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

A combinatorial approach to Ramana's exact dual for semidefinite programming

Gábor Pataki

Received: date / Accepted: date

Abstract Thirty years ago, in a seminal paper Ramana derived an exact dual for Semidefinite Programming (SDP). Ramana's dual has the following remarkable features: i) it assumes feasibility of the primal, but it does not make any regularity assumptions, such as strict feasibility ii) its optimal value is the same as the optimal value of the primal, so there is no duality gap. iii) it attains its optimal value when it is finite iv) it yields a number of complexity results in SDP, which are fundamental, and to date are still the best known. For example, it proves that SDP feasibility in the Turing model is not NP-complete, unless NP = co-NP.

In this work we extend and simplify previous analyses of Ramana's dual. First, we completely characterize the feasible set of Ramana's dual for inequality constrained SDPs. Second, we similarly analyze Ramana's dual for equality constrained SDPs. We do this by connecting it to a seemingly very different way of inducing strong duality: reformulating the SDP using elementary row operations inherited from Gaussian eliminination. Our characterizations yield a short and transparent derivation of Ramana's dual.

Our approach is combinatorial in the following sense: i) we use a minimum amount of continuous optimization theory ii) we show that feasible solutions in Ramana's dual are identified with regular facial reduction sequences, i.e., essentially discrete structures.

Keywords semidefinite programming · duality · Ramana's dual · facial reduction

Mathematics Subject Classification (2020) 90C22 · 90C46 · 49N15 · 52A40

Department of Statistics and Operations Research, University of North Carolina at Chapel Hill
E-mail: gabor@unc.edu

1 Introduction

1.1 Semidefinite programs and shortcomings of the usual dual

Semidefinite Programs (SDPs) – optimization problems with linear objective, linear constraints, and semidefiniteness constraints on matrix variables – are some of the most versatile and popular optimization problems to emerge in the last thirty years. SDPs appear in combinatorial optimization, polynomial optimization, engineering, and other application areas, and can be solved by efficient optimization algorithms. See, for example, [15,23] for the foundational theory of interior point methods, [25,24] for efficient implementations of such methods, and [27,2,6] for efficient algorithms based on different principles.

We formulate an SDP mathematically as

$$\begin{array}{ll} \sup & c^\top x \\ s.t. & \sum_{i=1}^m x_i A_i \preceq B, \end{array} \tag{P}$$

where the A_i and B are $n \times n$ symmetric matrices, $c \in \mathbb{R}^m$, and for symmetric matrices T and S we write $S \preceq T$ to say that $T - S$ is positive semidefinite (psd).

The problem (P), which we call the *primal*, has a natural dual problem

$$\begin{array}{lll} \inf & \langle B, Y \rangle & \\ s.t. & \langle A_i, Y \rangle & = c_i \ (i = 1, \dots, m) \\ & Y \succeq 0, & \end{array} \tag{D}$$

where for symmetric matrices S and T we write $\langle T, S \rangle := \operatorname{trace}(TS)$ to denote their inner product. The main role of (D) is to certify boundedness of the optimal value of (P) and to certify optimality of a feasible solution. For example, when x is feasible in (P), and Y in (D), then the weak duality inequality $c^\top x \leq \langle B, Y \rangle$ always holds. Thus, if we find a pair x and Y whose objective values are equal, then we know they must be both optimal.

Besides weak duality, a desirable property of (P) and of (D) is *strong duality,* which is said to hold when the optimal values of (P) and (D) agree, and the latter is attained, when it is finite. However, strong duality between (P) and (D) sometimes fails, as the following example shows:

Example 1 Consider the SDP

$$\begin{array}{l} \sup \ x_1 + x_2 + x_3 \\ s.t. \ x_1 \underbrace{\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{A_1} + x_2 \underbrace{\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}}_{A_2} + x_3 \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}}_{A_3} \preceq \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{B}. \end{array} \tag{1}$$

We claim that its optimal value is zero. Indeed, assume that x is a feasible solution, and let $S := B - \sum_i x_i A_i$. Since the upper left corner of S is zero, we have $x_1 = x_2 = 0$. Thus, the $(2,2)$ element of S is zero, hence $x_3 = 0$.

The dual is

$$\begin{aligned}
&\inf\ y_{33} \\
&s.t.\ y_{12} = 1/2 \\
&\qquad y_{22} + 2y_{13} = 1 \\
&\qquad y_{23} = 1/2 \\
&\qquad Y \succeq 0.
\end{aligned} \tag{2}$$

We claim that (2) has no solution with value zero. Indeed, to get a contradiction, suppose Y is such a solution. Then $y_{33} = 0,$ and $Y \succeq 0$ implies $y_{23} = 0,$ contradicting the constraint $y_{23} = 1.$ (With a bit more work one can show that the optimal value of (2) is 0, but it is not attained.)

Strong duality can be ensured if we assume certain regularity conditions. The best known such condition is strict feasibility: when (P) is strictly feasible, i.e., $B - \sum_i x_i A_i$ is positive definite for some $x,$ then strong duality holds between (P) and (D). An analogous result holds when (D) is strictly feasible, i.e., when it has a positive definite feasible $Y.$

However, assuming strict feasibility is not satisfactory from a theoretical perspective. Most importantly, it is of no help in finding an exact alternative system of (P), i.e., a semidefinite system which is feasible, exactly when (P) is infeasible. Indeed, the usual "Farkas type" system

$$\begin{aligned}
\langle A_i, Y\rangle &= 0\, (i = 1, \ldots, m) \\
\langle B, Y\rangle &= -1 \\
Y &\succeq 0
\end{aligned} \tag{alt-P}$$

of (P) is not an exact alternative system [1]. For a concise treatment of duality in conic linear programs, which include SDPs, see, e.g. Renegar [23, Chapter 3].

1.2 **Ramana's dual**

Thirty years ago, in a seminal paper Ramana [20] [2] constructed an elegant dual problem, which avoids the shortcomings of the traditional dual. Ramana's dual has the following striking properties: i) it assumes that (P) is feasible, but does not assume that it is strictly feasible ii) strong duality holds between (P) and Ramana's dual. Put simply, it has all desirable properties of (D) when (P) is strictly feasible, without actually assuming that (P) is strictly feasible!

Ramana's dual yields an exact alternative system of (P), and fundamental results in complexity theory. The most important of these are:

(1) In the real number model of computing, deciding feasibility of SDP is in NP ∩ co-NP.

[1] More precisely, (alt-P) is an exact alternative system of (P), if there is a positive definite Y such that $\langle A_i, Y\rangle = 0$ for all $i.$ However, this assumption is quite restrictive. We can of course also assume that all the A_i and B are diagonal, so (P) is just a linear program, but this assumption is even more restrictive.

[2] The first version of [20] was circulated in 1995.

(2) In the Turing model of computing, deciding feasibility of SDPs is not NP-complete, unless NP = co-NP.

These results are still the best known on SDP feasibility.

To state Ramana's dual, we assume that the primal (P) is feasible, and we denote by val() the optimal value of an optimization problem. We denote by $\mathcal{S}^n$ the set of $n \times n$ symmetric matrices, and by $\mathcal{S}^n_+$ the set of symmetric psd matrices. We also introduce the linear operator $\mathcal{A}$ and its adjoint $\mathcal{A}^*$ as

$$\mathcal{A}x := \sum_{i=1}^{m} x_i A_i,\ \mathcal{A}^* Y = (\langle A_1, Y, \ldots, \langle A_m, Y \rangle)^\top \text{ for } x \in \mathbb{R}^m,\ Y \in \mathcal{S}^n.$$

Theorem 1 *Consider the optimization problem called the* Ramana dual of (P):

$$\begin{array}{rrcll} \inf & \langle B, U_{n+1} + V_{n+1} \rangle & & & \\ s.t. & \mathcal{A}^*(U_{n+1} + V_{n+1}) & = & c & \\ & \mathcal{A}^*(U_i + V_i) & = & 0 & i = 1, \ldots, n \\ & \langle B, U_i + V_i \rangle & = & 0 & i = 1, \ldots, n \\ & U_i & \in & \mathcal{S}^n_+ & i = 1, \ldots, n+1 \\ & V_i & \in & \tan(U_{i-1}) & i = 1, \ldots, n+1 \\ & U_0 = V_0 & = & 0. & \end{array} \qquad (\mathrm{D_{Ram}})$$

Here for $U \in \mathcal{S}^n_+$ the set $\tan(U)$ *is defined as*

$$\tan(U) \ = \ \{ W + W^\top \,|\, U \succeq \beta\, W W^\top \textit{ for some}\, \beta > 0 \}. \qquad (3)$$

We have

$$\mathrm{val}\,(\mathrm{P}) = \mathrm{val}\,(\mathrm{D_{Ram}}),$$

and $\mathrm{val}\,(\mathrm{D_{Ram}})$ *is attained when finite.* □

To make ($\mathrm{D_{Ram}}$) into a proper SDP, we claim that for a psd matrix U we can represent $\tan(U)$ as

$$\tan(U) = \left\{ W + W^\top \ : \ \begin{pmatrix} U & W \\ W^\top & \lambda I \end{pmatrix} \succeq 0 \text{ for some } \lambda \geq 0 \right\}. \qquad (4)$$

To see why, we use the Schur complement condition for positive semidefiniteness. If (U, W, β) satisfy the condition in (3), then $(U, W, 1/\beta)$ satisfy the condition in (4). Conversely, if (U, W, λ) satisfy the conditions in (4), then we can assume $\lambda > 0$, and $(U, W, 1/\lambda)$ satisfy the condition in (3).

Thus, ($\mathrm{D_{Ram}}$) can be written as an SDP with polynomially many variables and constraints.

Example 2 (Example 1 continued) The Ramana dual of (1) does have a solution with value zero. We set $U_0 = V_0 = 0,$ and

$$U_1{+}V_1 = \begin{pmatrix} \textcircled{1} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, U_2{+}V_2 = \begin{pmatrix} \textcircled{1} & 0 & -1/2 \\ 0 & \textcircled{1} & 0 \\ -1/2 & 0 & 0 \end{pmatrix}, U_3{+}V_3 = \begin{pmatrix} 0 & 1/2 & -1/2 \\ 1/2 & 2 & 1/2 \\ -1/2 & 1/2 & \textcircled{0} \end{pmatrix}. \tag{5}$$

In (23) we show the U_i component of $U_i + V_i$ with circled entries. For example, U_2 is the matrix in which the $(1,1)$ and $(2,2)$ elements are 1 and all other elements are zero.

We see that $V_i \in \tan(U_{i-1}),$ with the decomposition $V_i = W_i + W_i^\top$ for $i = 2, 3,$ where

$$W_2 = \begin{pmatrix} 0 & 0 & -1/2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, W_3 = \begin{pmatrix} 0 & 1/2 & -1/2 \\ 0 & 1 & 1/2 \\ 0 & 0 & 0 \end{pmatrix}.$$

1.3 **Literature**

Ramana's dual is fundamental, however, the original proof of its correctness is somewhat lengthy and technical. Thus several papers gave shorter proofs, and explored connections to other work. Ramana, Tunçel and Wolkowicz [22] and [17, 18, 11] connected Ramana's dual to the facial reduction algorithm of Borwein and Wolkowicz [1]. Klep and Schweighofer [8] designed a dual with similar properties, based on algebraic geometry. Luo, Sturm, and Zhang [13] gave a different proof of the correctness of Ramana's dual; and Ramana and Freund [21] studied its dual.

Ramana's dual was used by DeKlerk et al [4] in self-dual embeddings. Due to its complexity implications it is often mentioned in the discrete mathematics and theoretical computer science literature, see for example, Lovász [12] and O' Donnell [16]. Ramana's dual is often cited in surveys and books: see for example, DeKlerk [3], Drusvyatskiy and Wolkowicz [5], Vandenberghe and Boyd [26], Nemirovski [14], and Laurent and Rendl [9].

Despite the importance of Ramana's dual and the many followup papers, one can make the case that we still need to understand it better. On the one hand, the cited references characterize its optimal value. However, it would also be very useful to characterize its feasible set, both from the theoretical, and possibly a practical perspective. Second, a simple correctness proof, accessible to the theoretical computer science community, is also desirable.

1.4 **Contributions**

First we completely characterize the feasible set of ($\mathrm{D_{Ram}}$). Second, we similarly analyze a problem ($\mathrm{P_{Ram}}$), which is the Ramana dual of (D). Third, we connect ($\mathrm{P_{Ram}}$) to a seemingly very different way of inducing strong duality from [10]: reformulating (D) using elementary row operations inherited from Gaussian elimination. Fourth, using our characterizations we give short and transparent derivations of Theorem 1 and its primal counterpart Theorem 3.

As we alluded to in the title, our approach is combinatorial. While a "combinatorial approach" is not perfectly defined, the main features of our proof are:

(1) We avoid the use of more technical concepts in convex analysis, such as relative interiors, and conjugate faces, which play an important role in the analysis of [22,17,18]. In fact, we only use a single ingredient from continuous optimization theory: when (P) is strictly feasible, then strong duality holds between (P) and (D).
(2) We show that feasible solutions in (D_{Ram}) are identified with regular facial reduction sequences, i.e., essentially discreet structures.

In our analysis of (D_{Ram}) we use ideas from [11], whose main contribution is to derive (D_{Ram}) using less machinery, than previous papers. However, our treatment is more extensive, as it characterizes the feasible *set* of (D_{Ram}), and it is considerably shorter. (Also, the material on (P_{Ram}) is not covered in [11]).

1.5 **Organization of the paper and guide to the reader**

In Subsection 1.6 we fix notation, prove two simple propositions, and define one of the main players of the paper, regular facial reduction sequences. In Section 2 we analyse (D_{Ram}):

– In Subsection 2.1 we define the maximum rank slack in (P), which measures "how far" (P) is from being strictly feasible. We then show how to certify the maximum rank slack using regular facial reducition sequences.
– In Subsection 2.2 we study the strong dual (D_{strong}) which has all the properties required from Ramana's dual. However, (D_{strong}) relies on knowing the maximum rank slack in (P), which is generally an unknown quantity.
– In Subsection 2.3 we present Theorem 2, which precisely characterizes the feasible solutions of (D_{Ram}).
– In Subsection 2.4 we prove Theorem 1.

Section 3 treats (P_{Ram}). In particular,

– In Subsection 3.1 we introduce (P_{Ram}) and state Theorem 3, which proves strong duality between (D) and (P_{Ram}).
– In Subsection 3.2 we recall reformulations of (D), and show how they certify the maximum rank feasible solution in (D).
– In Subsection 3.3 we define the strong primal of (D).
– In Subsection 3.4 we prove Theorem 4, which precisely characterizes feasible solutions of (P_{Ram}).
– In Subsection 3.4 we give a quick proof of Theorem 3.

Some readers of this paper may only want to see a quick and transparent derivation of (D_{Ram}). For them, reading only Subsection 1.6, and Section 2 will probably suffice.

1.6 Preliminaries

We denote by $\mathcal{S}^{n,k}$ the set of $n \times n$ symmetric matrices in which all nonzeroes are in the first k rows and columns. We let $\mathcal{S}_+^{n,k} = \mathcal{S}_+^n \cap \mathcal{S}^{n,k}$ i.e, the set of psd matrices in which only the upper left $k \times k$ block is nonzero. We denote by $\mathcal{S}_{++}^{n,k}$ the matrices in $\mathcal{S}^{n,k}$ in which the upper left $k \times k$ block is positive definite.

Next we state two basic propositions. The proof of Proposition 1 is straightforward from the properties of the trace, and the proof of Proposition 2 from the definition of $\tan(U)$.

Proposition 1 *Suppose Q is an $n \times n$ orthonormal matrix. Then*

$$\langle X, Y \rangle = \langle Q^\top X Q, Q^\top Y Q \rangle \tag{6}$$

for all $X, Y \in \mathcal{S}^n$. Further,

$$V \in \tan(U) \;\Leftrightarrow\; Q^\top V Q \in \tan(Q^\top U Q) \tag{7}$$

for all $U \in \mathcal{S}_+^n$, and $V \in \mathcal{S}^n$. □

Proposition 2 *The following hold:*

(1) If $U \in \mathcal{S}_+^{n,k}$ and $V \in \tan(U)$, then $V \in \mathcal{S}^{n,k}$.
(2) If $U \in \mathcal{S}_{++}^{n,k}$ and $V \in \mathcal{S}^{n,k}$, then $V \in \tan(U)$.

□

We visualize Proposition 2 in (8), in which $\oplus$ stands for a psd submatrix, and the $\times$ stands for a block with arbitrary elements. If U is as given on the left, and $V \in \tan(U)$, then V must be of the form given on the right. Further, if the $\oplus$ block in U is positive definite, then any V in the form on the right is in $\tan(U)$:

$$U = \begin{pmatrix} \overbrace{\oplus}^{k} & \overbrace{0}^{n-k} \\ 0 & 0 \end{pmatrix}, V = \begin{pmatrix} \overbrace{\times}^{k} & \overbrace{\times}^{n-k} \\ \times & 0 \end{pmatrix}. \tag{8}$$

We next introduce a main player of the paper:

Definition 1 *We say that $(Y_1, \ldots, Y_k)$ is a* regular facial reduction sequence for $\mathcal{S}_+^n$ [3] *if the Y_i are in $\mathcal{S}^n$ and are of the form*

$$Y_1 = \begin{pmatrix} \overbrace{\Lambda_1}^{r_1} & \overbrace{0}^{n-r_1} \\ 0 & 0 \end{pmatrix}, \ldots, Y_i = \begin{pmatrix} \overbrace{\times}^{r_1+\ldots+r_{i-1}} & \overbrace{\times}^{r_i} & \overbrace{\times}^{n-r_1-\ldots-r_i} \\ \times & \Lambda_i & 0 \\ \times & 0 & 0 \end{pmatrix}$$

for $i = 1, \ldots, k$. Here the r_i are nonnegative integers, the Λ_i diagonal positive definite matrices, and the $\times$ symbols correspond to blocks with arbitrary elements.

[3] Slightly different versions of regular facial reduction sequences have been defined in other papers, e.g. in [19].

Note that regular facial reduction sequences are essentially discrete structures. When we use them, we only use that the Λ_i are positive definite, and what their sizes are; however, we never refer to their actual entries.

2 Analysis of (D_{Ram})

As a convention, whenever we consider solutions of (D_{Ram}), we assume $U_0 = V_0 = 0$ without spelling this out.

2.1 Certificates for the maximum rank slack in (P)

Definition 2 *We say that a positive semidefinite matrix is a* slack *in* (P)*, if it is of the form $B - \mathcal{A}x$ for some $x \in \mathbb{R}^m$.*

Since the rank of a matrix is a nonnegative integer, (P) has a slack of maximum rank.

Definition 3 *We say that we* rotate *a set of matrices say $M_1, \dots, M_k$ by an orthonormal matrix Q, if we replace M_i by $Q^\top M_i Q$ for all i.*

Proposition 3 *Suppose we rotate all A_i and B by an orthonormal matrix Q. This operation does not change the optimal value and attainment in* (P)*,* (D)*, and* (D_{Ram})*.*

Proof The statement is obvious for (P), since $x \in \mathbb{R}^m$ is feasible in (P) before the rotation if and only if it is feasible afterwards. To prove it for (D), if Y is feasible before the rotation, then $Q^\top Y Q$ is feasible afterwards by part (6) of Proposition 1.

To deal with (D_{Ram}), let $A_i' = Q^\top A_i Q$ for all i and $B' = Q^\top B Q$. Suppose $(U_j, V_j)_{j=0}^{n+1}$ is feasible in (D_{Ram}) before the rotation, and let $U_j' = Q^\top U_j Q$ and $V_j' = Q^\top V_j Q$ for all j.

Then

$$\begin{aligned} \langle A_i, U_j + V_j \rangle &= \langle A_i', U_j' + V_j' \rangle \text{ for all } i, j \\ \langle B, U_{n+1} + V_{n+1} \rangle &= \langle B', U_{n+1}' + V_{n+1}' \rangle \\ V_i' &\in \tan(U_{i-1}') \text{ for all } i, \end{aligned}$$

where the first two equations follow from (6) and the last from (7). Thus, feasible solutions of the original and rotated problems are in one-to-one correspondence with the same objective value. □

Lemma 1 *Suppose* (P) *is not strictly feasible. Then there is $Y \in \mathcal{S}_+^n \setminus \{0\}$ such that*

$$\begin{aligned} \mathcal{A}^* Y &= 0 \\ \langle B, Y \rangle &= 0. \end{aligned} \tag{9}$$

Proof Suppose (P) is not strictly feasible. We claim that the optimal value of the SDP

$$\begin{array}{ll} \sup z & \\ s.t. \ \mathcal{A}x + zI \preceq B \end{array} \tag{10}$$

is zero. Indeed it is ≥ 0 since (P) is feasible. It is ≤ 0, since if it had a feasible solution $(\bar{x}, \bar{z})$ with $\bar{z} > 0$, then $B - \mathcal{A}\bar{x}$ would be a strictly feasible solution in (P). Thus the dual of (10) has a feasible solution $Y \succeq 0$ which satisfies the requirements of the lemma: Y is nonzero, due to the dual constraint $\langle I, Y \rangle = 1$. □

The following lemma shows that regular facial reduction sequences can be used to certify that the maximum rank slack in (P) indeed has maximum rank:

Lemma 2 *Suppose $r \in \{0, \dots, n\}$. Then the following statements are equivalent:*

(1) The rank of the maximum rank slack in (P) is r.
(2) There is an orthonormal matrix Q such that after rotating all A_i and B with Q there is
   (a) a slack in (P) of the form

$$Z = \begin{pmatrix} 0 & 0 \\ 0 & \Lambda \end{pmatrix}, \tag{11}$$

   where Λ is diagonal positive definite of order r.
   (b) a regular facial reduction sequence $Y_1, \dots, Y_k$ such that $k \leq n$ and

$$\mathcal{A}^* Y_i = 0 \ \ for\ i = 1, \dots, k \tag{12}$$

$$\langle B, Y_i \rangle = 0 \ \ for\ i = 1, \dots, k \tag{13}$$

$$r_1 + \dots + r_k = n - r, \tag{14}$$

   where the r_i are the sizes of the positive definite blocks in the Y_i.

Proof For both directions we will use that rotating the A_i and B does not change the rank of the maximum rank slack in (P).

(1) $\Leftarrow$ (2) : By (2a) the rank of the maximum rank slack in (P) is at least r. To show it is at most r, suppose S is any slack in (P). Then $\langle Y_1, S \rangle = 0$, so a positively weighted sum of the first r_1 diagonal elements of S is zero. Since these elements are nonnegative, they must all be zero. Since $S \succeq 0$, the first r_1 rows and columns are of S zero. Continuing, we deduce that the first $r_1 + \dots + r_k$ rows and columns of S are zero [4].

(1) $\Rightarrow$ (2) : If $r = n$, then we set $k = 0$ and $Y_1, \dots, Y_k$ the empty sequence. If $r < n$, then we construct $Y_1 \in \mathcal{S}^n_+ \setminus \{0\}$ to satisfy (9) in Lemma 1. After

[4] This argument explains the parlance "facial reduction sequence," since the set of such psd matrices is a face of $\mathcal{S}^n_+$. A convex subset F of a convex set $\mathcal{S}^n_+$ is a face, if for any $X, Y \in \mathcal{S}^n_+$ if the open line segment $\{\lambda X + (1-\lambda)Y : 0 < \lambda < 1\}$ intersects F, then both X and Y must be in F.

rotating Y_1 by a suitable orthonormal Q and all A_i and B by the same Q we have

$$Y_1 = \begin{pmatrix} \Lambda_1 & 0 \\ 0 & 0 \end{pmatrix},$$

where Λ_1 is order r_1 diagonal positve definite for some $r_1 > 0$. Also, by part (6) of Proposition 1 (with A_i, B, and I in place of X, and Y_1 in place of Y) the first two equations in (9) still hold.

Now if S is any slack in (P) (after the rotation), then $\langle Y_1, S\rangle = 0$. Hence, as we explained above, the first r_1 rows and columns of S are zero. We next construct an SDP

$$\sum_{i=2}^{m} x_i F_i + G \preceq 0, \tag{15}$$

where F_i is obtained from A_i by deleting the first r_1 rows and columns for $i = 2, \dots, m$ and G is obtained from B in the same manner. The maximum rank of a slack in (15) is still r, so we proceed in a similar manner with this smaller SDP. Once our process stops, we have the required $Y_1, \dots, Y_k$.

Finally, we fix a maximum rank slack Z and rotate the A_i and B by a matrix that affects only the lower right order r block of Z to put Z into the form (11).

□

Note that the constructions in Lemma 1 and 2 are theoretical. While the proofs are constructive, to actually compute the Y_i in Lemma 2 we would need to solve (10) (and its dual) in exact arithmetic.

Example 3 (Example 1 and 2 continued) In the SDP (1) all variables are zero in a feasible solution, so the right hand side is the maximum rank slack.

This SDP does not need any rotation. A regular facial reduction sequence Y_1, Y_2 that satisfies the conclusions of Lemma 2 is obtained by setting

$$Y_i := U_i + V_i$$

for $i = 1, 2$ in the formula (23).

From now on we assume

(1) We rotated all A_i and B so the Y_i in Lemma 2 exist. (2) There is a maximum rank slack Z in (P) of rank $r \in \{0, \dots, n\}$, of the form given in (11).

2.2 The strong dual

Lemma 3 can be proved by standard convex optimization techniques. To make the paper self contained, in Section A we give a proof that only uses the result "strict feasibility of (P) implies strong duality between (P) and (D)" and avoids the concept of relative interior. Our strong dual is essentially equivalent to the minimal cone based dual of [1].

Lemma 3 *Consider the optimization problem*

$$\begin{aligned} &\inf\ \langle B, Y\rangle \\ &s.t.\ \mathcal{A}^*Y = c \\ &\quad\ Y_{22} \in \mathcal{S}^r_+. \end{aligned} \tag{$\mathrm{D_{strong}}$}$$

Then

$$\mathrm{val}\,(\mathrm{P}) = \mathrm{val}\,(\mathrm{D_{strong}}),$$

and the optimal value of ($\mathrm{D_{strong}}$) *is attained when it is finite. Here* Y_{22} *stands for the lower right* $r \times r$ *block of* Y. □

2.3 Connecting regular facial reduction sequences to Ramana's dual

We now present the main result of the paper, Theorem 2, which connects regular facial reduction sequences and feasible solutions of ($\mathrm{D_{strong}}$) to feasible solutions of ($\mathrm{D_{Ram}}$). It shows that regular facial reduction sequences correspond to $U_1, V_1, \ldots, U_n, V_n$ in ($\mathrm{D_{Ram}}$) and feasible solutions of ($\mathrm{D_{strong}}$) correspond to U_{n+1} and V_{n+1} in ($\mathrm{D_{Ram}}$).

Before stating it, we note that in Lemma 2 we can assume $k = n$. Indeed, if $k < n$, then we can just add $Y_1 = \cdots = Y_{n-k} = 0$ at the start of the sequence, and renumber the remaining Y_i.

We also need a definition:

Definition 4 *We say that a* Q *orthonormal matrix is* admissible, *if*

$$Q = \begin{pmatrix} Q' & 0 \\ 0 & I_r \end{pmatrix}.$$

for some Q' *orthonormal matrix.*

Clearly, if we rotate the A_i and B by an admissible Q, then the maximum rank slack remains as in the form required in (11); hence the optimal value and attainment in ($\mathrm{D_{strong}}$) is unaffected.

Theorem 2 *The following hold:*

(1) Suppose $Y_1, \ldots, Y_n$ *is a regular facial reduction sequence constructed by Lemma 2 (with* $k = n$*), and* Y_{n+1} *is feasible in* ($\mathrm{D_{strong}}$)*. Then there is a decomposition*

$$Y_i = U_i + V_i \text{ for } i = 1, \ldots, n+1, \tag{16}$$

such that $(U_i, V_i)_{i=0}^{n+1}$ *is feasible in* ($\mathrm{D_{Ram}}$)*.*

(2) Suppose $(U_i, V_i)_{i=0}^{n+1}$ *is feasible in* ($\mathrm{D_{Ram}}$)*. Then we can rotate all* A_i, B, U_i, V_i *by an admissible matrix so that after the rotation*

$$U_1 + V_1, \ldots, U_n + V_n \text{ is a regular facial reduction sequence} \tag{17}$$

and

$$U_{n+1} + V_{n+1} \quad \text{is feasible in} \quad (\mathrm{D_{strong}}). \tag{18}$$

Further, if after this rotation the sizes of the positive definite blocks in $U_i + V_i$ are r_i for $i = 1, \dots, n$, then

$$U_i \in \mathcal{S}_+^{r_1+\dots+r_i,n} \text{ for } i = 1, \dots, n. \tag{19}$$

Proof of (1): For $i = 1, \dots, n$ let Λ_i be the positive definite block in Y_i, suppose the order of Λ_i is r_i, and define

$$U_i := \begin{pmatrix} \overbrace{\begin{matrix} I \\ 0 \\ 0 \end{matrix}}^{r_1+\dots+r_{i-1}} & \overbrace{\begin{matrix} 0 \\ \Lambda_i \\ 0 \end{matrix}}^{r_i} & \overbrace{\begin{matrix} 0 \\ 0 \\ 0 \end{matrix}}^{n-\sum_{\ell=1}^i r_\ell} \end{pmatrix}, V_i := Y_i - U_i = \begin{pmatrix} \overbrace{\begin{matrix} \times \\ \times \\ \times \end{matrix}}^{r_1+\dots+r_{i-1}} & \overbrace{\begin{matrix} \times \\ 0 \\ 0 \end{matrix}}^{r_i} & \overbrace{\begin{matrix} \times \\ 0 \\ 0 \end{matrix}}^{n-\sum_{\ell=1}^i r_\ell} \end{pmatrix}. \tag{20}$$

Also, let $\bar{Y}$ be the lower right order r block of Y_{n+1} and define

$$U_{n+1} = \begin{pmatrix} \overbrace{\begin{matrix} 0 \\ 0 \end{matrix}}^{n-r} & \overbrace{\begin{matrix} 0 \\ \bar{Y} \end{matrix}}^{r} \end{pmatrix}, V_{n+1} := Y_{n+1} - U_{n+1} = \begin{pmatrix} \overbrace{\begin{matrix} \times \\ \times \end{matrix}}^{n-r} & \overbrace{\begin{matrix} \times \\ 0 \end{matrix}}^{r} \end{pmatrix}.$$

The above notation stresses which blocks of the U_i and V_i can be nonzero.

By part (2) of Proposition 2 we deduce $V_i \in \tan(U_{i-1})$ for $i = 2, \dots, n+1$. Also, the U_i and V_i satisfy the equality constraints of (D_{Ram}) by (12)–(13). So $(U_i, V_i)_{i=0}^{n+1}$ is feasible in (D_{Ram}), as wanted.

Proof of (2): Let $Y_i = U_i + V_i$ for all i.

Since $Y_1 = U_1 \in \mathcal{S}_+^n$ and $\langle Y_1, Z \rangle = 0$, the last r rows and columns of Y_1 are zero. So we rotate all A_i, U_i, V_i, Y_i and B by an admissible matrix to ensure

$$Y_1 = \begin{pmatrix} \Lambda_1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{21}$$

where Λ_1 is diagonal positive definite.

Suppose next that $1 \leq i \leq n$ and $Y_1, \dots, Y_i$ is a regular facial reduction sequence, in which the positive definite blocks have size $r_1, \dots, r_i$, respectively. Also assume $U_1 \in \mathcal{S}_+^{r_1,n}$. Both of these statements are true, when $i = 1$.

We claim

$$\begin{aligned} U_1 \in \mathcal{S}_+^{n,r_1} \Rightarrow V_2 \in \mathcal{S}^{n,r_1} &\Rightarrow U_2 \in \mathcal{S}_+^{n,r_1+r_2} &&\Rightarrow V_3 \in \mathcal{S}^{n,r_1+r_2} \\ \dots \quad &\Rightarrow U_i \in \mathcal{S}_+^{n,r_1+\dots+r_i} &&\Rightarrow V_{i+1} \in \mathcal{S}^{n,r_1+\dots+r_i}. \end{aligned} \tag{22}$$

Indeed, the first implication is by $V_2 \in \tan(U_1)$ and part (1) of Proposition 2. The second implication follows, since $Y_2 = U_2 + V_2$. The third is by $V_3 \in \tan(U_2)$ and again by part (1) of Proposition 2. The other implications follow similarly.

Since $U_{i+1} \in \mathcal{S}_+^n$ and $V_{i+1} \in \mathcal{S}^{n,r_1+\dots+r_i}$, the lower right order $n - \sum_{\ell=1}^i r_\ell$ block of Y_{i+1}, which we call $\bar{Y}$, is psd. We now distinguish two cases.

First suppose $i < n$. Since $\langle Y_{i+1}, Z \rangle = 0$, the lower right order r block of $\bar{Y}$ is zero, so the last r rows and columns of $\bar{Y}$ are zero. Thus

$$Y_{i+1} = \begin{pmatrix} \overbrace{\times}^{r_1+\cdots+r_i} & \overbrace{\times}^{n-\sum_{\ell=1}^{i} r_\ell - r} & \overbrace{\times}^{r} \\ \times & Y' & 0 \\ \times & 0 & 0 \end{pmatrix}, \; Q^\top Y_{i+1} Q = \begin{pmatrix} \overbrace{\times}^{r_1+\cdots+r_i} & \overbrace{\times}^{r_{i+1}} & \overbrace{\times}^{n-\sum_{\ell=1}^{i+1} r_\ell} \\ \times & \Lambda_{i+1} & 0 \\ \times & 0 & 0 \end{pmatrix},$$

where Y' is psd, Q is an orthonormal matrix of the form

$$Q = \begin{pmatrix} I_{r_1+\cdots+r_i} & 0 & 0 \\ 0 & Q' & 0 \\ 0 & 0 & I_r \end{pmatrix},$$

[] and Λ_{i+1} is diagonal positive definite. So we rotate all A_i, U_i, V_i, Y_i and B by Q, and afterwards $Y_1, \ldots, Y_{i+1}$ is a regular facial reduction sequence. Further, $U_1 \in \mathcal{S}_+^{r_1,n}$ still holds. After we are done with the rotations for $i = 1, \ldots, n-1$ we have that $Y_1, \ldots, Y_n$ is a regular facial reduction sequence and (19) holds.

Second, assume $i = n$. Since $V_{n+1} \in \mathcal{S}^{n, r_1+\cdots+r_n}$ and $r_1 + \cdots + r_n \le n - r$, we see that the lower right order $n-r$ block of Y_{n+1} is psd, hence it is feasible in ($\mathrm{D}_{\mathrm{strong}}$).

□

2.4 Proof of Theorem 1

By Theorem 2 we see that ($\mathrm{D}_{\mathrm{strong}}$) and ($\mathrm{D}_{\mathrm{Ram}}$) are equivalent in the following sense: one is feasible iff the other is; when they are feasible, their optimal values are equal; and when they are feasible, one attains its optimal value if and only if the other one does. Combining this argument with Lemma 3 implies Theorem 1. □

Example 4 (Examples 1, 2 and 3 continued) As we saw in Example 2

$$U_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, U_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, V_2 = \begin{pmatrix} 0 & 0 & -1/2 \\ 0 & 0 & 0 \\ -1/2 & 0 & 0 \end{pmatrix}. \tag{23}$$

with the U_3 and V_3 listed in (23) is feasible in the Ramana dual of (1). However, the Ramana dual of (1) has other feasible solutions as well.

Part (2) of Theorem 2 helps us describe all feasible solutions. First, we see that $(Y_1 := U_1, Y_2 := U_2 + V_2)$ is a regular facial reduction sequence and the sizes of the positive definite blocks in Y_1 and Y_2 are $r_1 = 1$ and $r_2 = 1$, respectively. Suppose we decompose Y_2 as

$$Y_2 = U_2' + V_2' \text{ with } U_2' \succeq 0,\ V_2' \in \tan(U_1'). \tag{24}$$

Then by part (2) of Theorem 2 we have $U_2' \in \mathcal{S}_+^{3,2}$. Indeed, all decompositions of Y_2 which obey (24) have

$$U_2' = \begin{pmatrix} \alpha & \beta & 0 \\ \beta & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, V_2' = Y_2 - U_2', \tag{25}$$

where the upper left 2×2 block of U_2 is psd. If we choose α and β so this block is positive definite, then $U_1, U_2', V_2' U_3, V_3$ is an optimal solution of ($\mathrm{D}_{\mathrm{Ram}}$).

3 Ramana's primal

In this section we assume that (D) is feasible.

3.1 Statement of Ramana's primal

First we state the Ramana dual of (D), which, with some abuse of terminology we call Ramana's primal:

Theorem 3 *Consider the optimization problem called* Ramana's primal

$$\begin{array}{llll} \sup & c^\top x & & \\ s.t. & B - \mathcal{A}x \in \mathcal{S}^n_+ + \tan(U_n) & & \\ & \mathcal{A}x_i = U_i + V_i & i = 1, \dots, n & \\ & c^\top x_i = 0 & i = 1, \dots, n & \\ & U_i \in \mathcal{S}^n_+ & i = 1, \dots, n+1 & \\ & V_i \in \tan(U_{i-1}) & i = 1, \dots, n+1 & \\ & U_0 = V_0 = 0, & & \end{array} \qquad (\mathrm{P}_{\mathrm{Ram}})$$

We have

$$\mathrm{val}\,(\mathrm{D}) = \mathrm{val}\,(\mathrm{P}_{\mathrm{Ram}}),$$

and $\mathrm{val}\,(\mathrm{P}_{\mathrm{Ram}})$ *is attained when finite.* □

3.2 Reformulations to certify the maximum rank solution in (D)

In this section we recall reformulations of (D) and show how they certify the maximum rank solution of a feasible solution in (D).

Definition 5 *We say that* (D) *is in partial rank revealing form,or partial RR form, if for some* $0 \le k \le m$

(1) $A_1, \dots, A_k$ *a regular facial reduction sequence*
(2) $c_1 = \dots = c_k = 0$

We say that (D) *is in complete rank revealing form, or complete RR form, if it is in partial RR form, and* $n - \sum_{\ell=1}^{k}$ *is the rank of the maximum rank solution in* (D).

We next explain the terminology "rank revealing form" Suppose Y is feasible in (D), and (D) is in partial RR form with $(A_1, \dots, A_k)$ a regular facial reduction sequence, in which the sizes of the positive definite blocks are $r_1, \dots, r_k$, respectively, and $c_1 = \dots = c_k = 0$. Then $\langle A_1, Y \rangle = 0$ implies that the first r_1 rows and columns of Y are zero; $\langle A_2, Y \rangle = 0$ implies the next r_2 rows and columns of Y are zero; etc.

Definition 6 *We say that we reformulate* (D) *if we apply the following operations (in any order):*

(1) Elementary row operations on the equations: multiply an equation $\langle A_i, Y\rangle = c_i$ by a nonzero scalar; exchange two equations; add a nonzero multiple of an equation to another equation.
(2) rotations applied to $A_1, \dots, A_m$ and B.

We also say that by reformulating (D) we obtain a reformulation.

Lemma 4 *We can always reformulate (D) to put it into complete RR form.* □

Lemma 4 was proved in Theorem 2 of [10]. We can give a simpler proof, by relying only on "strong duality holds under strict feasibility," similarly to how we proved Lemma 2. We do not give this proof (for the sake of brevity), but invite the reader to work out its details.

Example 5 To illustrate reformulations of (D), we consider a semidefinite system with data

$$A_1 = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} A_2 = \begin{pmatrix} 0 & 0 & -1 & 1 \\ 0 & 1 & -1 & 0 \\ -1 & -1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}, A_3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 2 \end{pmatrix},$$
$$A_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad c = (1, 1, 2, 1). \tag{26}$$

Suppose we perform the operations

$$\begin{aligned} (A_1, c_1) &:= (A_1, c_1) + (A_2, c_2) - (A_3, c_3) \\ (A_2, c_2) &:= (A_2, c_2) - (A_3, c_3) + (A_4, c_4) \end{aligned}$$

and obtain the problem with data

$$A_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, A_3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 2 \end{pmatrix},$$
$$A_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad c = (0, 0, 2, 1). \tag{27}$$

This system is in partial RR form, since (A_1, A_2) is a regular facial reduction sequence, and $c_1 = c_2 = 0$ [5]. Now the sizes of the positive definite blocks in

[5] Of course (A_1, A_2, A_3) is also a regular facial reduction sequence, but $c_3 \neq 0$.

A_1 and A_2 are $r_1 = r_2 = 1$. However, it is not in complete RR form, since there is no feasible Y with rank 2.

If we further perform

$$(A_3, c_3) := (A_3, c_3) - 2(A_4, c_4),$$

then we get the SDP with data

$$A_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \; A_2 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \; A_3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix},$$

$$A_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad c = (0, 0, 0, 1). \tag{28}$$

This system is in complete RR form, since the matrix Y in which the lower right element is $1,$ and all other elements are zero satisfies its constraints.

Proposition 4 *Reformulating (D) does not change the optimal value and attainment in (D), (P), and ($\mathrm{P}_{\mathrm{Ram}}$).*

Proof First consider (P) and (D). Rotations do not change their optimal value and attainment by Proposition 3. The same is true of elementary row operations by a straightforward argument.

To deal with ($\mathrm{P}_{\mathrm{Ram}}$), suppose $x_1, \dots, x_n, x$ is feasible in it (with some suitable U_i and V_i), before the reformulation. First suppose we perform an elementary row operation on (D). This amounts to replacing the equations $\mathcal{A}^* Y = c$ by $M \mathcal{A}^* Y = Mc$ where $M : \mathbb{R}^m \to \mathbb{R}^m$ is an invertible linear map. Thus, $M^\top x_1, \dots, M^\top x_n, M^\top x$ is feasible afterwards in ($\mathrm{P}_{\mathrm{Ram}}$), with the same objective value.

Next suppose we rotate the A_i and B by an orthonormal matrix Q. Also suppose we rotate all U_i and V_i by the same Q. Then by part (7) of Proposition 1 we see that $x, x_1, \dots, x_n$ with the rotated U_i and V_i is feasible in ($\mathrm{P}_{\mathrm{Ram}}$) after the rotation. □

Example 6 (Example 5 continued) Proposition 4 will help us analyze ($\mathrm{P}_{\mathrm{Ram}}$). It also helps to analyze pathological SDPs. For example, suppose we seek the infimum of y_{12} subject to the constraints of (26). We claim this SDP has optimal value zero, which is attained; however, the primal has no solution with value zero.

How do we prove this statement? For the original problem (26) this argument is actually not easy to carry out. However, for the reformulated version (28) it is straightforward.

3.3 The relaxed primal and the strong primal

Inn this section we introduce the relaxed and strong primals of (D). The latter is a counterpart of the strong dual (34).

Definition 7 *Suppose* (D) *is in partial RR form with* $A_1, \ldots, A_k$ *a regular facial reduction sequence. Let* r_i *be the order of the positive definite block in* A_i *for* $i = 1, \ldots, k$*. We call the optimization problem*

$$\begin{aligned} &\sup\ c^\top x \\ &s.t.\ S = B - \mathcal{A}x \\ &\qquad S_{22} \in \mathcal{S}^r_+. \end{aligned} \tag{29}$$

the relaxed primal of (D)*. Here* S_{22} *stands for the lower right order* $n - (r_1 + \cdots + r_k)$ *block of* S*. If* (D) *is in complete RR form, then we call* (29) *the* strong primal *of* (D)*.*

Lemma 5 *Suppose* (D) *is in partial RR form and* x *is feasible in its relaxed primal. Then*

$$\operatorname{val}(D) \geq c^\top x. \tag{30}$$

If (D) *is in complete RR form, then equality holds in* (30) *for some* x *feasible in the strong primal of* (D)*.* □

Lemma 5 can be proved similarly to Lemma 3.

Example 7 (Example 5 continued) Suppose again we set up an SDP which seeks inf y_{12} in the semidefinite system of (28). Thus the objective matrix B of this SDP has 1/2 in the $(1,2)$ and $(2,1)$ positions, and zeros everywhere else.

In the strong primal of this SDP only the lower right 1×1 corner of $B - \sum_{i=1}^4 x_i A_i$ must be psd (i.e., nonnegative). Thus the strong primal has value 0, which is attained.

3.4 Connecting reformulations to Ramana's primal

This section presents Theorem 4, which connects reformulations of (D) to Ramana's primal. Its proof has some common elements with the proof of Theorem 2, so we only sketch these portions.

Before stating it, we note that in a complete RR form of Definition 5 we may assume $k = n$. Indeed, a complete RR form with $r_1 + \cdots + r_k \leq n$ exists, since we can discard the A_i in which the size of the positive definite block is zero. If $k < n$, then we can just add $n - k$ redundant constraints of the form $\langle 0, Y \rangle = 0$ in the beginning of (D), and renumber the other A_i.

Theorem 4 *The following hold:*

(1) Suppose (D) *is in complete RR form with* $k = n$*, and* x *is feasible in its strong primal. Then* $e_1, \ldots, e_n, x$ *is feasible in* ($\mathrm{P}_{\mathrm{Ram}}$) *with some suitable* U_i *and* V_i*.*

(2) Suppose $x, x_1, \ldots, x_n$ *and* $(U_i, V_i)_{i=0}^{n+1}$ *is feasible in* ($\mathrm{P}_{\mathrm{Ram}}$)*. Then we can rotate all* A_i, B, U_i, V_i *so that after the rotation the following hold:*

(a) The SDP

$$\begin{array}{rl} \inf & \langle B, Y\rangle \\ \text{s.t.} & \langle \mathcal{A}x_i, Y\rangle = 0 \;\; \text{for } i = 1, \ldots, n \\ & \langle A_j, Y\rangle = c_j \text{ for } j \in J \end{array} \tag{31}$$

is in partial RR form for some suitable $J \subseteq \{1, \ldots, m\}$*,*

(b) $B - \mathcal{A}x$ *is feasible in its relaxed primal.*

Further, if after this rotation the sizes of the positive definite blocks in $\mathcal{A}x_1, \ldots, \mathcal{A}x_n$ *are* r_i *for* $i = 1, \ldots, n$*, then*

$$U_i \in \mathcal{S}_+^{r_1+\cdots+r_i, n} \text{ for } i = 1, \ldots, n. \tag{32}$$

Proof of (1) Let us make the assumptions. For $i = 1, \ldots, n$ we do the following. We let Λ_i be the positive definite block in A_i, assume the order of Λ_i is r_i and define U_i as in (20). We also define $V_i = A_i - U_i$. Using the same argument as in the proof of Theorem 2 we deduce $V_i \in \tan(U_{i-1})$ for all i.

Also, let $S := B - \mathcal{A}x$. Since x is feasible in the strong primal of (D), the lower right order $n - \sum_{\ell=1}^n r_i$ corner of S is psd. Thus $S \in \mathcal{S}_+^n + \tan(U_n)$. Summarizing, $x, e_1, \ldots, e_n$ with the U_i and V_i is feasible in ($\mathrm{P}_{\mathrm{Ram}}$), as required.

Proof of (2) Let us make the assumptions. First we prove (2a), i.e., that after a suitable rotation $(\mathcal{A}x_1, \ldots, \mathcal{A}x_n)$ is a regular facial reduction sequence and (32). We can do this just how we proved (2) and (19) in Theorem 2, so we omit the details. Since we also have $c^\top x_i = 0$ for $i = 1, \ldots, n,$ we can replace some of the equations in (D) by $\langle \mathcal{A}x_i, Y\rangle = 0$ to put it into partial RR form.

Also, after the rotation we still have $B - \mathcal{A}x \in \mathcal{S}_+^n + \tan(U_n)$. Since $U_n \in \mathcal{S}_+^{n, r_1+\cdots+r_n}$, we see that the lower right order $n - \sum_{\ell=1}^n r_\ell$ block of $B - \mathcal{A}x$ is psd. So it is feasible in the relaxed primal of (31), as required.

3.5 **Proof of Theorem 3**

By Lemma 4 we can put (D) into complete RR form. Hence for some x feasible in the strong primal we have

$$\mathrm{val}\,(\mathrm{D}) = c^\top x \le \mathrm{val}\,(\mathrm{P}_{\mathrm{Ram}}),$$

where the equality is from Lemma 5 and the inequality is from part (1) in Theorem 4. Also, ($\mathrm{P}_{\mathrm{Ram}}$) has a solution with value $c^\top x$ by part (1) in Theorem 4.

On the other hand, suppose $x, x_1, \ldots, x_n$ is feasible in ($\mathrm{P}_{\mathrm{Ram}}$) with some U_i and V_i. Then

$$\mathrm{val}\,(\mathrm{D}) \ge c^\top x \ge \mathrm{val}\,(\mathrm{P}_{\mathrm{Ram}}),$$

where the second inequality is obvious. The first inequality follows since by part (2) of Theorem 4 x is feasible in the relaxed primal of (31), hence Lemma 5 applies.

3.6 Discussion

The characterizations of Theorem **??** and 4 can turn out to be useful if (or hopefully when) Ramana's dual is implemented. On the one hand, a full implementation with all the U_i and V_i may be cumbersome. On the other hand, even an implementation using just a few U_i and V_i can close the duality gap, or lead to an attained dual optimal value in some pathological SDPs.

A Proof of Lemma 3

Suppose

$$Z = B - \mathcal{A}\bar{x} \text{ for some } \bar{x} \in \mathbb{R}^m,$$

and we replace B by Z in (P), i.e. subtract $\mathcal{A}\bar{x}$ from B. It is straightforward that by doing so we subtract $c^\top \bar{x}$ from the optimal value of (P) and (D_{strong}). Also, if the optimal value of (D_{strong}) is attained before this replacement, then it is also attained afterwards. Thus, for this proof we assume w.l.o.g. that the right hand side in (P) is Z.

For $i = 1, \ldots, m$ let $D_i \in \mathcal{S}^r$ be the lower right $r \times r$ submatrix of A_i and $E_i \in \mathcal{S}^n$ the matrix obtained by replacing D_i by the all zero matrix. Also, define the linear operators $\mathcal{D} : \mathbb{R}^m \to \mathcal{S}^r$ and $\mathcal{E} : \mathbb{R}^m \to \mathcal{S}^n$ as

$$\mathcal{D}x = \sum_{i=1}^m x_i D_i, \; \mathcal{E}x = \sum_{i=1}^m x_i E_i \text{ for } x \in \mathbb{R}^m.$$

Suppose x is feasible in (P). Note that the Y_i constructed in Lemma 2 satisfy equations (12)–(13) with Z in place of B. Thus, using the argument after the statement of Lemma 2 we see that the first $n - r$ rows and columns of $Z - \mathcal{A}x$ are zero, hence $\mathcal{E}x = 0$.

Let $\mathcal{G} : \mathcal{S}^n \to \mathbb{R}^m$ be a linear operator from $\mathbb{R}^m$ to $\mathcal{S}^n$ such that $\mathcal{N}(\mathcal{E}) = \mathcal{R}(\mathcal{G})$. Thus we can write any x feasible solution in (P) as $x = \mathcal{G}U$ for some $U \in \mathcal{S}^n$. So (P)–(D) are equivalent to the primal-dual pair

$$\begin{array}{llcll} & \sup\ c^\top(\mathcal{G}U) = \langle \mathcal{G}^* c, U\rangle & & \inf\ \langle \Lambda, Y_{22}\rangle & \\ (33) & s.t.\ \mathcal{D}\mathcal{G}U \preceq \Lambda & & s.t.\ \mathcal{G}^*\mathcal{D}^*Y_{22} = \mathcal{G}^* c & (34) \\ & \quad U \in \mathcal{S}^n & & \quad Y \succeq 0. & \end{array}$$

Since (33) is strictly feasible, the optimal values of (33) and (34) agree, and the latter is attained when finite. The constraints in (34) are equivalent to

$$\mathcal{A}^*Y - c \in \mathcal{N}(\mathcal{G}^*) = \mathcal{R}(\mathcal{E}^*).$$

Thus, Y is feasible in (34) if and only if there is $V \in \mathcal{S}^n$ such that

$$\langle D_i, Y_{22}\rangle + \langle E_i, V\rangle = c_i \text{ for all } i. \tag{35}$$

Thus (34) is equivalent to (D_{strong}), as wanted. □

References

1. Borwein, J.M., Wolkowicz, H.: Regularizing the abstract convex program. J. Math. Anal. App. **83**, 495–530 (1981) 5, 10
2. Burer, S., Monteiro, R.D.: A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. Mathematical programming **95**(2), 329–357 (2003) 2
3. De Klerk, E.: Aspects of semidefinite programming: interior point algorithms and selected applications, vol. 65. Springer Science & Business Media (2006) 5
4. De Klerk, E., Terlaky, T., Roos, K.: Self-dual embeddings. In: Handbook of Semidefinite Programming, pp. 111–138. Springer (2000) 5
5. Drusvyatskiy, D., Wolkowicz, H., et al.: The many faces of degeneracy in conic optimization. Foundations and Trends® in Optimization **3**(2), 77–170 (2017) 5
6. Helmberg, C., Rendl, F.: A spectral bundle method for semidefinite programming. SIAM Journal on Optimization **10**(3), 673–696 (2000) 2
7. Henrion, D., Naldi, S., El Din, M.S.: Exact algorithms for linear matrix inequalities. SIAM Journal on Optimization **26**(4), 2512–2539 (2016)
8. Klep, I., Schweighofer, M.: An exact duality theory for semidefinite programming based on sums of squares. Math. Oper. Res. **38**(3), 569–590 (2013) 5
9. Laurent, M., Rendl, F.: Semidefinite programming and integer programming. Handbooks in Operations Research and Management Science **12**, 393–514 (2005) 5
10. Liu, M., Pataki, G.: Exact duality in semidefinite programming based on elementary reformulations. SIAM J. Opt. **25(3)**, 1441–1454 (2015) 6, 15
11. Lourenço, B.F., Pataki, G.: A simplified treatment of Ramana's exact dual for semidefinite programming. Optimization Letters **17**(2), 219–243 (2023) 5, 6
12. Lovász, L.: Graphs and geometry, vol. 65. American Mathematical Soc. (2019) 5
13. Luo, Z.Q., Sturm, J., Zhang, S.: Duality results for conic convex programming. Tech. Rep. Report 9719/A, Erasmus University Rotterdam, Econometric Institute, The Netherlands (1997) 5
14. Nemirovski, A.: Advances in convex optimization: conic programming. In: International Congress of Mathematicians, vol. 1, pp. 413–444 (2007) 5
15. Nesterov, Y.E., Nemirovskii, A.S.: Interior Point Polynomial Algorithms in Convex Programming. SIAM Publications. SIAM, Philadelphia, USA (1994) 2
16. O'Donnell, R.: SOS is not obviously automatizable, even approximately. In: 8th Innovations in Theoretical Computer Science Conference (ITCS 2017). Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik (2017) 5
17. Pataki, G.: A simple derivation of a facial reduction algorithm and extended dual systems. Tech. rep., Columbia University (2000) 5, 6
18. Pataki, G.: Strong duality in conic linear programming: facial reduction and extended duals. In: D. Bailey, H.H. Bauschke, F. Garvan, M. Théra, J.D. Vanderwerff, H. Wolkowicz (eds.) Proceedings of Jonfest: a conference in honour of the 60th birthday of Jon Borwein. Springer, also available from http://arxiv.org/abs/1301.7717 (2013) 5, 6
19. Pataki, G., Touzov, A.: An echelon form of weakly infeasible semidefinite programs and bad projections of the psd cone. Foundations of Computational Mathematics pp. 1–38 (2022) 7
20. Ramana, M.V.: An exact duality theory for semidefinite programming and its complexity implications. Math. Program. Ser. B **77**, 129–162 (1997) 3
21. Ramana, M.V., Freund, R.: On the ELSD duality theory for SDP. Tech. rep., MIT (1996) 5
22. Ramana, M.V., Tunçel, L., Wolkowicz, H.: Strong duality for semidefinite programming. SIAM J. Opt. **7**(3), 641–662 (1997) 5, 6
23. Renegar, J.: A Mathematical View of Interior-Point Methods in Convex Optimization. MPS-SIAM Series on Optimization. SIAM, Philadelphia, USA (2001) 2, 3
24. Sturm, J.F.: Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones. Optimization methods and software **11**(1-4), 625–653 (1999) 2
25. Toh, K.C., Todd, M.J., Tütüncü, R.H.: Sdpt3?a matlab software package for semidefinite programming, version 1.3. Optimization methods and software **11**(1-4), 545–581 (1999) 2

26. Vandenberghe, L., Boyd, S.: Semidefinite programming. SIAM review **38**(1), 49–95 (1996) 5
27. Yang, L., Sun, D., Toh, K.C.: Sdpnal+: a majorized semismooth newton-cg augmented lagrangian method for semidefinite programming with nonnegative constraints. Mathematical Programming Computation **7**(3), 331–366 (2015) 2