IRREDUCIBLE INFEASIBLE SUBSYSTEMS OF SEMIDEFINITE SYSTEMS

KAI KELLNER, MARC E. PFETSCH, AND THORSTEN THEOBALD

ABSTRACT. Farkas' lemma for semidefinite programming characterizes semidefinite feasibility of linear matrix pencils in terms of an alternative spectrahedron. In the well-studied special case of linear programming, a theorem by Gleeson and Ryan states that the index sets of irreducible infeasible subsystems are exactly the supports of the vertices of the corresponding alternative polyhedron.

We show that one direction of this theorem can be generalized to the nonlinear situation of extreme points of general spectrahedra. The reverse direction, however, is not true in general, which we show by means of counterexamples. On the positive side, an irreducible infeasible block subsystem is obtained whenever the extreme point has minimal block support. Motivated by results from sparse recovery, we provide a criterion for the uniqueness of solutions of semidefinite block systems.

1. Introduction

The structure of infeasible linear inequality systems is quite well understood. In particular, Farkas' Lemma, also called Theorem of the Alternative, gives a characterization of infeasibility (see, e.g., [1]). Moreover, the basic building blocks are so-called Irreducible Infeasible Systems (IISs, also called Irreducible Inconsistent Systems), i.e., infeasible subsystems such that every proper subsystem is feasible. An extension of the Theorem of the Alternative, due to Gleeson and Ryan [2], states that the IISs of an infeasible linear inequality system correspond exactly to the vertices of a so-called alternative polyhedron (see Theorem 3.4). These IISs provide a means to analyze infeasibilities of a system, see, e.g., [3, 4, 5] and the book [6]. Today, standard optimization software can compute (hopefully) small IISs. Further investigations include the mixed-integer case [7] and the application within Benders' decomposition [8].

In this article, we consider infeasible systems in semidefinite form. There are well-known generalizations of the Theorem of the Alternative to this setting (see, e.g., [9]), although one has to be more careful, since feasibility might only be attained in the limit – see Proposition 2.1 for a more precise statement. As in the linear case, solutions of certain alternative systems give a certificate of the (weak) infeasibility of a semidefinite system.

In this context, the following natural questions arise: How can infeasible semidefinite systems be analyzed? What can be said about the structure of irreducible infeasible semidefinite systems? Moreover, is there a generalization of the theorem of Gleeson and Ryan to this setting?

Date: January 29, 2019.

These questions are motivated by solving mixed-integer semidefinite programs using branch-and-bound in which an SDP is solved in every node (see, e.g., [10]). Then, it often happens that these SDPs turn out to be infeasible. One would now like to learn from this infeasibility in order to strengthen the relaxations of other nodes. This is done in mixed-integer and SAT solvers, see, e.g., [11, 12].

To come up with an appropriate definition of an IIS for a semidefinite system it appears to be very natural to consider block systems. Then, an IIS is given by an inclusion minimal set of infeasible block subsystems (see Definition 2.6). We will show in Section 3 that one direction of the above mentioned connection can be generalized: there always exists an extreme point of the alternative system that corresponds to a given IIS, see Theorem 3.5. The reverse direction, however, is not true in general, which we show and discuss via two counterexamples, see Examples 3.6 and 3.7. On the positive side, whenever an extreme point has (inclusionwise) minimal block support, the corresponding subsystem forms indeed an IIS. This leads to the general task to compute such points.

In the particular case in which the alternative semidefinite system has a unique solution, this algorithmic challenge simplifies to solving one semidefinite program. Motivated by results from sparse recovery, we provide a criterion for the uniqueness of solutions of semidefinite block systems. In Section 4, we generalize the results in [13, 14, 15, 16] to give unique recovery characterizations for a block semidefinite system in Theorem 4.1. Further perspectives and open questions are given in Section 5.

2. Infeasible Systems and Block Structure

We use the following notation. Let $\mathcal{S}^n$ be the set of all (real) symmetric $n \times n$ matrices, and "$\succeq$" denotes that a symmetric matrix is positive semidefinite (psd). For a matrix $A \in \mathcal{S}^n$ and $I \subseteq [n] := \{1, \ldots, n\}$, let A_I be the submatrix containing the rows and columns of A indexed by I. For A, $B \in \mathcal{S}^n$, we denote the inner product by

$$A \bullet B = \operatorname{tr}(A^\top B) = \sum_{i,j=1}^{n} A_{ij}\, B_{ij},$$

where $\operatorname{tr}(\cdot)$ denotes the trace. $\|A\|_2$ denotes the operator norm $\|A\|_2 = \max_{1 \leq j \leq n} |\lambda_j(A)|$, where $\lambda_1(A), \ldots, \lambda_n(A)$ are the eigenvalues of A.

Throughout the paper, let $A_0, \ldots, A_m \in \mathcal{S}^n$. For $y \in \mathbb{R}^m$, we consider the linear *(matrix) pencil*

$$A(y) := A_0 - \sum_{i=1}^{m} y_i A_i$$

and the *linear matrix inequality* (LMI) $A(y) \succeq 0$. With respect to infeasibility, we will use the following result, where $\mathbb{I}$ denotes the identity matrix.

Proposition 2.1. *Either $A(y) + \varepsilon \mathbb{I} \succeq 0$ is feasible for every $\varepsilon > 0$ or there exists $X \succeq 0$ with $A_i \bullet X = 0$, $i \in [m]$, and $A_0 \bullet X = -1$.*

This statement is equivalent to Sturm's Farkas' Lemma for semidefinite programming (see Lemmas 3.1.1 and 3.1.2 in [17]) and a variation of Theorem 2.21 in [9], and its proof is provided for completeness.

Proof. Consider the following dual pair of semidefinite programs (SDPs):

$$\inf\,\{\eta\;:\;A(y)+\eta\,\mathbb{I}\succeq 0,\;\eta\geq 0\},\tag{2.1}$$

$$\sup\,\{-A_0\bullet X\;:\;A_i\bullet X=0,\;i\in[m],\;\operatorname{tr}(X)\leq 1,\;X\succeq 0\}.\tag{2.2}$$

Setting $y = 0$, $\eta = \|A_0\|_2 + 1$ shows that (2.1) has a Slater point. Moreover, $X = 0$ is feasible for (2.2). The strong duality theorem (see, e.g., Theorem 2.14 in [9]) implies that (2.2) attains its optimal value and the objective values are the same.

Suppose that no $X \succeq 0$ with $A_i \bullet X = 0$, $i \in [m]$, $A_0 \bullet X = -1$ exists. By scaling, this implies that no such X exists with $A_0 \bullet X < 0$. And since the zero matrix is feasible for (2.2), the optimal value of (2.2) is 0. By the strong duality theorem, the optimal value of (2.1) is also 0. Either (2.1) attains this value and we are done, or there exists a sequence (y^k, η_k) such that $A(y^k) + \eta_k\,\mathbb{I} \succeq 0$ and $\eta_k \searrow 0$. This implies the theorem. ☐

Remark 2.2. In slight deviation from parts of the literature, we call $A(y) \succeq 0$ *weakly feasible*, if for every $\varepsilon > 0$ the system $A(y) + \varepsilon\mathbb{I} \succeq 0$ is feasible; compare this, for instance, to the definition in [9], which requires $\tilde{A}_0 - \sum_{i=1}^m y_i\, A_i \succeq 0$ to be feasible for some $\tilde{A}_0$ such that $\|A_0 - \tilde{A}_0\|_2 < \varepsilon$. Moreover, $A(y) \succeq 0$ is *weakly infeasible* if it is not weakly feasible. Note the slight inaccuracy of this naming convention, which should, however, not lead to confusion in the present paper.

Corollary 2.3. *Assume that there exists $\bar{X} \succ 0$ with $A_i \bullet \bar{X} = 0$, $i \in [m]$. Then, either $A(y) \succeq 0$ is feasible or there exists $X \succeq 0$ with $A_i \bullet X = 0$, $i \in [m]$, and $A_0 \bullet X = -1$.*

Proof. By scaling $\bar{X}$ to satisfy $\operatorname{tr}(\bar{X}) \leq 1$, the assumption guarantees that (2.2) above has a Slater point and therefore that the optimal value of (2.1) is attained, see, e.g., Corollary 2.17 in [9]. The remaining part of the proof is as for the one of Proposition 2.1. ☐

Our subsequent definition of an alternative spectrahedron will allow to handle structured semidefinite systems. To motivate this viewpoint, consider a simple example where the goal is to check whether two given halfplanes $H_i = \{y \in \mathbb{R}^2 \,:\, \alpha_i y_1 + \beta_i y_2 + \gamma_i \geq 0\}$ ($i \in [2]$) and a given disc $D = \{y \in \mathbb{R}^2 \,:\, \|y - c\|_2^2 \leq r^2\}$ in the Euclidean plane $\mathbb{R}^2$ with center $c \in \mathbb{R}^2$ have a common point. The smallest LMI-representation (w.r.t. matrix size) of D is

$$K(r,c;y) \coloneqq \begin{pmatrix} r + c_1 - y_1 & y_2 - c_2 \\ y_2 - c_2 & r - c_1 + y_1 \end{pmatrix} \succeq 0,$$

and thus the existence of a point in $H_1 \cap H_2 \cap D$ is equivalent to the feasibility of the LMI

$$A(y) = \begin{pmatrix} \alpha_1 y_1 + \beta_1 y_2 + \gamma_1 & & \\ & \alpha_2 y_1 + \beta_2 y_2 + \gamma_2 & \\ & & K(r,c;y) \end{pmatrix} \succeq 0.\tag{2.3}$$

In order to capture such natural structure within semidefinite systems, one arrives at block systems. In particular, already in the simple example this allows then to consider the 2×2-subsystem of the disc as an entity. Formally, this yields the following.

Definition 2.4. Let $k \geq 1$ and $B_1, \dots, B_k \neq \emptyset$ a partition of the set $[n]$. A linear pencil $A(y)$ is in *block-diagonal form* with blocks $B_1, \dots, B_k$ if and only if each A_i is 0 outside of the blocks $B_1, \dots, B_k$, i.e., $(A_i)_{st} = 0$ for all $(s,t) \notin (B_1 \times B_1) \cup \cdots \cup (B_k \times B_k)$ and all $i \in [m]$. Note that the blocks might be *decomposable*, i.e., at least one block consists of blocks of smaller size while still retaining the block structure of $A(y)$.

Assumption 2.5. To avoid trivial infeasibilities, we will assume that for each block B_i, $i \in [k]$, there exists $y \in \mathbb{R}^m$ such that $A(y)_{B_i} \succeq 0$ is weakly feasible.

Definition 2.6. Let $A(y)$ be in block-diagonal form with blocks $B_1, \dots, B_k$.

(a) For $I \subseteq [k]$, the *block subsystem of* $A(y)$ *with respect to* I is given by $A(y)_{B(I)}$ for the index set $B(I) \coloneqq \bigcup_{i \in I} B_i$. By convention, $B(\emptyset) = \emptyset$ and $A(y)_\emptyset$ is a feasible system.

(b) A block subsystem with respect to some $I \subseteq [k]$ is an *irreducible infeasible subsystem (IIS)* iff $A(y)_{B(I)} \succeq 0$ is weakly infeasible, but $A(y)_{B(I')} \succeq 0$ is weakly feasible for all $I' \subsetneq I$.

(c) Given a matrix $X \in \mathcal{S}^n$, its *block support* $\mathrm{BS}(X)$ is defined as

$$\mathrm{BS}(X) \coloneqq \{i \in [k] \,:\, X_{B_i} \neq 0\}.$$

Remark 2.7. Linear inequality systems arise if all matrices $A_0, \dots, A_m$ of $A(y)$ are diagonal. In this case, each inequality is of the form

$$(A_0)_{jj} - \sum_{i=1}^{m} y_i \, (A_i)_{jj} \geq 0, \; j \in [n].$$

If this system is written as $Dx \leq d$, then IISs correspond to infeasible subsystems of $Dx \leq d$ such that each proper subsystem is feasible.

The linear case arises, in particular, if the block system satisfies $k = n$ (and hence $|B_i| = 1$); then the blocks are not decomposable. However, it is also possible that the blocks are decomposable. In this case, the system consists of k linear inequality systems $D_1 x \leq d_1, \dots, D_k x \leq d_k$, each defining a polyhedron. If the intersection of these polyhedra is empty, the original LMI is infeasible; see Example 3.6 below.

Remark 2.8. An alternative way to define IISs would be to consider subsets $S \subseteq [n]$ such that $A(y)_S \succeq 0$ is (weakly) infeasible, but $A(y)_{\hat{S}} \succeq 0$ is (weakly) feasible for every proper subset $\hat{S}$ of S. However, this definition would not retain the structure within semidefinite systems such as (2.3). As a consequence, we currently do not know to which extent our subsequent investigations can be transferred to that model.

3. Alternative Systems

In view of Proposition 2.1, we define the following, where the abbreviation Σ for the LMI $A(y) \succeq 0$ will allow for a convenient notation. For general background on spectrahedra, we refer to [18, 19].

Definition 3.1. The *alternative spectrahedron* for $\Sigma : A(y) \succeq 0$ is

$$S(\Sigma) \coloneqq \{X \succeq 0 \,:\, A_i \bullet X = 0, \; i \in [m], \; A_0 \bullet X = -1\}.$$

Assumption 3.2. By standard polarity theory, a block structure of the system can also be assumed for $X \in S(\Sigma)$. Thus, we only consider matrices $X \in S(\Sigma)$ in block-diagonal form, where the blocks are indexed by $\mathrm{BS}(X)$.

The definition of the alternative spectrahedron immediately implies:

Lemma 3.3. *Let $\Sigma : A(y) \succeq 0$ be a weakly infeasible semidefinite system with blocks $B_1, \ldots, B_k$.*

(a) For any $X \in S(\Sigma)$, there exists an infeasible subsystem of Σ with block support contained in $\mathrm{BS}(X)$.

(b) For any $X \in S(\Sigma)$ with inclusion-minimal block support, the index set $\mathrm{BS}(X)$ defines an IIS of Σ.

As mentioned in the introduction, in the linear case there exists a characterization of IISs:

Theorem 3.4 (Gleeson and Ryan [2])**.** *Consider an infeasible system $\Sigma : Ax \le b$, where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$. The index sets of the IISs of Σ are exactly the support sets of the vertices of the alternative polyhedron*

$$P(\Sigma) = \{y \in \mathbb{R}^m \,:\, y^\top A = 0,\ y^\top b = -1,\ y \ge 0\}.$$

A proof can be found in [2] and [20]. Note that in the non decomposable linear case, $P(\Sigma)$ is equivalent to the alternative spectrahedron $S(\Sigma)$.

One goal of this paper is to investigate whether/how far Theorem 3.4 generalizes to the spectrahedral situation. We can show that one of the directions can be generalized.

Theorem 3.5. *Let $\Sigma : A(y) \succeq 0$ be a weakly infeasible LMI with blocks $B_1, \ldots, B_k$. For each index set I of an IIS, there exists an extremal point of $S(\Sigma)$ with block support I.*

The following proof proceeds by revealing the convex-geometric structure of the alternative spectrahedron.

Proof. Without any loss of generality, we can assume that $I = \{1, \ldots, t\}$ for some $t \in [k]$. By Proposition 2.1, the alternative spectrahedron $S(\Sigma)$ contains a feasible point X supported exactly on the blocks $B_1, \ldots, B_t$. In order to show that the alternative spectrahedron contains an extremal point with block support $\{1, \ldots, t\}$, we first observe that $S(\Sigma)$ has at least one extremal point. This follows from the fact that the positive semidefinite cone is pointed and thus any slice of a subspace with this cone cannot have a nontrivial lineality space either.

By Theorem 18.5 in [21], the alternative spectrahedron can be written in the form

$$S(\Sigma) = \mathrm{conv}(E \cup F),$$

where E is the set of its extremal points and F is the set of extremal directions of $S(\Sigma)$. Hence, by a general version of Carathéodory's Theorem (see Theorem 17.1 in [21]), there exist $r \ge 1$, $s \ge 0$, extremal points $V^{(1)}, \ldots, V^{(r)}$ and extremal rays $W^{(1)}, \ldots, W^{(s)}$ of the alternative spectrahedron such that

$$X \;=\; \sum_{i=1}^{r} \lambda_i \, V^{(i)} + \sum_{j=1}^{s} \mu_j \, W^{(j)}$$

with λ_i, $\mu_j > 0$ and $\sum_{i=1}^{r} \lambda_i = 1$. Since $V^{(i)}, W^{(j)}$ are positive semidefinite and λ_i, $\mu_j > 0$, the block support of each $V^{(i)}$, $W^{(j)}$ must be contained in the block support of X. Due to

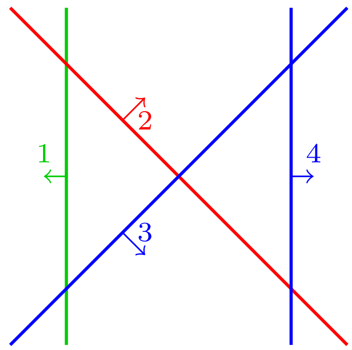


FIGURE 1. Illustration for Example 3.6.

the minimality of I, all $V^{(i)}$ must have the same block support. Hence, the block support of $V^{(1)}$ is exactly I, so that it is an extremal point with the desired property. □

We also provide the following shorter proof, which, however, reveals less structural insights.

Alternative proof. Consider the intersection

$$S' := S(\Sigma) \cap \{X \,:\, X_{B_i} = 0,\ i \notin I\}.$$

Then, S' has an extreme point $X' \in S'$, since it is the intersection of the pointed positive semidefinite cone with an affine space and therefore also pointed. Now let $I' := \mathrm{BS}(X')$ be the block support of X'. Then, $I' \subseteq I$ by construction. If $I' = I$, then we are done, since X' is an extreme point of $S(\Sigma)$ as well: Assume $X' = \lambda Z + (1-\lambda)Y$, $0 < \lambda < 1$, would be the strict convex combination of two other feasible points Y and Z, such that w.l.o.g. Z has a support in a block B outside of I'. Then,

$$\underbrace{\mathrm{tr}(X'_B)}_{=0} = \lambda \underbrace{\mathrm{tr}(Z_B)}_{>0} + (1-\lambda) \underbrace{\mathrm{tr}(Y_B)}_{\geq 0},$$

would give a contradiction.

Moreover, if $I' \subsetneq I$, then X' shows that $A(y)_{B(I')} \succeq 0$ is infeasible. Thus, I would not be minimal. □

The converse of this theorem is, however, not true in general. This direction may already fail in the presence of blocks of size 2. We will demonstrate this by two counterexamples. The first one is linear, but decomposable. The second one is not decomposable, but nonlinear.

Example 3.6. Let $m = 2$, $k = 3$ and

$$A_0 = \begin{pmatrix} 0 & & & \\ & -1 & & \\ & & -1 & 0 \\ & & 0 & -2 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & 0 \\ & & 0 & -1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & & & \\ & -1 & & \\ & & 1 & 0 \\ & & 0 & 0 \end{pmatrix}.$$

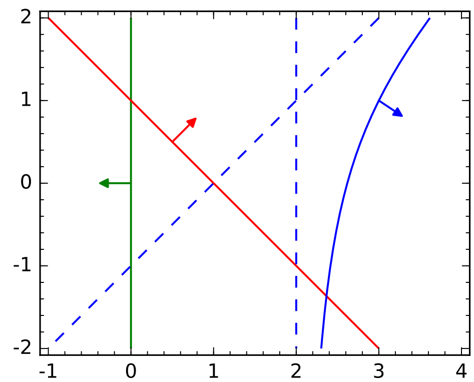


FIGURE 2. Deformation of Example 3.6, $\varepsilon = 1$. For $\varepsilon = 0$, the quadratic curve degenerates to Figure 1, as illustrated by the dashed blue lines.

The blocks are $B_1 = \{1\}$, $B_2 = \{2\}$, $B_3 = \{3, 4\}$, and this example corresponds to the three polyhedra

$$P_1 := \{y \in \mathbb{R}^2 \,:\, y_1 \le 0\}, \quad P_2 := \{y \in \mathbb{R}^2 \,:\, y_1 + y_2 \ge 1\},$$
$$P_3 := \{y \in \mathbb{R}^2 \,:\, -y_1 + y_2 \le -1,\ y_1 \ge 2\},$$

see Figure 1 for an illustration. In this case, only the diagonal elements of the points X in the alternative spectrahedron are relevant, which can be formulated as the polyhedron

$$S(\Sigma) := \left\{ x \in \mathbb{R}^{1+1+2} \,:\, \begin{pmatrix} 0 & -1 & -1 & -2 \\ 1 & -1 & -1 & -1 \\ 0 & -1 & 1 & 0 \end{pmatrix} x = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix},\ x \ge 0 \right\}.$$

$S(\Sigma)$ is a one-dimensional polytope with the two vertices

$$(1, \tfrac{1}{2}, \tfrac{1}{2}, 0)^\top \quad \text{and} \quad (\tfrac{1}{2}, 0, 0, \tfrac{1}{2})^\top.$$

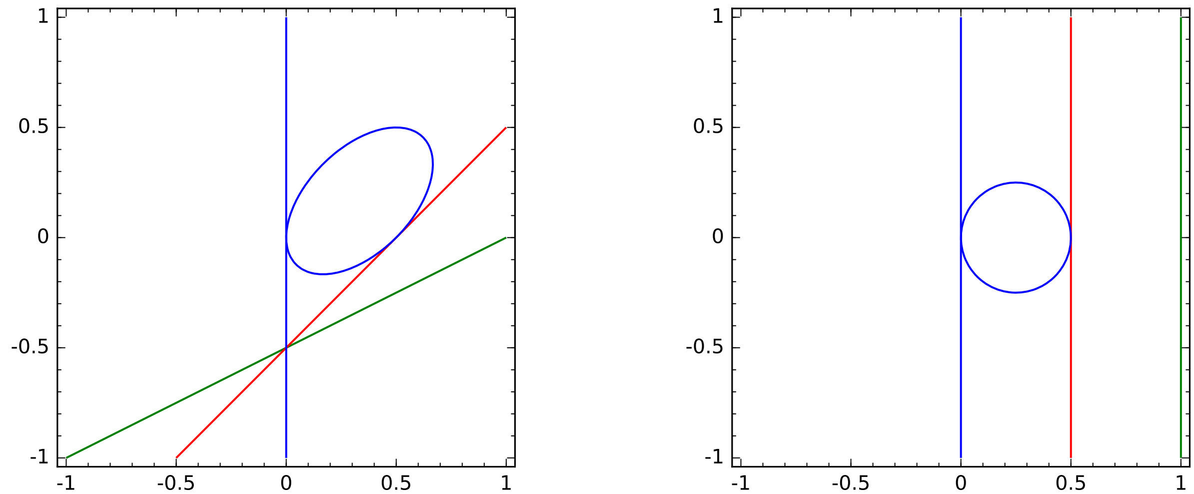
For the vertex $\tilde{x} = (1, \frac{1}{2}, \frac{1}{2}, 0)^\top$ of $S(\Sigma)$, we have $\mathrm{BS}(\tilde{x}) = \{1, 2, 3\}$. However, this does not correspond to an IIS, since $\{1, 3\}$ gives a proper subsystem that is infeasible.

To come up with non-decomposable blocks, the next counterexample deals with a deformed version.

Example 3.7. For $\varepsilon \ge 0$, consider the linear matrix pencil given by

$$A_0 = \begin{pmatrix} 0 & & & \\ & -1 & & \\ & & -1 & \varepsilon \\ & & \varepsilon & -2 \end{pmatrix}$$

and the matrices A_1 and A_2 of Example 3.6. For $\varepsilon = 0$, the system $\Sigma : A(y) \succeq 0$ specializes to Example 3.6. For $\varepsilon > 0$ the two lines in Figure 1 indexed by 3 and 4 deform to a quadratic curve; see Figure 2. Note that the quadratic curve has a second component corresponding to the lower right block being negative definite.



(A) $\varepsilon = 1$. (B) $\varepsilon = 0$.

FIGURE 3. Alternative spectrahedron for Example 3.7 in (X_{44}, X_{34})-coordinates: X_{44} is on the horizontal axis and X_{34} on the vertical axis. The straight lines in green, red and blue correspond to to (3.1), (3.2) and (3.3), respectively.

The alternative spectrahedron $S(\Sigma)$ is given by the set of symmetric block matrices

$$X = \operatorname{diag}\left(\begin{bmatrix} X_{11} \end{bmatrix}, \begin{bmatrix} X_{22} \end{bmatrix}, \begin{bmatrix} X_{33} & X_{34} \\ X_{34} & X_{44} \end{bmatrix} \right)$$

satisfying

$$X_{11} = 1 - X_{44} + 2\,\varepsilon X_{34} \geq 0, \tag{3.1}$$

$$X_{22} = X_{33} = \tfrac{1}{2} - X_{44} + \varepsilon X_{34} \geq 0, \tag{3.2}$$

$$X_{44} \geq 0, \tag{3.3}$$

$$(\tfrac{1}{2} - X_{44} + \varepsilon X_{34}) \cdot X_{44} - X_{34}^2 \geq 0. \tag{3.4}$$

In (X_{44}, X_{34})-coordinates, $S(\Sigma)$ is the set bounded by the ellipse in Figure 3 (for $\varepsilon = 1$). For $\varepsilon = 0$, the ellipse becomes a circle. Independent of ε, i.e., for any $\varepsilon \geq 0$, there are two distinguished extreme points, namely $(X_{44}, X_{34}) = (0, 0)$ and $(X_{44}, X_{34}) = (\frac{1}{2}, 0)$, corresponding to the matrices

$$\begin{pmatrix} 1 & & & \\ & \frac{1}{2} & & \\ & & \frac{1}{2} & 0 \\ & & 0 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \frac{1}{2} & & & \\ & 0 & & \\ & & 0 & 0 \\ & & 0 & \frac{1}{2} \end{pmatrix}$$

in $S(\Sigma)$. The diagonals of these matrices are exactly the two vertices of the alternative polyhedron as in Example 3.6. While the right matrix corresponds to an IIS, the left matrix does not.

These two examples motivate the question of how to compute IISs. By Lemma 3.3 it would suffice to compute a solution with minimal block support. This can be obtained by a greedy approach in which one iteratively solves semidefinite programs and fixes blocks to 0. Note, however, that computing an IIS with minimal cardinality block support is NP-hard already in the linear case, see [22].

In the particular case in which the alternative semidefinite system has a unique solution, this algorithmic challenge simplifies to solving one semidefinite program. In the next section we discuss universal conditions under which the alternative semidefinite system has a unique solution.

4. Universal Unique Solutions of Alternative Semidefinite Systems

For a block matrix $X \in \mathcal{S}^n$ with blocks $B_1, \dots, B_k$, denote by $\sigma_+^B(X)$ the number of blocks with at least one positive eigenvalue of X and by $\sigma_-^B(X)$ the number of blocks with at least one negative eigenvalue. Note that in case of a positive semidefinite matrix X, the value $\sigma_+^B(X)$ coincides with $|\mathrm{BS}(X)|$.

The following statement is a generalization of Theorem 1 in [15] to the case of block semidefinite systems. Our proof employs a block semidefinite generalization of standard techniques from linear systems. See also [14] for a variant for linear systems and Theorem 5 in [16] for a different (and non-block) generalization of that theorem to the semidefinite case.

Theorem 4.1. *For all psd block matrices $X^0 \in \mathcal{S}^n$ with $\sigma_+^B(X^0) \leq t$, the set*

$$\{X \succeq 0 \,:\, A_i \bullet X = A_i \bullet X^0,\, i \in [m]\}$$

is a singleton if and only if for all symmetric $V \neq 0$, with $A_i \bullet V = 0$, $i \in [m]$, we have $\sigma_+^B(V) > t$ and $\sigma_-^B(V) > t$.

Proof. Assume w.l.o.g. that there exists a symmetric $V \neq 0$ with $A_i \bullet V = 0$, $i \in [m]$, and $\sigma_-^B(V) \leq t$. The proof of the case $\sigma_+^B(V) \leq t$ is analogous, since the mapping $V \mapsto -V$ exchanges positive and negative eigenvalues and we have $A_i \bullet (-V) = 0$, $i \in [m]$, as well. For simplicity we further assume $\sigma_-^B(V) = t$. Then, there exists a decomposition

$$V \;=\; S^\top D S,$$

where S is a regular block matrix (with respect to the blocks $B_1, \dots, B_k$) and $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ where λ_i are the eigenvalues of V. In fact S can be assumed to be orthonormal ($S^\top = S^{-1}$) by performing a principal axis transformation for each block and combining the parts.

By reordering we can assume that the negative eigenvalues appear in the first t blocks. We then define the diagonal matrices D^1, $D^2 \in \mathbb{R}^{n\times n}$ with

$$D^1_{ii} \coloneqq \begin{cases} -\lambda_i & \text{if } \lambda_i < 0,\\ 0 & \text{otherwise,}\end{cases} \qquad D^2_{ii} \coloneqq \begin{cases} \lambda_i & \text{if } \lambda_i > 0,\\ 0 & \text{otherwise,}\end{cases} \qquad i \in [n]\,.$$

Then, $D^2 - D^1 = D$. We now obtain the block matrices

$$X^1 = S^\top D^1 S, \qquad X^2 = S^\top D^2 S,$$

with $X^1 \neq X^2$, $X^1 \succeq 0$, $X^2 \succeq 0$, and $\sigma_+^B(X^1) = t$. By construction $X^1_{B_i} = 0$ for all $i = t+1, \dots, k$. Moreover, for $i \in [m]$, we have

$$A_i \bullet X^1 = A_i \bullet (S^\top D^1 S) = A_i \bullet (S^\top (D^2 - D) S).$$

Since $A_i \bullet (S^\top D S) = A_i \bullet V = 0$, this implies

$$A_i \bullet X^1 = A_i \bullet (S^\top D^2 S) = A_i \bullet X^2.$$

Hence, the set $\{X \succeq 0 : A_i \bullet X = A_i \bullet X^1,\ i \in [m]\}$ also contains X^2 and is thus not a singleton.

Conversely, assume that there exists a psd matrix X^0 with $\sigma_+^B(X^0) \le t$ such that

$$\{X \succeq 0 : A_i \bullet X = A_i \bullet X^0,\ i \in [m]\}$$

is not a singleton. That is, there exists a matrix $\bar{X} \succeq 0$ with

$$A_i \bullet \bar{X} = A_i \bullet X^0, \quad i \in [m]$$

and $\bar{X} \neq X^0$. By the principal axis transformation, X^0 can be written as

$$X^0 = S^\top D^0 S$$

with an orthonormal block matrix S (w.r.t. the blocks $B_1, \dots, B_k$) and $D^0_{B_i} \ge 0$, $D^0_{B_i} \neq 0$ for $i = 1, \dots, t$ and $D^0_{B_i} = 0$ for $i = t+1, \dots, k$. Setting $\bar{Y} = S \bar{X} S^\top$, we have $\bar{Y} \succeq 0$ and $\bar{X} = S^\top \bar{Y} S$.

The block matrix $V = \bar{X} - X^0$ then satisfies $A_i \bullet V = 0$ and

$$V \;=\; S^\top (\bar{Y} - D^0) S.$$

Then, in $\bar{Y} - D^0$ only the first t blocks can have negative eigenvalues. Since the transformation matrix S respects the block structure, we have $\sigma_-^B(V) \le t$. ☐

Remark 4.2. In the special case in which all blocks have size 1, Theorem 4.1 can be stated as follows: In this case all matrices A_i, $i = 0, \dots, m$ are diagonal. Let $a^i = ((A_1)_{ii}, \dots, (A_m)_{ii})$, and let A be the matrix formed by the rows a^i. Then, $A_i \bullet X = A_0 \bullet X^0$ is equivalent to $Ax = Ax^0$. The condition states that for all $v \neq 0$ with $Av = 0$ we have $|\{i : v_i < 0\}| > t$ and $|\{i : v_i > 0\}| > t$, which is Theorem 1 in [15].

Example 4.3. Consider again the matrices A_0, A_1, A_2 from Example 3.6. Setting

$$X^0 = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{pmatrix}$$

yields $A_0 \bullet X^0 = -1$ and $A_1 \bullet X^0 = A_2 \bullet X^0 = 0$. In this case $\sigma_+^B(X^0) = 2$. For the corresponding system of equations to have X^0 as the unique solution, we would need $\sigma_+^B(V) > 2$ and $\sigma_-^B(V) > 2$ for all symmetric $V \neq 0$ with $A_0 \bullet V = A_1 \bullet V = A_2 \bullet V = 0$.

However,

$$V = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

satisfies the equality constraints, but has $\sigma_-^B(V) = 1$, which is in accordance with Example 3.6, in which two extreme point solutions arise.

Example 4.4. Let n be even and consider the $(n-1) \times n$ linear system of equations

$$Dv \coloneqq \begin{pmatrix} 1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & \cdots & 0 & 1 & 1 & 0 \\ 0 & \cdots & 0 & 0 & 1 & 1 \end{pmatrix} v = 0. \tag{4.1}$$

Then, form a symmetric matrix $V = \operatorname{diag}(v)$. Let A_i, $i = 1, \dots, n-1$, be appropriate symmetric $n \times n$ matrices such that

$$A_i \bullet V = 0, \quad i = 1, \dots, n-1,$$

is equivalent to $Dv = 0$. Without any loss of generality, these are block matrices with respect to the blocks $B_1 = \{1\}, \dots, B_n = \{n\}$. In the notation of Theorem 4.1, $m \coloneqq n-1$ and $V \neq 0$ with $A_i \bullet V = 0$ for all $i \in [m]$ is equivalent to $v \neq 0$ with $Dv = 0$.

Then, $v_1 = -v_2 = v_3 = -v_4 = \cdots = v_{n-1} = -v_n$. We can assume w.l.o.g. (by possible multiplication with -1) that $v_1 > 0$. Then, $v_1, v_3, \dots, v_{n-1}$ will be positive, while $v_2, v_4, \dots, v_n$ will be negative. Thus, any solution $V \neq 0$ to $A_i \bullet V = 0$, $i \in [m]$, satisfies $\sigma_+^B(V) = \sigma_-^B(V) = n/2$. By Theorem 4.1, the system $A_i \bullet X = A_i \bullet X^0$, $i \in [m]$, has the unique (symmetric) solution $X^0 \succeq 0$ if $\sigma_+^B(X^0) < n/2$. Note that the rank of the matrix D is $n-1$, which shows that the system has infinitely many solutions if X^0 is an arbitrary matrix.

Remark 4.5. Consider the condition on V in Theorem 4.1. The total number of blocks is at most n, and if a block contributes both to $\sigma_+^B(V)$ and $\sigma_-^B(V)$, the block has to have at least size 2. Therefore, $\sigma_+^B(V) + \sigma_-^B(V) \leq n$. This implies that the largest t for which $\sigma_+^B(V) > t$ and $\sigma_-^B(V) > t$ can hold is $\lfloor n/2 \rfloor - 1$. Example 4.4 shows that this bound is tight (if n is odd, one can ignore a single variable in v and use the construction on the remaining part). Note that for even n this bound can only be attained in the LP-case, i.e., if all matrices are diagonal.

Example 4.6. Let n be divisible by 3, define $k \coloneqq n/3$, and consider the 2×2 blocks $B_1 = \{1, 2\}, \dots, B_k = \{2k-1, 2k\}$. Take the same $(n-1) \times n$ linear system of equations as in Example 4.4 and fill in the variables of a solution v into the symmetric $2k \times 2k$ block matrix V as follows:

$$V = \begin{pmatrix} v_1 & v_3 & & & & & \\ v_3 & v_2 & & & & & \\ & & v_4 & v_6 & & & \\ & & v_6 & v_5 & & & \\ & & & & \ddots & & \\ & & & & & v_{3k-2} & v_{3k} \\ & & & & & v_{3k} & v_{3k-1} \end{pmatrix}.$$

Let A_i, $i = 1, \dots, m := n-1$, be symmetric $n \times n$ matrices such that

$$A_i \bullet V = 0, \quad i = 1, \dots, n-1$$

is equivalent to $Dv = 0$ from (4.1). We can assume that the A_i are block matrices for the above blocks. As in Example 4.4, assuming that $v_1 > 0$, the equations imply that $v_1 = v_3 = \cdots = v_{n-1} > 0$, while $v_2 = v_4 = \cdots = v_n < 0$. Thus, denoting $\lambda = v_1$, each 2×2 block has the following structure:

$$\begin{pmatrix} \lambda & \lambda \\ \lambda & -\lambda \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} -\lambda & -\lambda \\ -\lambda & \lambda \end{pmatrix}.$$

In both cases, the eigenvalues are $\pm\sqrt{2}\,\lambda$. Therefore, each block is counted both in $\sigma_+^B(V)$ and in $\sigma_-^B(V)$. Thus, any solution $V \neq 0$ to $A_i \bullet V = 0$, $i \in [m]$, will satisfy $\sigma_+^B(V) = \sigma_-^B(V) = k = n/3$. By Theorem 4.1, the system $A_i \bullet X = A_i \bullet X^0$, $i \in [m]$, has the unique (symmetric) solution $X^0 \succeq 0$ if $\sigma_+^B(X^0) < n/3$.

Remark 4.7. Note that the uniqueness conditions of Example 4.6 include matrices X^0 with negative entries, which would not be allowed in the LP-case (as in Example 4.4); for example, if $k-1$ blocks of X^0 consist of the positive definite matrices

$$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

and 0-blocks otherwise. This shows that while the size of t is possibly smaller than in the LP-case, the general spectrahedron case allows for a wider range of cases of X^0 in which uniqueness appears.

5. Perspectives and Open Problems

In Section 4, we have provided a criterion for particular semidefinite block systems to have a unique feasible solution. If this particular situation does not arise, it is an open question whether one can obtain an IIS by solving a single semidefinite program.

By Lemma 3.3 it would suffice to find solutions of minimal block support. For a matrix X the number of nonzero blocks can be written as

$$\|X\|_{2,0} := \big\|(\|X_{B_1}\|_2, \dots, \|X_{B_k}\|_2)\big\|_0,$$

where $\|x\|_0$ denotes the number of nonzeros in a vector x. Thus, it would suffice to solve the following problem to find an IIS:

$$\min \big\{ \|X\|_{2,0} \,:\, X \in S(\Sigma) \big\}. \tag{5.1}$$

Unfortunately, the $\|\cdot\|_{2,0}$ "norm" is nonconvex and thus hard to handle, for instance, (5.1) is NP-hard. However, for linear systems recent developments, see, e.g., [23, 13, 24], suggest to replace $\|X\|_{2,0}$ by

$$\|X\|_{2,1} \coloneqq \big\|(\|X_{B_1}\|_2, \dots, \|X_{B_k}\|_2)\big\|_1 = \sum_{i=1}^{k} \|X_{B_i}\|_2,$$

which leads to the following convex optimization problem:

$$\min \Big\{ \sum_{i=1}^{k} \|X_{B_i}\|_2 \,:\, X \in S(\Sigma) \Big\}. \tag{5.2}$$

Lemma 5.1. *Problem (5.2) can be formulated as SDP.*

Proof. Use the second order-cone condition

$$\Big\{(x,t) \,:\, \big(\sum x_i^2\big)^{1/2} \le t\Big\}$$

to represent $\|X_{B_i}\|_2 \le t_i$ with new variables t_i and minimize the objective function $\sum_{i=1}^{k} t_i$. It is well-known that second order conditions are special cases of semidefinite conditions (see, e.g., [9]). Since $X \in S(\Sigma)$ is already a positive semidefinite condition, this concludes the proof. □

An interesting line of future research would investigate conditions under which (5.2) provides an optimal solution for (5.1), which would try to generalize the above mentioned results from the linear to the block semidefinite case.

As pointed out by an anonymous referee, it would also be interesting to understand properties, under which an infeasible semidefinite system satisfies the converse of Theorem 3.5, as well as properties, under which every extreme point of the alternative spectrahedron has minimal block support.

Finally, it remains as a natural question to study the combination of our methods with exact duality theory versions for semidefinite programming, such as the reformulation technique of [25].

6. Conclusions

We have shown that one direction of the Gleeson-Ryan-Theorem for infeasible linear systems generalizes to infeasible block semidefinite systems, but the other direction does not. To overcome the situation to identify IISs, we have given a unique recovery characterization. Both the algorithmic question touched in Section 5 and the practical question of how to effectively exploit IISs of semidefinite systems within semidefinite integer programming solvers deserve further study.

Acknowledgment. We thank the anonymous referees for helpful suggestions.

References

[1] Schrijver, A.: Theory of Linear and Integer Programming. Wiley-Interscience Series in Discrete Mathematics. John Wiley & Sons, Ltd., Chichester (1986)

[2] Gleeson, J., Ryan, J.: Identifying minimally infeasible subsystems of inequalities. ORSA J. Comput. **2**(1), 61–63 (1990)

[3] Chinneck, J.W.: Finding a useful subset of constraints for analysis in an infeasible linear program. INFORMS J. Comput. **9**(2), 164–174 (1997)

[4] Chinneck, J.W., Dravnieks, E.W.: Locating minimal infeasible constraint sets in linear programs. ORSA J. Comput. **3**(2), 157–168 (1991)

[5] van Loon, J.N.M.: Irreducibly inconsistent systems of linear inequalities. Eur. J. Oper. Res. **8**(3), 283–288 (1981)

[6] Chinneck, J.W.: Feasibility and Infeasibility in Optimization: Algorithms and Computational Methods, *International Series in Operations Research and Management Sciences*, vol. 118. Springer (2008)

[7] Guieu, O., Chinneck, J.W.: Analyzing infeasible mixed-integer and integer linear programs. INFORMS J. Comput. **11**(1), 63–77 (1999)

[8] Codato, G., Fischetti, M.: Combinatorial Benders' cuts. In: D. Bienstock, G. Nemhauser (eds.) Proc. 10th International Conference on Integer Programming and Combinatorial Optimization (IPCO), New York, *LNCS*, vol. 3064, pp. 178–195. Springer-Verlag, Berlin (2004)

[9] Tunçel, L.: Polyhedral and Semidefinite Programming Methods in Combinatorial Optimization. Fields Institute Monographs. American Mathematical Society, Providence, RI (2010)

[10] Gally, T., Pfetsch, M.E., Ulbrich, S.: A framework for solving mixed-integer semidefinite programs. Optim. Methods and Software **33**(3), 594–632 (2018)

[11] Achterberg, T.: Conflict analysis in mixed integer programming. Discrete Opt. **4**(1), 4–20 (2007)

[12] Witzig, J., Berthold, T., Heinz, S.: Experiments with conflict analysis in mixed integer programming. In: D. Salvagnin, M. Lombardi (eds.) Integration of AI and OR Techniques in Constraint Programming, *LNCS*, vol. 10335, pp. 211–222. Springer-Verlag, Berlin (2017)

[13] Elhamifar, E., Vidal, R.: Block-sparse recovery via convex optimization. IEEE Trans. Signal Process. **60**(8), 4094–4107 (2012)

[14] Khajehnejad, M.A., Dimakis, A.G., Xu, W., Hassibi, B.: Sparse recovery of nonnegative signals with minimal expansion. IEEE Trans. Signal Process. **59**(1), 196–208 (2011)

[15] Wang, M., Tang, A.: Conditions for a unique non-negative solution to an underdetermined system. In: 47th Annual Allerton Conf. on Communication, Control, and Computing, Monticello IL. IEEE (2009)

[16] Wang, M., Xu, W., Tang, A.: A unique "nonnegative" solution to an underdetermined system: From vectors to matrices. IEEE Trans. Signal Process. **59**(3), 1007–1016 (2011)

[17] Klep, I., Schweighofer, S.: An exact duality theory for semidefinite programming based on sums of squares. Math. of Oper. Res. **38**(3), 569–590 (2013)

[18] Blekherman, G., Parrilo, P.A., Thomas, R.R.: Semidefinite Optimization and Convex Algebraic Geometry. SIAM, Philadelphia, PA (2013)

[19] Theobald, T.: Some recent developments in spectrahedral computation. In: G. Böckle, W. Decker, G. Malle (eds.) Algorithmic and Experimental Methods in Algebra, Geometry, and Number Theory, pp. 717–739. Springer (2017)

[20] Pfetsch, M.E.: The maximum feasible subsystem problem and vertex-facet incidences of polyhedra. Ph.D. thesis, TU Berlin (2003)

[21] Rockafellar, R.T.: Convex Analysis. Princeton University Press (1997)

[22] Amaldi, E., Pfetsch, M.E., Trotter, Jr., L.E.: On the maximum feasible subsystem problem, IISs, and IIS-hypergraphs. Math. Program. **95**(3), 533–554 (2003)

[23] Eldar, Y.C., Kuppinger, P., Bölcskei, H.: Block-sparse signals: Uncertainty relations and efficient recovery. IEEE Trans. Signal Process. **58**(6), 3042–3054 (2010)

[24] Lin, J.H., Li, S.: Block sparse recovery via mixed l_2/l_1 minimization. Acta Mathematica Sinica, English Series **29**(7), 1401–1412 (2013)

[25] Liu, M., Pataki, G.: Exact duality in semidefinite programming based on elementary reformulations. SIAM J. Optim. **25**(3), 1441–1454 (2015)

(Kai Kellner) Frankfurt am Main, Germany
E-mail address: kellner.kai@gmx-topmail.de

(Marc E. Pfetsch) Department of Mathematics, TU Darmstadt, Dolivostr. 15, 64293 Darmstadt, Germany
E-mail address: pfetsch@mathematik.tu-darmstadt.de

(Thorsten Theobald) Goethe-Universität, FB 12 – Institut für Mathematik, Postfach 11 19 32, 60054 Frankfurt am Main, Germany
E-mail address: theobald@math.uni-frankfurt.de