,\ y = e,\ Y_{12} = 1 \right\},$$

where PSD denotes the cone of positive semidefinite matrices. In the sequel we will also need

$$\mathrm{RLT}_y := \left\{ (y, Y_{12}) \in \mathbb{R}^2 \times \mathbb{R} : \max\{0, y_1 + y_2 - 1\} \le Y_{12} \le \min\{y_1, y_2\} \right\},$$

which gives the convex hull of $(y, y_1 y_2)$ over all four $y \in \{0,1\}^2$. Note that RLT_y is a polytope, unlike PER, RLT_x and PSD, which are convex cones.

In many applications, the product $y_1 y_2$ is not of interest, so it is also natural to consider the convex hull $\mathscr{H}'$ that ignores this product. Based on the known representations for $\mathscr{H}_{e_1}$, $\mathscr{H}_{e_2}$ and $\mathscr{H}_e$, $\mathscr{H}'$ is certainly contained in the set of (x, X, y) satisfying the constraints

$$\begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \in \mathrm{PSD} \cap \mathrm{RLT}_x$$
$$(x_j, X_{jj}, y_j) \in \mathrm{PER}, \quad y_j \le 1 \ \ \forall\, j = 1, 2.$$

However it is easy to generate examples that satisfy these constraints but are not in $\mathscr{H}'$. In the next theorem we will focus on $\mathscr{H}$, but we will return to a discussion of $\mathscr{H}'$ in Section 6.

Theorem 2 *For $n = 2$, let $\mathscr{H}^+$ be the set of all $(x, X, y, Y_{12}, \alpha)$ satisfying the convex constraints*

$$x \le y \tag{2a}$$
$$\begin{pmatrix} Y_{12} & (x-\alpha)^T \\ x-\alpha & X - \mathrm{Diag}(\beta) \end{pmatrix} \in \mathrm{PSD} \cap \mathrm{RLT}_x \tag{2b}$$
$$(\alpha_j, \beta_j, y_j - Y_{12}) \in \mathrm{PER} \ \ \forall\, j = 1, 2 \tag{2c}$$
$$(y, Y_{12}) \in \mathrm{RLT}_y \tag{2d}$$

where $\alpha \in \mathbb{R}^2$, $\beta \in \mathbb{R}^2$ are auxiliary variables. Then $\mathscr{H}$ equals the projection of $\mathscr{H}^+$ onto the variables (x, X, y, Y_{12}).

Proof We first argue that (2) is a relaxation of $\mathscr{H}$ in the lifted space that includes α and β. It suffices to show that each "rank-1" solution $(x, xx^T, y, y_1 y_2)$ for $y \in \{0,1\}^2$ can be extended in (α, β) to a feasible solution of (2). We handle the four cases for $y \in \{0,1\}^2$ separately. We clearly always have $x \le y$ and $(y, Y_{12}) \in \mathrm{RLT}_y$, so it remains to check that (2b) and (2c) hold in each case.

We introduce the notation

$$Z := \begin{pmatrix} Y_{12} & (x-\alpha)^T \\ x-\alpha & X - \mathrm{Diag}(\beta) \end{pmatrix}.$$

First, let $y = 0 \Rightarrow x = 0$. Then $(x, xx^T, y, y_1 y_2) = (0,0,0,0)$, and we choose $(\alpha, \beta) = (0,0)$. Since all variables are zero, it is straightforward to check that (2b) and (2c) are satisfied. Second, let $y = e \Rightarrow 0 \le x \le e$. Then $(x, xx^T, y, y_1 y_2) = (x, xx^T, e, 1)$, and we choose $(\alpha, \beta) = (0,0)$ for this case also, which yields $(\alpha_j, \beta_j, y_j - Y_{12}) = (0,0,0) \in$ PER for $j = 1, 2$. Moreover,

$$Z = \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} = \begin{pmatrix} 1 & x^T \\ x & xx^T \end{pmatrix} \in \mathrm{PSD} \cap \mathrm{RLT}_x,$$

as desired.

Next we consider the case $y = e_1$, which implies $x_1 \le 1$ and $x_2 = 0$. Then $(x, xx^T, y, y_1y_2) = (x_1e_1, x_1^2 e_1 e_1^T, e_1, 0)$, and we choose $(\alpha, \beta) = (x_1e_1, x_1^2 e_1)$. Hence,

$$Z = \begin{pmatrix} 0 & (x - x_1e_1)^T \\ x - x_1e_1 & X - x_1^2 e_1 e_1^T \end{pmatrix} = 0 \in \mathrm{PSD} \cap \mathrm{RLT}_x,$$

satisfying (2b). Moreover, $(\alpha_1, \beta_1, y_1 - y_1y_2) = (x_1, x_1^2, 1) \in \mathrm{PER}$ and $(\alpha_2, \beta_2, y_2 - y_1y_2) = (0, 0, 0) \in \mathrm{PER}$, so that (2c) is satisfied. The final case $y = e_2$ is similar. We have thus shown that (2) is a relaxation of $\mathscr{H}$.

To complete the proof, we show the reverse containment, i.e., that any member $(x, X, y, Y_{12}, \alpha, \beta)$ of $\mathscr{H}^+$ also satisfies $(x, X, y, Y_{12}) \in \mathscr{H}$. Define the four scalars

$$\lambda_0 := 1 - y_1 - y_2 + Y_{12}, \;\; \lambda_{e_1} := y_1 - Y_{12}, \;\; \lambda_{e_2} := y_2 - Y_{12}, \;\; \lambda_e := Y_{12}, \qquad (3)$$

satisfying $\lambda_0 + \lambda_{e_1} + \lambda_{e_2} + \lambda_e = 1$ and note that $(y, Y_{12}) \in \mathrm{RLT}_y$ implies each of these scalars is nonnegative. So $(\lambda_0, \lambda_{e_1}, \lambda_{e_2}, \lambda_e)$ is a convex combination. Next, letting $0/0 := 0$, define

$$Z_0 := \lambda_0^{-1} \begin{pmatrix} \lambda_0 & 0^T \\ 0 & 0 \end{pmatrix} \qquad Z_{e_2} := \lambda_{e_2}^{-1} \begin{pmatrix} \lambda_{e_2} & \alpha_2 e_2^T \\ \alpha_2 e_2 & \beta_2 e_2 e_2^T \end{pmatrix}$$
$$Z_{e_1} := \lambda_{e_1}^{-1} \begin{pmatrix} \lambda_{e_1} & \alpha_1 e_1^T \\ \alpha_1 e_1 & \beta_1 e_1 e_1^T \end{pmatrix} \qquad Z_e := \lambda_e^{-1} \begin{pmatrix} \lambda_e & (x - \alpha)^T \\ x - \alpha & X - \mathrm{Diag}(\beta) \end{pmatrix}.$$

Writing

$$Z_y =: \begin{pmatrix} 1 & x_y^T \\ x_y & X_y \end{pmatrix},$$

we note that $(x_y, X_y, y, y_1y_2) \in \mathscr{H}_y$ for each $y \in \{0,1\}^2$; for $y = e_1$ and $y = e_2$ we use the representation from Theorem 1, and for $y = e$ we use the result from [2] stated above this theorem. Hence, the easily verified equations $(y, Y_{12}) = \lambda_0(0,0) + \lambda_{e_1}(e_1, 0) + \lambda_{e_2}(e_2, 0) + \lambda_e(e, 1)$ and

$$\begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} = \lambda_0 Z_0 + \lambda_{e_1} Z_{e_1} + \lambda_{e_2} Z_{e_2} + \lambda_e Z_e,$$

establish that $(x, X, y, Y_{12}) \in \mathscr{H}$. □

Since $\mathscr{H} = \mathrm{conv}(\mathscr{H}_0 \cup \mathscr{H}_{e_1} \cup \mathscr{H}_{e_2} \cup H_e)$, where there is a known convex representation for each $\mathscr{H}_y := \mathrm{conv}\left\{(x, xx^T, y, y_1y_2) : 0 \le x \le y\right\}$, standard disjunctive programming techniques [6] could be used to obtain an extended formulation for $\mathscr{H}$. The representation given in Theorem 2, which is particulary compact, is convenient for our subsequent analysis.

4 Eliminating β

System (2) captures $\mathscr{H}$ by projection from a lifted space, which includes the additional variables $\alpha \in \mathbb{R}^2$, $\beta \in \mathbb{R}^2$. In this section, we eliminate the β variables from (2), but the price we pay is to replace the semidefinite constraint in (2b) with PSD conditions on four matrices. In Section 5 we will show that, in order to obtain a characterization of $\mathscr{H}$, it is in fact only necessary to impose one of these four PSD conditions.

We begin by introducing some notation. First, define the matrix function $M : \mathbb{R}^2 \times \mathbb{S}^2 \times \mathbb{R} \times \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{S}^3$ by

$$M(\beta) := M(x, X, Y_{12}, \alpha, \beta) := \begin{pmatrix} Y_{12} & (x-\alpha)^T \\ x-\alpha & X - \mathrm{Diag}(\beta) \end{pmatrix}. \tag{4}$$

The simplified notation $M(\beta)$ will be convenient because instances of M will only differ in the values of β; note also that M does not depend on y. We also define four different functions $\beta_{pq} : \mathbb{R}^2 \times \mathbb{S}^2 \times \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$ depending on $(x, X, y, Y_{12}, \alpha)$ for the indices $(p,q) \in \{1,2\}^2$, where $0/0 := 0$:

$$\begin{aligned}
\beta_{11} &:= \beta_{11}(x, X, y, Y_{12}, \alpha) := (X_{11} - x_1 + \alpha_1, X_{22} - x_2 + \alpha_2) \\
\beta_{12} &:= \beta_{12}(x, X, y, Y_{12}, \alpha) := \left((y_1 - Y_{12})^{-1}\alpha_1^2, X_{22} - x_2 + \alpha_2\right) \\
\beta_{21} &:= \beta_{21}(x, X, y, Y_{12}, \alpha) := \left(X_{11} - x_1 + \alpha_1, (y_2 - Y_{12})^{-1}\alpha_2^2\right) \\
\beta_{22} &:= \beta_{22}(x, X, y, Y_{12}, \alpha) := \left((y_1 - Y_{12})^{-1}\alpha_1^2, (y_2 - Y_{12})^{-1}\alpha_2^2\right).
\end{aligned}$$

As with $M(\beta)$, the shorter notation β_{pq} will prove more convenient. Note also that p and q are only index labels to designate the four functions. The result below replaces the PSD condition in (2b) with the four conditions $M(\beta_{pq}) \succeq 0$, $p,q \in \{1,2\}$.

Theorem 3 *For $n = 2$, let $\mathscr{H}^+$ be the set of all $(x, X, y, Y_{12}, \alpha)$ satisfying the convex constraints*

$$\begin{aligned}
&\mathrm{diag}(X) \le x \le y && \text{(5a)} \\
&\max\{0, x_1 - \alpha_1 + x_2 - \alpha_2 - Y_{12}\} \le X_{12} \le \min\{x_1 - \alpha_1, x_2 - \alpha_2\} && \text{(5b)} \\
&0 \le \alpha_j \le y_j - Y_{12} \ \ \forall\, j = 1,2 && \text{(5c)} \\
&(y, Y_{12}) \in \mathrm{RLT}_y && \text{(5d)} \\
&M(\beta_{11}) \succeq 0 && \text{(5e)} \\
&M(\beta_{12}) \succeq 0 && \text{(5f)} \\
&M(\beta_{21}) \succeq 0 && \text{(5g)} \\
&M(\beta_{22}) \succeq 0. && \text{(5h)}
\end{aligned}$$

Then $\mathscr{H}$ equals the projection of $\mathscr{H}^+$ onto the variables (x, X, y, Y_{12}).

Proof The proof is based on reformulating (2), which using $M(\beta)$ can be restated as

$$\begin{aligned}
&x \le y \\
&M(\beta) \in \text{PSD} \cap \text{RLT}_x \\
&(\alpha_j, \beta_j, y_j - Y_{12}) \in \text{PER} \ \ \forall\ j = 1,2 \\
&(y, Y_{12}) \in \text{RLT}_y.
\end{aligned}$$

In particular, considering $(x, X, y, Y_{12}, \alpha)$ fixed, the above system includes four linear conditions on β:

$$\beta_j \ge \max\left\{(y_j - Y_{12})^{-1}\alpha_j^2, X_{jj} - x_j + \alpha_j\right\} \quad \forall\ \ j = 1,2.$$

Moreover, since decreasing β_1 and β_2 while holding all other variables constant does not violate $M(\beta) \succeq 0$, we may define β_1 and β_2 by

$$\beta_j(x, X, y, Y_{12}, \alpha) := \max\left\{(y_j - Y_{12})^{-1}\alpha_j^2, X_{jj} - x_j + \alpha_j\right\} \quad \forall\ \ j = 1,2$$

without affecting the projection onto (x, X, y, Y_{12}). It follows that values $(x, X, y, Y_{12}, \alpha)$, which are feasible for (5a)–(5d), are feasible for the constraints (2) if and only if $M(\beta_{pq}) \succeq 0$, $(p,q) \in \{1,2\}^2$. □

In Section 5, we will show that in order to obtain an exact representation of $\mathscr{H}$ only the condition $M(\beta_{22}) \succeq 0$ is required. For clarity in the exposition it is helpful to write out the conditions $M(\beta_{pq}) \succeq 0$ explicitly. In particular, (5e) can be written

$$\begin{pmatrix} Y_{12} & x_1 - \alpha_1 & x_2 - \alpha_2 \\ x_1 - \alpha_1 & x_1 - \alpha_1 & X_{12} \\ x_2 - \alpha_2 & X_{12} & x_2 - \alpha_2 \end{pmatrix} \succeq 0. \tag{5e$'$}$$

In the remaining cases we can utilize the well-known Schur complement condition to conclude that (5f) is equivalent to

$$\begin{pmatrix} y_1 - Y_{12} & 0 & \alpha_1 & 0 \\ 0 & Y_{12} & x_1 - \alpha_1 & x_2 - \alpha_2 \\ \alpha_1 & x_1 - \alpha_1 & X_{11} & X_{12} \\ 0 & x_2 - \alpha_2 & X_{12} & x_2 - \alpha_2 \end{pmatrix} \succeq 0, \tag{5f$'$}$$

(5g) is equivalent to

$$\begin{pmatrix} y_2 - Y_{12} & 0 & 0 & \alpha_2 \\ 0 & Y_{12} & x_1 - \alpha_1 & x_2 - \alpha_2 \\ 0 & x_1 - \alpha_1 & x_1 - \alpha_1 & X_{12} \\ \alpha_2 & x_2 - \alpha_2 & X_{12} & X_{22} \end{pmatrix} \succeq 0, \tag{5g$'$}$$

and (5h) is equivalent to

$$\begin{pmatrix} y_1 - Y_{12} & 0 & 0 & \alpha_1 & 0 \\ 0 & y_2 - Y_{12} & 0 & 0 & \alpha_2 \\ 0 & 0 & Y_{12} & x_1 - \alpha_1 & x_2 - \alpha_2 \\ \alpha_1 & 0 & x_1 - \alpha_1 & X_{11} & X_{12} \\ 0 & \alpha_2 & x_2 - \alpha_2 & X_{12} & X_{22} \end{pmatrix} \succeq 0. \tag{5h$'$}$$

In the statement of results in the sequel we will always refer to the conditions (5e)–(5h), but these statements may be easier to understand if the reader refers to (5e$'$)–(5h$'$).

5 Reducing to a single semidefinite condition

Theorem 3 establishes that $\mathscr{H}$ is described in part by the four PSD conditions (5e)–(5h)—one of size 3×3, two of size 4×4, and one of size 5×5. In this section, we show that Theorem 3 holds even if (5e)–(5g) are not enforced. We show this in several steps. In section 5.1 we show that if (5a)–(5d) and (5h) hold then condition (5e) is redundant and at most one of the conditions (5f) and (5g) can fail to hold. In the sequel, when a point $(x, X, y, Y_{12}, \alpha)$ satisfies (5a)–(5d) and (5g)–(5h) but not (5f), we say that such an $(x, X, y, Y_{12}, \alpha)$ *lacks only* (5f).

In section 5.2 we show that if $(x, X, y, Y_{12}, \alpha)$ lacks only (5f) then it is always possible to construct another $(\bar{x}, \bar{X}, y, Y_{12}, \bar{\alpha})$ that lacks only (5f) but with $\bar{\alpha}_1 = 0$. In section 5.3 we consider how conditions (5f) and (5h) depend on α_2 when $(x, X, y, Y_{12}, \alpha)$ lacks only (5f) and $\alpha_1 = 0$. In section 5.4 we use the results of section 5.3 to show that if $(x, X, y, Y_{12}, \alpha)$ lacks only (5f) and $\alpha_1 = 0$, then it is always possible to construct $\hat{\alpha}$ with $\hat{\alpha}_1 = 0$ so that $(x, X, y, Y_{12}, \hat{\alpha})$ satisfies all of the conditions of (3). Finally in section 5.5 we give the final desired result that Theorem 3 holds even if (5e)–(5g) are not enforced.

5.1 Condition (5e) is redundant and at most one of (5f) and (5g) can fail

Lemma 1 *If $(x, X, y, Y_{12}, \alpha)$ satisfies (5a)–(5d), then it satisfies (5e).*

Proof Consider the linear conditions (5a)–(5d) of (5). In terms of the remaining variables, the constraints (5b) on X_{12} are simple bounds:

$$l := \max\{0, x_1 - \alpha_1 + x_2 - \alpha_2 - Y_{12}\} \leq X_{12} \leq \min\{x_1 - \alpha_1, x_2 - \alpha_2\} =: u.$$

We claim that (5e) is satisfed at both endpoints $X_{12} = l$ and $X_{12} = u$, which will prove the theorem since the determinant of every principal submatrix of $M(\beta_{11})$ that includes X_{12} is a concave quadratic function of X_{12}.

So we need $M(\beta_{11}) \succeq 0$ at both $X_{12} = l$ and $X_{12} = u$, i.e.,

$$\begin{pmatrix} Y_{12} & x_1 - \alpha_1 & x_2 - \alpha_2 \\ x_1 - \alpha_1 & x_1 - \alpha_1 & l \\ x_2 - \alpha_2 & l & x_2 - \alpha_2 \end{pmatrix} \succeq 0 \quad \text{and} \quad \begin{pmatrix} Y_{12} & x_1 - \alpha_1 & x_2 - \alpha_2 \\ x_1 - \alpha_1 & x_1 - \alpha_1 & u \\ x_2 - \alpha_2 & u & x_2 - \alpha_2 \end{pmatrix} \succeq 0.$$

The two matrices above share several properties necessary for positive semidefiniteness. Both have nonnegative diagonals, and all 2×2 principal minors are nonnegative:

- For each, the $\{1,2\}$ principal minor is nonnegative if and only if $Y_{12}(x_1 - \alpha_1) - (x_1 - \alpha_1)^2 \geq 0$. This follows from (5b):

  $$Y_{12} \geq (x_1 - \alpha_1) + (x_2 - \alpha_2 - X_{12}) \geq (x_1 - \alpha_1) + 0 = x_1 - \alpha_1, \tag{6}$$

  which implies $Y_{12}(x_1 - \alpha_1) \geq (x_1 - \alpha_1)^2$.
- For each, the $\{1,3\}$ principal minor is similarly nonnegative.

– The respective $\{2,3\}$ minors are nonnegative if $(x_1 - \alpha)(x_2 - \alpha_2) - l^2 \geq 0$ and $(x_1 - \alpha_1)(x_2 - \alpha_2) - u^2 \geq 0$, which hold because $0 \leq l \leq u \leq x_1 - \alpha_1$ and $0 \leq l \leq u \leq x_2 - \alpha_2$.

It remains to show that the determinants of both matrices are nonnegative. Let us first examine the case for $X_{12} = l$, which itself breaks into two subcases depending on whether the maximum in $l = \max\{0, x_1 - \alpha_1 + x_2 - \alpha_2 - Y_{12}\}$ is achieved in the first or second term. When the maximum is achieved in the first term, the determinant equals $(x_1 - \alpha_1)(x_2 - \alpha_2)(Y_{12} - x_1 + \alpha_1 - x_2 + \alpha_2)$, which is the product of three nonnegative terms. When the maximum is achieved in the second term, the determinant equals

$$(Y_{12} - x_2 + \alpha_2)(Y_{12} - x_1 + \alpha_1)(x_1 - \alpha_1 + x_2 - \alpha_2 - Y_{12})$$

which is also the product of three nonnegative terms; in particular, see (6). The case for $X_{12} = u$ similarly breaks down into two subcases, which mirror (i) and (ii) above. □

From Lemma 1 it is not necessary to enforce (5e). We next show that at most one of (5f) and (5g) can fail to hold.

Lemma 2 *Assume that* $(x, X, y, Y_{12}, \alpha)$ *satisfies* (5a)–(5d) *and* (5h). *Then at most one of (5f) and (5g) fails to hold.*

Proof Note that the matrices $M(\beta)$, $\beta \in \{\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}\}$ differ only in the components of $\text{Diag}(\beta)$. As a result, if $X_{11} - \alpha_1^2/(y_1 - Y_{12}) \geq x_1 - \alpha_1$ then (5e) implies (5f), and if $X_{22} - \alpha_2^2/(y_2 - Y_{12}) \geq x_2 - \alpha_2$ then (5e) implies (5g). On the other hand, if $X_{11} - \alpha_1^2/(y_1 - Y_{12}) \leq x_1 - \alpha_1$ then (5h) implies (5g), and if $X_{22} - \alpha_2^2/(y_2 - Y_{12}) \leq x_2 - \alpha_2$ then (5h) implies (5f). Since (5e) holds by Lemma 1 and (5h) holds by assumption, this means that if (5f) fails then $X_{11} - \alpha_1^2/(y_1 - Y_{12}) < x_1 - \alpha_1$ and $X_{22} - \alpha_2^2/(y_2 - Y_{12}) > x_2 - \alpha_2$, while if (5g) fails then $X_{22} - \alpha_2^2/(y_2 - Y_{12}) < x_2 - \alpha_2$ and $X_{11} - \alpha_1^2/(y_1 - Y_{12}) > x_1 - \alpha_1$. It is then obviously impossible for both (5f) and (5g) to fail. □

Our goal is to show that in Theorem 3 it is not necessary to enforce (5f) or (5g) once we enforce (5a)–(5d) and (5h). If $(x, X, y, Y_{12}, \alpha)$ satisfies (5a)–(5d) and (5f)–(5h) then by Lemma 1 there is nothing left to show, and by Lemma 2 it is not possible for both (5f) and (5g) to fail. In the sequel we will assume without loss of generality that (5g) holds but (5f) fails to hold, implying that $X_{11} - \alpha_1^2/(y_1 - Y_{12}) < x_1 - \alpha_1$ and $X_{22} - \alpha_2^2/(y_2 - Y_{12}) \geq x_2 - \alpha_2$ (see the proof of Lemma 2). To describe this situation, recall the following terminology regarding system (5): we say that a point $(x, X, y, Y_{12}, \alpha)$ *lacks only* (5f) when the point satisfies all conditions in (5) except that it violates (5f).

5.2 Reduction to $\alpha_1 = 0$

We next show that given $(x, X, y, Y_{12}, \alpha)$ that lacks only (5f) with $\alpha_1 > 0$ it is always possible to construct another $(\bar{x}, \bar{X}, y, Y_{12}, \bar{\alpha})$ that lacks only (5f), but with $\bar{\alpha}_1 = 0$. Most of the subsequent analysis will be based on such a $(\bar{x}, \bar{X}, y, Y_{12}, \bar{\alpha})$.

Lemma 3 *Suppose that* (x,X,y,Y_{12},α) *lacks only (5f), and suppose* $\alpha_1 > 0$. *Then* $y_1 - Y_{12} > 0$ *and* $(\bar{x},\bar{X},y,Y_{12},\bar{\alpha})$ *lacks only (5f), where*

$$\bar{x} := \begin{pmatrix} x_1-\alpha_1 \\ x_2 \end{pmatrix}, \quad \bar{X} := \begin{pmatrix} X_{11}-\alpha_1^2/(y_1-Y_{12}) & X_{12} \\ X_{12} & X_{22} \end{pmatrix}, \quad \bar{\alpha} := \begin{pmatrix} 0 \\ \alpha_2 \end{pmatrix}.$$

Proof If $\alpha_1 > 0$ then (5h) implies that $y_1 - Y_{12} > 0$. For notational convenience, define $v := (x,X,y,Y_{12},\alpha)$ and $\bar{v} := (\bar{x},\bar{X},y,Y_{12},\bar{\alpha})$. We need to check that $\bar{v}$ satisfies all conditions in (5) except (5f). Since only $\bar{x}_1$, $\bar{X}_{11}$, and $\bar{\alpha}_1$ differ between v and $\bar{v}$, and since $\bar{x}_1 - \bar{\alpha}_1 = x_1 - \alpha_1$, we need to verify $\bar{X}_{11} \le \bar{x}_1 \le y_1$, $0 \le \bar{\alpha}_1 \le y_1 - Y_{12}$, and (5h) at $\bar{v}$, and we need to show (5f) does *not* hold at $\bar{v}$. Clearly $0 \le \bar{\alpha}_1 \le y_1 - Y_{12}$ because $\bar{\alpha}_1 = 0$, and $\bar{x}_1 \le x_1 \le y_1$.

With $\bar{\alpha}_1 = 0$ and $\bar{x}_1 = x_1 - \alpha_1$, conditions (5e) and (5f) at $\bar{v}$ are respectively equivalent to

$$\begin{pmatrix} Y_{12} & \bar{x}_1 & x_2-\alpha_2 \\ \bar{x}_1 & \bar{x}_1 & X_{12} \\ x_2-\alpha_2 & X_{12} & x_2-\alpha_2 \end{pmatrix} = \begin{pmatrix} Y_{12} & x_1-\alpha_1 & x_2-\alpha_2 \\ x_1-\alpha_1 & x_1-\alpha_1 & X_{12} \\ x_2-\alpha_2 & X_{12} & x_2-\alpha_2 \end{pmatrix} \succeq 0,$$

and

$$\begin{pmatrix} Y_{12} & \bar{x}_1 & x_2-\alpha_2 \\ \bar{x}_1 & \bar{X}_{11} & X_{12} \\ x_2-\alpha_2 & X_{12} & x_2-\alpha_2 \end{pmatrix} = \begin{pmatrix} Y_{12} & x_1-\alpha_1 & x_2-\alpha_2 \\ x_1-\alpha_1 & X_{11}-\alpha_1^2/(y_1-Y_{12}) & X_{12} \\ x_2-\alpha_2 & X_{12} & x_2-\alpha_2 \end{pmatrix} \succeq 0.$$

These conditions both match the conditions of (5e) and (5f) at v, showing that (5e) holds at v if and only if (5e) holds at $\bar{v}$, and similarly for (5f). In particular, this implies $\bar{v}$ does not satisfy (5f), as desired. In addition, we conclude $\bar{X}_{11} \le \bar{x}_1$ because, if $\bar{X}_{11}$ were greater than $\bar{x}_1$, then (5e) holding at v would imply (5f) holds at v by just comparing the diagonal elements above, but this would violate our assumptions.

Finally, using again the relationship between $\bar{v}$ and v, (5h) holds at $\bar{v}$ if and only if

$$\begin{pmatrix} y_2-Y_{12} & 0 & 0 & \alpha_2 \\ 0 & Y_{12} & x_1-\alpha_1 & x_2-\alpha_2 \\ 0 & x_1-\alpha_1 & X_{11}-\alpha_1^2/(y_1-Y_{12}) & X_{12} \\ \alpha_2 & x_2-\alpha_2 & X_{12} & X_{22} \end{pmatrix} \succeq 0,$$

which is true by applying the Schur complement, using the fact that (5h) holds at v. □

5.3 Characterizing (5f) and (5h) in terms of α_2

Given (x,X,y,Y_{12},α) with $\alpha_1 = 0$ that lacks only (5f), in Section 5.4 our goal will be to modify α_2 to a new value $\hat{\alpha}_2$ so as to satisfy all the constraints of (5). To facilitate this analysis, we now examine how conditions (5f) and (5h) depend on α_2. Define

$$\theta := Y_{12}X_{11} - x_1^2 \ge 0. \tag{7}$$

Lemma 4 *Suppose $(x, X, y, Y_{12}, \alpha)$ lacks only (5f), where $\alpha_1 = 0$. Then $X_{11} > 0$ and $y_2 - Y_{12} > 0$. Moreover, if $(x, X, y, Y_{12}, \hat{\alpha})$ satisfies (5f) where $\hat{\alpha} := (0, \hat{\alpha}_2)$, then $\hat{\alpha}_2 \in [\alpha_2^-, \alpha_2^+]$, where*

$$\alpha_2^- := x_2 - \frac{X_{12}x_1}{X_{11}} - \frac{\theta + \sqrt{\theta(\theta + 4X_{12}(x_1 - X_{12}))}}{2X_{11}} \le x_2 - \frac{X_{12}x_1}{X_{11}} - \frac{\theta}{X_{11}} \quad \text{(8a)}$$

$$\alpha_2^+ := x_2 - \frac{X_{12}x_1}{X_{11}} - \frac{\theta - \sqrt{\theta(\theta + 4X_{12}(x_1 - X_{12}))}}{2X_{11}} \ge x_2 - \frac{X_{12}x_1}{X_{11}}. \quad \text{(8b)}$$

Proof Note that if $(x, X, y, Y_{12}, \alpha)$ with $\alpha_1 = 0$ satisfies (5h), then $X_{11} = 0$ implies that $x_1 = X_{12} = 0$. In this case (5f) follows immediately from (5b). In addition, if $y_2 - Y_{12} = 0$ then (5h) implies that $\alpha_2 = 0$, in which case (5f) would follow immediately from $X_{22} \le x_2$. Thus if $(x, X, y, Y_{12}, \alpha)$ with $\alpha_1 = 0$ lacks only (5f) we must have $X_{11} > 0$ and $y_2 - Y_{12} > 0$.

Because $y_1 - Y_{12} \ge 0$ and $\alpha_1 = 0$, (5f) is equivalent to

$$Z := \begin{pmatrix} Y_{12} & x_1 & x_2 - \alpha_2 \\ x_1 & X_{11} & X_{12} \\ x_2 - \alpha_2 & X_{12} & x_2 - \alpha_2 \end{pmatrix} \succeq 0. \quad (9)$$

Letting $\bar{x}_2 := x_2 - \alpha_2$, we have $\det(Z) = -X_{11}\bar{x}_2^2 + (2X_{12}x_1 + Y_{12}X_{11} - x_1^2)\bar{x}_2 - Y_{12}X_{12}^2$. As a function of $\bar{x}_2$, this is a strictly concave quadratic since $X_{11} > 0$. Moreover, the discriminant for this quadratic is

$$\begin{aligned} &(Y_{12}X_{11} - x_1^2 + 2x_1X_{12})^2 - 4Y_{12}X_{11}X_{12}^2 \\ &= (Y_{12}X_{11} - x_1^2)^2 + 4x_1X_{12}(Y_{12}X_{11} - x_1^2) + 4x_1^2X_{12}^2 - 4Y_{12}X_{11}X_{12}^2 \\ &= (Y_{12}X_{11} - x_1^2)^2 + 4x_1^2X_{12}(X_{12} - x_1) + 4Y_{12}X_{11}X_{12}(x_1 - X_{12}) \\ &= (Y_{12}X_{11} - x_1^2)^2 + 4X_{12}(x_1 - X_{12})(Y_{12}X_{11} - x_1^2) \\ &= \theta(\theta + 4X_{12}(x_1 - X_{12})). \end{aligned}$$

It follows that $\det(Z) \ge 0$ if and only if $\bar{x}_2$ is contained in the interval bounded by the roots

$$\frac{X_{12}x_1}{X_{11}} + \frac{\theta \pm \sqrt{\theta(\theta + 4X_{12}(x_1 - X_{12}))}}{2X_{11}},$$

or equivalently, if and only if $\alpha_2 \in [\alpha_2^-, \alpha_2^+]$. Therefore if $(x, X, y, Y_{12}, \hat{\alpha})$ satisfies (5f) we must have $\hat{\alpha}_2 \in [\alpha_2^-, \alpha_2^+]$. The inequalities in (8a) and (8b) are used in the sequel; note that the inequality in (8b) follows from the fact that $\theta(\theta + 4X_{12}(x_1 - X_{12})) \ge \theta^2$ since $X_{12} \le x_1$. □

From the above lemma, if $(x, X, y, Y_{12}, \alpha)$ lacks only (5f), where $\alpha_1 = 0$, then to have $(x, X, y, Y_{12}, \hat{\alpha})$ satisfy (5f) with $\hat{\alpha}_1 = 0$ we certainly require that $\hat{\alpha}_2 \in [\alpha_2^-, \alpha_2^+]$. In the next lemma we show that in fact this condition is necessary and sufficient.

Lemma 5 *Suppose $(x, X, y, Y_{12}, \alpha)$ lacks only (5f), where $\alpha_1 = 0$, and let $\hat{\alpha} := (0, \hat{\alpha}_2)$. Then $(x, X, y, Y_{12}, \hat{\alpha})$ satisfies (5f) if and only if $\hat{\alpha}_2 \in [\alpha_2^-, \alpha_2^+]$.*

Proof We consider $\hat{Z}$ defined in (9) but with $\hat{\alpha}_2$ substituted for α_2; we wish to show $\hat{Z} \succeq 0$ if and only if $\hat{\alpha}_2 \in [\alpha_2^-, \alpha_2^+]$. From Lemma 4 we know that $\det(\hat{Z}) \geq 0$ for such $\hat{\alpha}_2$, but it could happen that $\hat{Z} \not\succeq 0$ even when $\det(\hat{Z}) \geq 0$. Note that, since $(x, X, y, Y_{12}, \alpha)$ satisfies (5h) by assumption, then by the eigenvalue interlacing theorem (see, for example, Theorem 4.3.8 of Horn and Johnson [12]), $\hat{Z}$ has at most one negative eigenvalue.

We consider two cases based on whether $\theta \geq 0$ in (7) is positive or zero. If $\theta > 0$, then by the determinant and discriminant formulas above we have $\det(\hat{Z}) > 0 \Rightarrow \hat{Z} \succ 0$ for $\hat{\alpha}_2 \in (\alpha_2^-, \alpha_2^+)$, and $\hat{Z} \succeq 0$ with $\det(\hat{Z}) = 0$ when $\hat{\alpha}_2 = \alpha_2^-$ or $\hat{\alpha}_2 = \alpha_2^+$. The latter follows, for example, by continuity of the determinants of all principal submatrices. On the other hand, if $\theta = 0$, then $\alpha_2^- = \alpha_2^+ = x_2 - X_{12}x_1/X_{11}$, $\det(\hat{Z}) = 0$ when $\hat{\alpha}_2 = x_2 - X_{12}x_1/X_{11}$ and $\det(\hat{Z}) < 0$ for any other value of $\hat{\alpha}_2$. Focusing then on $\hat{\alpha}_2 = x_2 - X_{12}x_1/X_{11}$, we have

$$\hat{Z} = \begin{pmatrix} Y_{12} & x_1 & X_{12}x_1/X_{11} \\ x_1 & X_{11} & X_{12} \\ X_{12}x_1/X_{11} & X_{12} & X_{12}x_1/X_{11} \end{pmatrix}.$$

In this case $\operatorname{diag}(\hat{Z}) \geq 0$ and $\det(\hat{Z}) = 0$, so to demonstrate $\hat{Z} \succeq 0$, we need to show that the 2×2 principal submatrices are positive semidefinite or equivalently have nonnegative determinants. The $\{1,2\}$ submatrix is positive semidefinite since (5h) is satisfied; the determinant of the $\{1,3\}$ submatrix is nonnegative because $Y_{12}X_{11} \geq x_1^2 \geq X_{12}x_1$; and the determinant of the $\{2,3\}$ submatrix is nonnegative because $x_1 \geq X_{12}$. □

The next lemma considers how (5h) depends on α_2 when $\alpha_1 = 0$ and $(x, X, y, Y_{12}, \alpha)$ lacks only (5f).

Lemma 6 *Suppose* $(x, X, y, Y_{12}, \alpha)$ *lacks only* (5f)*, where* $\alpha_1 = 0$*. Define*

$$\alpha_2^* := \frac{(y_2 - Y_{12})(x_2 X_{11} - x_1 X_{12})}{y_2 X_{11} - x_1^2} = \left(x_2 - \frac{X_{12}x_1}{X_{11}}\right) \frac{y_2 - Y_{12}}{y_2 - x_1^2/X_{11}}, \tag{10}$$

and let $\hat{\alpha} := (0, \hat{\alpha}_2) \in \mathbb{R}^2$ *denote a vector variable with first entry equal to 0. Then if* $\alpha_2^* \geq \alpha_2$*,* $(x, X, y, Y_{12}, \hat{\alpha})$ *satisfies (5h) for any* $\hat{\alpha}_2 \in [\alpha_2, \alpha_2^*]$*, and if* $\alpha_2^* \leq \alpha_2$*,* $(x, X, y, Y_{12}, \hat{\alpha})$ *satisfies (5h) for any* $\hat{\alpha}_2 \in [\alpha_2^*, \alpha_2]$*.*

Proof When $\alpha_1 = 0$, (5h) is equivalent to

$$U := \begin{pmatrix} y_2 - Y_{12} & 0 & 0 & \alpha_2 \\ 0 & Y_{12} & x_1 & x_2 - \alpha_2 \\ 0 & x_1 & X_{11} & X_{12} \\ \alpha_2 & x_2 - \alpha_2 & X_{12} & X_{22} \end{pmatrix} \succeq 0,$$

implying that $\det(U) \geq 0$. When $X_{11} > 0$ and $y_2 - Y_{12} > 0$, which both hold from Lemma 4, it is straightforward to show that $\det(U)$ is a strictly concave quadratic function of α_2, and the maximizer of this determinant is α_2^*. Note that the denominator $y_2 X_{11} - x_1^2$ in (10) is strictly positive since $Y_{12}X_{11} \geq x_1^2$ and $y_2 > Y_{12}$.

Assume that $\alpha_2 < \alpha_2^*$. Let $\hat{\alpha}_2 \in (\alpha_2, \alpha_2^*]$ and consider $\hat{U}$ equal to U but with $\hat{\alpha}_2$ in place of α_2. Then $\det(\hat{U}) > 0$ since $(x, X, y, Y_{12}, \alpha)$ satisfies (5h) and $\det(\hat{U})$ is a strictly concave quadratic function of $\hat{\alpha}_2$. However this implies that $\hat{U} \succeq 0$, since by eigenvalue interlacing $\hat{U}$ can have at most one negative eigenvalue. The argument when $\alpha_2 > \alpha_2^*$ is similar. □

Finally, for $\alpha_1 = 0$, Lemma 7 considers conditions under which (5f) ⇒ (5h), and (5h) ⇒ (5f).

Lemma 7 *Let $(x, X, y, Y_{12}, \alpha)$ be given with $\alpha_1 = 0$, $y_2 - Y_{12} > 0$ and $0 \le x_2 - X_{22} \le \frac{1}{4}(y_2 - Y_{12})$. Define $\rho := \sqrt{1 - 4(x_2 - X_{22})/(y_2 - Y_{12})} \le 1$. Also define*

$$\lambda^- := \tfrac{1}{2}(1-\rho)(y_2 - Y_{12}) \le \tfrac{1}{2}(1+\rho)(y_2 - Y_{12}) =: \lambda^+.$$

Then $\lambda^- \le \alpha_2 \le \lambda^+$ ensures (5f) ⇒ (5h), and $\alpha_2 \le \lambda^-$ or $\lambda^+ \le \alpha_2$ ensures (5h) ⇒ (5f).

Proof Comparing diagonal elements of $M(\beta)$ for $\beta = \beta_{12}$ and $\beta = \beta_{22}$, similar to the proof of Lemma 2, we see that: (i) (5f) ⇒ (5h) is ensured when $x_2 - \alpha_2 \le X_{22} - \alpha_2^2/(y_2 - Y_{12})$; and (ii) (5h) ⇒ (5f) is ensured when the reverse inequality $x_2 - \alpha_2 \ge X_{22} - \alpha_2^2/(y_2 - Y_{12})$ holds. Note that λ^- and λ^+ are the roots of the quadratic equation $x_2 - \alpha_2 = X_{22} - \alpha_2^2/(y_2 - Y_{12})$ in α_2. In particular, the assumption $0 \le x_2 - X_{22} \le \frac{1}{4}(y_2 - Y_{12})$ guarantees that the discriminant is nonnegative and that $x_2 - \alpha_2 \le X_{22} - \alpha_2^2/(y_2 - Y_{12})$ is satisfied at the midpoint $\frac{1}{2}(y_2 - Y_{12})$ of λ^- and λ^+. Then the final statement of the lemma is just the restatement of (i) and (ii). □

5.4 Adjusting α_2 when $\alpha_1 = 0$

Assume that $(x, X, y, Y_{12}, \alpha)$ lacks only (5f) with $\alpha_1 = 0$. Then by Lemma 5 either $\alpha_2 < \alpha_2^-$ or $\alpha_2 > \alpha_2^+$; see (8) for the definitions of α_2^- and α_2^+. The next two lemmas show that $(x, X, y, Y_{12}, \hat{\alpha})$ then satisfies (5), where in the first case $\hat{\alpha} = (0, \alpha_2^-)$ and in the second case $\hat{\alpha} = (0, \alpha_2^+)$.

Lemma 8 *Assume that $(x, X, y, Y_{12}, \alpha)$ lacks only (5f) with $\alpha_1 = 0$, and $\alpha_2 < \alpha_2^-$. Then $(x, X, y, Y_{12}, \hat{\alpha})$ satisfies (5) with $\hat{\alpha} = (0, \alpha_2^-)$.*

Proof From Lemma 5 we know that $X_{11} > 0$, $y_2 - Y_{12} > 0$ and $(x, X, y, Y_{12}, \hat{\alpha})$ satisfies (5f). Since (5a)–(5d) ⇒ (5e) by Proposition 1 and (5h) ⇒ (5g) when $\alpha_1 = 0$ by inspection, we need to establish just (5a)–(5d) and (5h). Since $(x, X, y, Y_{12}, \alpha)$ satisfies (5a)–(5d) and we have increased α_2 to α_2^- to form $\hat{\alpha}$, we need only show $\alpha_2^- \le x_2 - X_{12}$ and $\alpha_2^- \le y_2 - Y_{12}$ to establish that (5a)–(5d) hold for $(x, X, y, Y_{12}, \hat{\alpha})$. In fact, we will show $\alpha_2^- \le x_2 - X_{12}$ as well as the stronger inequality $\alpha_2^- \le \lambda^+$, where $\lambda^+ = \frac{1}{2}(1+\rho)(y_2 - Y_{12})$ and $0 \le \rho \le 1$ are defined in Lemma 7. Indeed, the conditions of Lemma 7 hold here because, as (5h) is satisfied but (5f) is violated at α_2, we have $x_2 - \alpha_2 \le X_{22} - \alpha_2^2/(y_2 - Y_{12})$, which ensures $0 \le x_2 - X_{22} \le \frac{1}{4}(y_2 - Y_{12})$ and $\alpha_2 \le \lambda^+$. Hence, proving $\alpha_2^- \le \lambda^+$ will ensure (5f) ⇒ (5h).

To begin, we have

$$\alpha_2^- \leq x_2 - \frac{X_{12}x_1}{X_{11}} \leq x_2 - X_{12},$$

where the first inequality uses (8a) and the second uses $x_1 \geq X_{11}$. Next, to prove $\alpha_2^- \leq \lambda^+$, assume for contradiction that $\alpha_2 \leq \lambda^+ < \alpha_2^-$. Consider α_2^* as defined in (10). We claim $\lambda^+ < \alpha_2^*$, which from (10) is equivalent to

$$x_2 - \frac{X_{12}x_1}{X_{11}} > \tfrac{1}{2}(1+\rho)\left(y_2 - \frac{x_1^2}{X_{11}}\right).$$

We then have the chain

$$\begin{aligned}
x_2 - \frac{X_{12}x_1}{X_{11}} &\geq \alpha_2^- + \frac{\theta}{X_{11}} \\
&> \tfrac{1}{2}(1+\rho)(y_2 - Y_{12}) + \left(Y_{12} - \frac{x_1^2}{X_{11}}\right) \\
&\geq \tfrac{1}{2}(1+\rho)(y_2 - Y_{12}) + \tfrac{1}{2}(1+\rho)\left(Y_{12} - \frac{x_1^2}{X_{11}}\right) \\
&= \tfrac{1}{2}(1+\rho)\left(y_2 - \frac{x_1^2}{X_{11}}\right),
\end{aligned}$$

where the first inequality holds due to (8a), the second due to the definition (7) of θ and the assumption that $\lambda^+ < \alpha_2^-$, and the third because $\rho \leq 1$. Then (5h) holds with α_2 replaced by λ^+, by Lemma 6. However Lemma 7 then implies that (5f) also then holds with α_2 replaced by λ^+, and therefore $\alpha_2^- \leq \lambda^+$ from Lemma 5. This is the desired contradiction of $\lambda^+ < \alpha_2^-$. We must therefore have $\alpha_2^- \leq \lambda^+$, which completes the proof. □

Lemma 9 *Assume $(x, X, y, Y_{12}, \alpha)$ lacks only (5f) with $\alpha_1 = 0$, and $\alpha_2 > \alpha_2^+$. Then $(x, X, y, Y_{12}, \hat{\alpha})$ satisfies (5) with $\hat{\alpha} = (0, \alpha_2^+)$.*

Proof We follow a similar proof as for the preceding lemma. In this case, however, since we are decreasing α_2 to α_2^+, we need to show $\alpha_2^+ \geq x_1 + x_2 - X_{12} - Y_{12}$ and $\alpha_2^+ \geq \lambda^-$, where $\lambda^- = \frac{1}{2}(1-\rho)(y_2 - Y_{12})$ as defined in Lemma 7. Note that $\alpha_2 \geq \lambda^-$ because $(x, X, y, Y_{12}, \alpha)$ lacks only (5f), just as in the preceding lemma.

For the first inequality, from (8b) it suffices to show

$$x_2 - \frac{X_{12}x_1}{X_{11}} \geq x_1 + x_2 - X_{12} - Y_{12}$$

which is equivalent to

$$X_{12}x_1 + X_{11}x_1 - X_{11}X_{12} \leq X_{11}Y_{12}.$$

Since $\theta = Y_{12}X_{11} - x_1^2 \geq 0$ by (7), it thus suffices to show

$$\begin{aligned}
X_{12}x_1 + X_{11}x_1 - X_{11}X_{12} &\leq x_1^2 \\
X_{12}(x_1 - X_{11}) &\leq x_1(x_1 - X_{11}),
\end{aligned}$$

which certainly holds because $X_{12} \le x_1$ and $X_{11} \le x_1$.

For the second inequality, assume by contradiction that $\alpha_2^+ < \lambda^-$. We claim $\alpha_2^* < \lambda^-$, which by (10) is equivalent to

$$x_2 - \frac{X_{12}x_1}{X_{11}} < \tfrac{1}{2}(1-\rho)\left(y_2 - \frac{x_1^2}{X_{11}}\right).$$

From (8b), the assumption $\alpha_2^+ < \lambda^-$, and the inequality $Y_{12}X_{11} \ge x_1^2$, we have

$$x_2 - \frac{X_{12}x_1}{X_{11}} \le \alpha_2^+ < \lambda^- = \tfrac{1}{2}(1-\rho)(y_2 - Y_{12}) \le \tfrac{1}{2}(1-\rho)\left(y_2 - \frac{x_1^2}{X_{11}}\right),$$

as desired. Since $\alpha_2^* \le \lambda^- \le \alpha_2$, Lemma 6 implies that (5h) holds with α_2 replaced by λ^-. Lemma 7 then implies that (5f) also then holds with α_2 replaced by λ^-, and therefore $\alpha_2^+ \ge \lambda^-$ from Lemma 5. This contradicts the assumption that $\alpha_2^+ < \lambda^-$, as desired. □

5.5 Removing (5f) and (5g) does not affect the projection

We can now prove the following streamlined version of Theorem 3, which requires only one of the four PSD conditions (5e)–(5h).

Theorem 4 *For $n = 2$, let $\mathscr{H}^+$ be the set of all $(x, X, y, Y_{12}, \alpha)$ satisfying the convex constraints (5a)–(5d) and (5h). Then $\mathscr{H}$ equals the projection of $\mathscr{H}^+$ onto the variables (x, X, y, Y_{12}).*

Proof We must show that if $(x, X, y, Y_{12}, \alpha)$ satisfies (5a)–(5d) and (5h), then $(x, X, y, Y_{12}) \in \mathscr{H}$. By Theorem 3 this is equivalent to showing that there is an α' so that $(x, X, y, Y_{12}, \alpha')$ satisfies all of the constraints in (5).

If (5a)–(5d) are satisfied, then (5e) is redundant by Proposition 1. Moreover, as described above Lemma 3, if (5h) also holds then at most one of (5f)–(5g) can fail to hold. If both (5f)–(5g) hold then there is nothing to show, so we assume without loss of generality that (5f) fails to hold; that is, $(x, X, y, Y_{12}, \alpha)$ lacks only (5f).

Assume first that $\alpha_1 = 0$. If $\alpha_2 < \alpha_2^-$, then by Lemma 8 we know that $(x, X, y, Y_{12}, \hat{\alpha})$ satisfies (5), where $\hat{\alpha} = (0, \alpha_2^-)$. Similarly, if $\alpha_2 > \alpha_2^+$, then by Lemma 9 we have the same conclusion using $\hat{\alpha} = (0, \alpha_2^+)$. Therefore $(x, X, y, Y_{12}) \in \mathscr{H}$.

If $\alpha_1 > 0$ we apply the transformation in Lemma 3 to obtain $(\bar{x}, \bar{X}, y, Y_{12}, \bar{\alpha})$, with $\bar{\alpha} = (0, \alpha_2)$, that lacks only (5f). We then apply either Lemma 8 or Lemma 9 to obtain $\hat{\alpha} = (0, \hat{\alpha}_2)$ so that $(\bar{x}, \bar{X}, y, Y_{12}, \hat{\alpha})$ satisfies (5). Let $\alpha' = (\alpha_1, \hat{\alpha}_2)$. We claim that $(x, X, y, Y_{12}, \alpha')$ satisfies (5) as well. For the linear conditions (5a)–(5d) this is immediate from the facts that both $(x, X, y, Y_{12}, \alpha)$ and $(\bar{x}, \bar{X}, y, Y_{12}, \hat{\alpha})$ satisfy (5a)–(5d), and $\bar{x}_1 - \bar{\alpha}_1 = x_1 - \alpha_1$. Therefore (5e) is also satisfied at $(x, X, y, Y_{12}, \alpha')$. The fact that the remaining PSD conditions (5f)–(5h) are satisfied at $(x, X, y, Y_{12}, \alpha')$ follows from the facts that these conditions are satisfied at $(\bar{x}, \bar{X}, y, Y_{12}, \hat{\alpha})$, $\bar{x}_1 - \bar{\alpha}_1 = x_1 - \alpha_1$, the definition of $\bar{X}_{11}$ and the Schur complement condition. □

6 Another interpretation

The representation for $\mathscr{H}$ in Theorem 4 was obtained by starting with the representation in Theorem 3 and then arguing that only the single semidefiniteness constraint (5h) was necessary. In this section we describe an alternative derivation for the representation in Theorem 4. This derivation provides another interpretation for the conditions of Theorem 4 and also leads to a simple conjecture for a representation of $\mathscr{H}'$ as defined in Section 1.

The alternative derivation is based on replacing the variables y with $t = e - y$ and introducing the slack variables s such that $x+s+t = e$, as was done for the case $n = 1$ in the proof of Theorem 1. Note that each y_i is binary if and only if t_i is binary, and $(y, Y_{12}) \in \text{RLT}_y$ if and only if $(t, T_{12}) \in \text{RLT}_y$ where $T_{12} = 1 + Y_{12} - y_1 - y_2$. In fact the linear constraints (5a)–(5d) can be obtained by considering the equations $x_i + s_i + t_i = 1$, $i = 1, 2$, generating RLT constraints by multiplying each equation in turn by the variables (x_j, s_j, t_j), $i = 1, 2$, and then projecting onto the variables $(x, X, t, T_{12}, \alpha)$, where $\alpha_1 \approx x_1 t_2 = x_1(1-y_2)$, $\alpha_2 \approx x_2 t_1 = x_2(1-y_1)$, $T_{12} = 1 + Y_{12} - y_1 - y_2 \approx t_1 t_2$. Let

$$S = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix},$$

and note that S is nonsingular since it is a row permutation of a lower triangular matrix with every diagonal entry equal to one. Substituting variables and applying a symmetric transformation that preserves semidefiniteness, the PSD condition (5h′) can be written

$$S \begin{pmatrix} t_1 - T_{12} & 0 & 0 & \alpha_1 & 0 \\ 0 & t_2 - T_{12} & 0 & 0 & \alpha_2 \\ 0 & 0 & 1 + T_{12} - t_1 - t_2 & x_1 - \alpha_1 & x_2 - \alpha_2 \\ \alpha_1 & 0 & x_1 & X_{11} & X_{12} \\ 0 & \alpha_2 & x_2 & X_{12} & X_{22} \end{pmatrix} S^T \succeq 0,$$

which is

$$\begin{pmatrix} 1 - T_{12} & x_1 & x_2 & t_1 - T_{12} & t_2 - T_{12} \\ x_1 & X_{11} & X_{12} & 0 & \alpha_1 \\ x_2 & X_{12} & X_{22} & \alpha_2 & 0 \\ t_1 - T_{12} & 0 & \alpha_2 & t_1 - T_{12} & 0 \\ t_2 - T_{12} & \alpha_1 & 0 & 0 & t_2 - T_{12} \end{pmatrix} \succeq 0. \tag{11}$$

The PSD constraint (11) has a simple interpretation as a strengthening of the natural PSD condition

$$V = \begin{pmatrix} 1 & x_1 & x_2 & t_1 & t_2 \\ x_1 & X_{11} & X_{12} & 0 & \alpha_1 \\ x_2 & X_{12} & X_{22} & \alpha_2 & 0 \\ t_1 & 0 & \alpha_2 & t_1 & T_{12} \\ t_2 & \alpha_1 & 0 & T_{12} & t_2 \end{pmatrix} \succeq 0, \tag{12}$$

where V is a relaxation of the rank-one matrix vv^T with $v = (1, x_1, x_2, t_1, t_2)^T$. The matrix in (11) is obtained from V in (12) by subtracting $T_{12}uu^T$, where $u = (1, 0, 0, 1, 1)^T$. This can be interpreted as removing the portion of V corresponding to $t = e$, or equivalently $y = 0$, if V is decomposed into a convex combination of four matrices corresponding to $t \in \{0, e_1, e_2, e\}$, similar to the decomposition of $\mathscr{H}$ into a convex combination of $\mathscr{H}_y$, $y \in \{0, e_1, e_2, e\}$ in Section 3. Note in particular that $T_{12} = \lambda_0$, as defined in (3). Alternatively, as suggested by a referee, the condition (11) can be viewed as a relaxation of the condition that $vv^T \succeq 0$ for $v = (1 - T_{12}, x_1, x_2, t_1 - T_{12}, t_2 - T_{12})^T$.

We know that to obtain a representation of $\mathscr{H}$ the condition (11) cannot be replaced by (12); there are solutions $(x, X, y, Y_{12}, \alpha)$ that are feasible with the weaker PSD condition but where $(x, X, y, Y_{12}) \notin \mathscr{H}$. However it appears that the condition (12) is sufficient to obtain a representation of $\mathscr{H}'$. The following conjecture regarding $\mathscr{H}'$ is supported by extensive numerical computations, but remains unproved.

Conjecture 1 For $n = 2$, let $\mathscr{H}^+$ be the set of $(x, X, y, Y_{12}, \alpha)$ satisfying the constraints (5a)–(5d) and (12), where $t_1 = 1 - y_1$, $t_2 = 1 - y_2$ and $T_{12} = 1 + Y_{12} - y_1 - y_2$. Then $\mathscr{H}'$ equals the projection of $\mathscr{H}^+$ onto (x, X, y).

Note that (5a)–(5d) and (12) amount to the relaxation of (x, xx^T, y), which enforces PSD and RLT in the (x, X, y, Y_{12}) space and also exploits the binary nature of y. In other words, the standard approach for creating a strong SDP relaxation would be sufficient to capture the convex hull of (x, X, y) in this case, similar to the case of $n = 1$ as shown in the proof of Theorem 1, as well as the characterization of QPB for $n = 2$ from [2]. Finally, the RLT constraints and PSD condition (12) extend in a straightforward way to $n > 2$. For arbitrary n the constraint

$$\begin{pmatrix} 1 & x^T & t^T \\ x & X & \aleph \\ t & \aleph^T & T \end{pmatrix} \succeq 0 \tag{13}$$

where $\mathrm{diag}(\aleph) = 0$ and $\mathrm{diag}(T) = t$ relaxes the condition $vv^T \succeq 0$ for $v = (1, x^T, t^T)^T$ where $x \in \mathbb{R}^n_+$ and $t \in \{0, 1\}^n$. If Conjecture 1 is true, the one PSD condition (13), together with linear constraints constraints is sufficient to exactly represent the projection of $\mathscr{H}'$ onto any $(x_i, x_j, X_{ii}, X_{jj}, X_{ij}, y_i, y_j)$ corresponding to a subset of two variables $0 \le x_i \le y_i = 1 - t_i$ and $0 \le x_j \le y_j = 1 - t_j$.

Acknowlegement

We are grateful to two anonymous referees for their very careful readings of the paper and numerous insightful comments. We also thank Rekha Thomas and Pablo Parrilo for pointing out the applicability of [17, Lemma 3.14]

References

1. Aktürk, M.S., Atamtürk, A., Gürel, S.: A strong conic quadratic reformulation for machine-job assignment with controllable processing times. Oper. Res. Lett. **37**, 187–191 (2009)

2. Anstreicher, K.M., Burer, S.: Computable representations for convex hulls of low-dimensional quadratic forms. Math. Program. B **124**, 33–43 (2010)
3. Atamtürk, A., Gómez, A.: Strong formulations for quadratic optimization with M-matrices and indicator variables. Math. Program. B **170**, 141–176 (2018). DOI 10.1007/s10107-018-1301-5
4. Burer, S.: On the copositive representation of binary and continuous nonconvex quadratic programs. Math. Program. **120**, 479–495 (2009). DOI http://dx.doi.org/10.1007/s10107-008-0223-z
5. Burer, S., Letchford, A.N.: On non-convex quadratic programming with box constraints. SIAM J. Optim. **20**, 1073–1089 (2009)
6. Ceria, S., Soares, J.: Convex programming for disjunctive convex optimization. Math. Program. **86**, 595–614 (1999). DOI 10.1007/s101070050106
7. D'Ambrosio, C., Linderoth, J., Luedtke, J.: Valid inequalities for the pooling problem with binary variables. In: Integer Programming and Combinatorial Optimization, *Lecture Notes in Comput. Sci.*, vol. 6655, pp. 117–129. Springer, Heidelberg (2011)
8. Dong, H., Linderoth, J.: On valid inequalities for quadratic programming with continuous variables and binary indicators. In: M. Goemans, J. Correa (eds.) IPCO: International Conference on Integer Programming and Combinatorial Optimization, *Lecture Notes in Computer Science*, vol. 7801, pp. 169–180. Springer Verlag, Berlin (2013)
9. Frangioni, A., Gentile, C.: Perspective cuts for a class of convex 0-1 mixed integer programs. Math. Program. **106**, 225–236 (2006)
10. Günlük, O., Linderoth, J.: Perspective reformulations of mixed integer nonlinear programming with indicator variables. Math. Program. B **124**, 183–205 (2010)
11. Hijazi, H., Coffrin, C., Van Hentenryck, P.: Convex quadratic relaxations for mixed-integer nonlinear programs in power systems. Math. Program. Comput. **9**, 321–367 (2017)
12. Horn, R., Johnson, C.: Matrix Analysis. Cambridge University Press (1985)
13. Jabr, R.: Tight polyhedral approximation for mixed-integer linear programming unit commitment formulations. IET Gener. Transm. Distrib. **6**, 1104–1111(7) (2012)
14. Jach, M., Michaels, D., Weismantel, R.: The convex envelope of $(n-1)$-convex functions. SIAM. J. Optim. **19**, 1451–1466 (2008)
15. Linderoth, J.: A simplicial branch-and-bound algorithm for solving quadratically constrained quadratic programs. Math. Program. **103**, 251–282 (2005)
16. Maxfield, J.E., Minc, H.: On the matrix equation $X'X = A$. Proc. Edinburgh Math. Soc. (2) **13**, 125–129 (1962/1963)
17. Parrilo, P.A.: Polynomial optimization, sums of squares, and applications. In: G. Blekherman, P.A. Parrilo, R.R. Thomas (eds.) Semidefinite Optimization and Convex Algebraic Geometry, *MOS-SIAM Series on Optimization*, vol. 13, pp. 47–158 (2012)
18. Sherali, H.D., Adams, W.P.: A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems. Kluwer (1997)
19. Zheng, X., Sun, X., Li, D.: Improving the performance of MIQP solvers for quadratic programs with cardinality and minimum threshold constraints: A semidefinite program approach. INFORMS J. Comp. **26**, 690–703 (2014)