# Reduction methods in semidefinite and conic optimization

by

Frank Noble Permenter

BS Electrical Engineering, The University of Houston, 2006

MS Electrical Engineering, Stanford, 2010

Submitted to the Department of Electrical Engineering and Computer Science in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Electrical Engineering and Computer Science
at the Massachusetts Institute of Technology

September 2017



Signature of Author: ______________________________

Department of Electrical Engineering and Computer Science
August 11, 2017

Certified by: ______________________________

Pablo A. Parrilo
Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Certified by: ______________________________

Russ Tedrake
Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by: ______________________________

Leslie A. Kolodziejski
Professor of Electrical Engineering and Computer Science
Chair, Department Committee on Graduate Students

# Reduction methods in semidefinite and conic optimization

by Frank Noble Permenter
BS Electrical Engineering, The University of Houston, 2006
MS Electrical Engineering, Stanford, 2010

Submitted to the Department of Electrical Engineering
and Computer Science on August 11, 2017
in Partial Fulfillment of the Requirements for the Degree
of Doctor of Philosophy in Electrical Engineering and Computer Science

## Abstract

Conic optimization, or cone programming, is a subfield of convex optimization that includes linear, second-order cone, and semidefinite programming as special cases. While conic optimization problems arise in a diverse set of fields (including machine learning, robotics, and finance), efficiently solving them remains an active area of research. Developing methods that detect and exploit useful *structure*—such as symmetry, sparsity, or degeneracy—is one research topic. Such methods include facial and symmetry reduction, which have been successful in several applications, often reducing solve time by orders of magnitude. Nevertheless, theoretical and practical barriers preclude their general purpose use: to our knowledge, no solver uses facial or symmetry reduction as an automatic preprocessing step. This thesis addresses some of these barriers in three parts: the first develops more practical facial reduction techniques, the second proposes a more powerful and computationally efficient generalization of symmetry reduction (which we call *Jordan reduction*), and the third specializes techniques to convex relaxations of polynomial optimization problems. Throughout, we place emphasis on semidefinite programs and, more generally, optimization problems over symmetric cones. We also present computational results.

Thesis Supervisor: Pablo A. Parrilo
Professor of Electrical Engineering and Computer Science

Thesis Supervisor: Russ Tedrake
Professor of Electrical Engineering and Computer Science

# Acknowledgments

To Russ Tedrake and Pablo Parrilo, you two are a dream team of academic advisors. It would be any student's privilege to work with one of you; somehow, I got to work with you both. Thank you for the support. Thank you for the mentorship. Thank you most of all for the freedom to work on problems that interested me. To Rob Freund, thank you for serving on my committee. Your feedback and advice both improved this thesis and led to new research ideas. To coauthors Charles Wampler, Scott Kuindersma, Henrik Friberg, and Erling Andersen, thank you for the collaboration. Working with you made me a better writer and a better thinker. To my labmates—especially James Saunderson, Michael Posa, Anirudha Majumdar, Robin Deits, and Mark Tobenkin—thank you for being a sounding board. Our conversations helped me filter out the bad ideas and refine the good. Finally, to the optimization research community—especially Etienne de Klerk, Monique Laurent, and Gábor Pataki—thank you for the feedback, the talk invitations, and the interest in my work.

# Contents

# Preface

Conic optimization, or cone programming, is a subfield of convex optimization that includes linear, second-order cone, and semidefinite programming as special cases. While conic optimization problems arise in a diverse set of fields (including machine learning, robotics, and finance), efficiently solving them remains an active area of research. One research topic—the topic of this thesis—is detecting and exploiting useful *structure* such as symmetry, sparsity, or degeneracy. This makes problems easier to solve, often reducing solve time by orders of magnitude. Indeed, without structure exploitation, some problems may be impossible to solve due to memory and time constraints.

Towards giving a concrete example of structure, consider polynomial interpolation with nonnegativity constraints, i.e., consider the problem of finding a vector of coefficients $c \in \mathbb{R}^{d+1}$ such that the polynomial

$$f(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_d x^d$$

satisfies $f(x) \geq 0$ for all $x \in \mathbb{R}$ and $f(x_i) = y_i$ for a given set of points

$$(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N).$$

In the context of data fitting, nonnegativity of $f$ is desired if $y_i$ represents some nonnegative physical quantity (e.g., energy or power). This interpolation problem also arises in robot path planning [43], where nonnegativity of $f(x)$ represents a collision avoidance constraint and the data $(x_i, y_i)$ represents a set of way points.

It turns out that finding such a polynomial $f$ is a conic optimization problem over the *cone* of polynomials that are nonnegative. For this problem, useful structure includes the presence of symmetry or the presence of zeros in the data (Figure 1). In these cases, one can restrict the problem to a special subset of nonnegative polynomials in advance of solving, yielding a new problem of reduced dimension. Specifically, if the data is symmetric about the $y$-axis, one can restrict to even functions, i.e., to polynomials that satisfy $f(x) = f(-x)$. If the data contains a zero, i.e., a point $(x_i, y_i)$ with $y_i = 0$, one can restrict to nonnegative polynomials that have $(x - x_i)^2$ as a factor, i.e., to

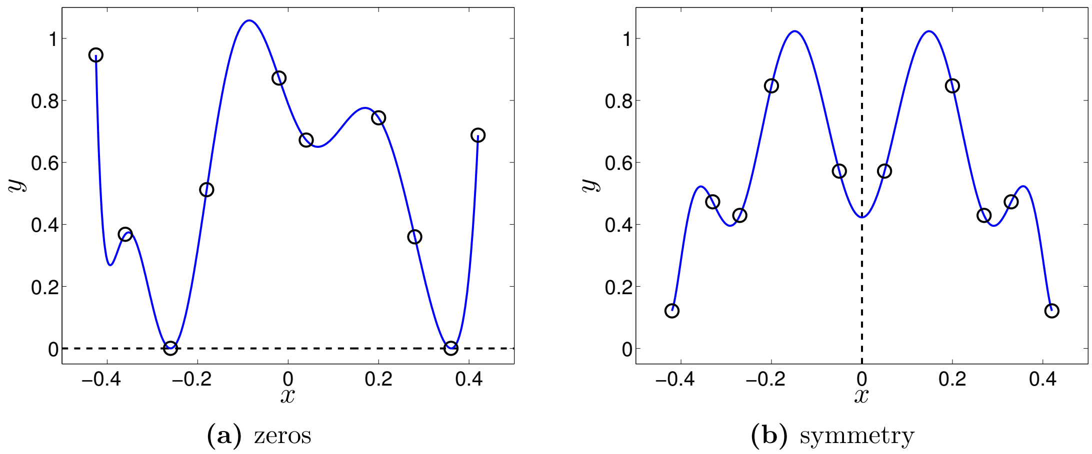
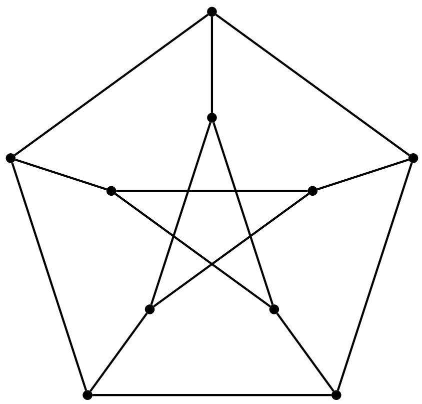



**Figure 1:** Useful structure in nonnegative polynomial interpolation, a conic optimization problem over the cone of nonnegative polynomials.

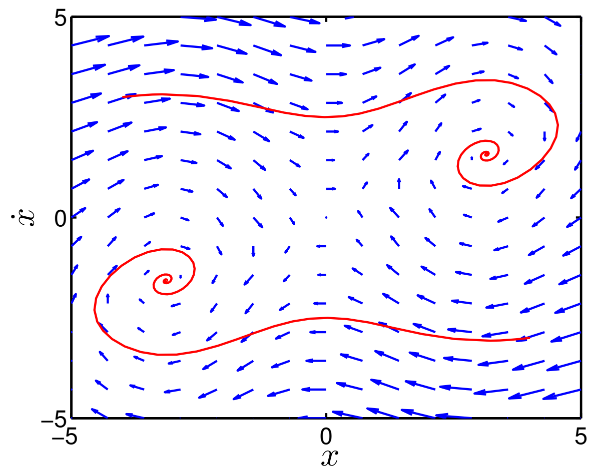
polynomials that equal $(x - x_i)^2 g(x)$ for some $g(x) \geq 0$.

Exploiting zeros and symmetry in nonnegative polynomial interpolation are special cases of two general methods: facial reduction [46, 20, 102] and symmetry reduction [59, 134, 37]. These methods simplify structured conic optimization problems arising in many areas, including graph theory [144, 39, 146, 23, 6], control of differential equations [141, 36, 34, 9], matrix completion [81, 47], distance geometry [2, 27], coding theory [82], and polynomial optimization [139, 106, 59]. For instance, they are fruitfully applied for problems involving graphs with nontrivial automorphisms (i.e., permutations of the node set that preserve edges and nonedges) and control of differential equations with symmetries or multiple equilibria (Figure 2). Despite their broad success, however, facial and symmetry reduction are not automated preprocessing steps taken by any publicly available solver, at least to our knowledge. In fact, both methods face significant theoretical and practical barriers precluding such automated use. This thesis attacks these barriers.

To explain in more detail, we first overview the high-level geometric picture of symmetry and facial reduction, which is simple despite considerable mathematical sophistication under the hood. We then summarize the benefits of these methods (which go beyond problem size reduction) and the challenges users of these methods face. We highlight certain limitations of current facial and symmetry reduction algorithms. Finally, we outline how chapters of this thesis address these challenges and remove these limitations.



**(a)** A graph with nontrivial automorphisms.




**(b)** Phase diagram of a symmetric differential equation with multiple equilibria.

**Figure 2:** Useful structure for conic optimization problems arising in graph theory and control of differential equations.

## Symmetry and facial reduction

To explain the geometric picture behind symmetry and facial reduction, we must first fix the form of our optimization problem. Therefore, consider the following *cone program* in decision variable $x \in \mathbb{R}^n$

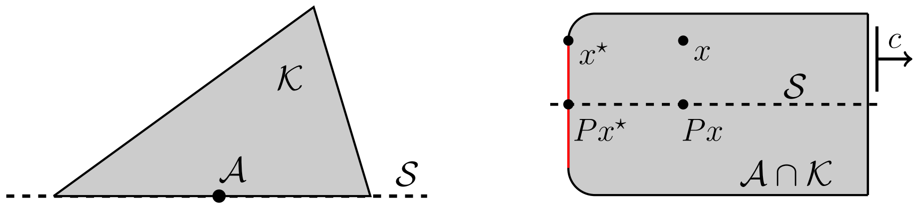
$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b, \\ & x \in \mathcal{K}, \end{array} \tag{1}$$

where $Ax = b$ are linear equations, $c \in \mathbb{R}^n$ is a fixed cost vector, and $\mathcal{K} \subseteq \mathbb{R}^n$ is a special type of set called a *convex cone*. Such cones include the nonnegative orthant, the set of positive semidefinite matrices and, as just mentioned, the set of degree $d$ polynomials that are nonnegative:

$$\{c \in \mathbb{R}^{d+1} : c_0 + \sum_{i=1}^{d} c_i x^i \geq 0 \;\; \forall x \in \mathbb{R}\}.$$

Given a cone program, facial and symmetry reduction perform two steps:

1. find a subspace $\mathcal{S} \subseteq \mathbb{R}^n$ containing solutions;




**(a)** Facial reduction **(b)** Symmetry reduction

**Figure 3:** Facial reduction finds a subspace $\mathcal{S}$ containing $\mathcal{A} := \{x \in \mathbb{R}^n : Ax = b\}$ that intersects the cone $\mathcal{K}$ on the boundary. Symmetry reduction takes $\mathcal{S}$ equal to the range of a projection matrix $P \in \mathbb{R}^{n \times n}$ that maps each feasible point $x$ (resp., optimal point $x^\star$) to a feasible point $Px$ (resp., optimal point $Px^\star$).

2. write $\mathcal{K} \cap \mathcal{S}$ as a linear transformation of a simpler cone $\mathcal{C} \subseteq \mathbb{R}^m$

$$\mathcal{K} \cap \mathcal{S} = \{\Phi z : z \in \mathcal{C}\},$$

where $m < n$ and $\Phi \in \mathbb{R}^{n \times m}$. (We will make "simpler" precise in later chapters.)

Completing these steps allows one to instead solve

$$\begin{array}{ll} \text{minimize} & (\Phi^T c)^T z \\ \text{subject to} & A\Phi z = b, \\ & z \in \mathcal{C}, \end{array}$$

a cone program in decision variable $z \in \mathbb{R}^m$ formulated over a lower dimensional vector space $\mathbb{R}^m$. Further, from a solution $z^\star$ of this cone program, one obtains a solution $x^\star$ to the original (1) simply by taking $x^\star = \Phi z^\star$.

Facial and symmetry reduction differ in how they identify $\mathcal{S}$. Figure 3 illustrates the different techniques. Here, we see that facial reduction finds $\mathcal{S}$ by exploiting a certain form of geometric degeneracy. Specifically, it exploits an empty intersection between $\mathcal{A} := \{x \in \mathbb{R}^n : Ax = b\}$ and the interior of the cone $\mathcal{K}$—a so-called failure of *Slater's condition*. Symmetry reduction, on the other hand, takes $\mathcal{S}$ equal to the range of a projection matrix $P \in \mathbb{R}^{n \times n}$ that maps feasible points to feasible points and optimal points to optimal points. It constructs $P$ from joint symmetries of the constraint sets $\mathcal{A}$ and $\mathcal{K}$ and the objective function $c^T x$.

Both facial and symmetry reduction write $\mathcal{K} \cap \mathcal{S}$ as a linear transformation of the cone $\mathcal{C}$ using special structure of $\mathcal{S}$ and $\mathcal{K}$. In facial reduction, the intersection of $\mathcal{S}$ with $\mathcal{K}$ is a special subset of $\mathcal{K}$ called a *face*. (Hence, the name facial reduction.) In symmetry reduction, $\mathcal{S}$ is the *fixed-point subspace* of a *group action*, a mathematical object that formalizes notions of symmetry.

### Example: Non-negative interpolating polynomials

For the interested reader, we concretely illustrate the steps taken by facial and symmetry reduction on the aforementioned nonnegative polynomial interpolation problem. This (somewhat lengthy) example can be skipped, as no other material depends on it.

**Interpolation via cone programming** We first illustrate that nonnegative polynomial interpolation is a cone program. To begin, recall that a polynomial $f(x)$ interpolates a set

$$D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\} \subset \mathbb{R}^2$$

if $f(x_i) = y_i$ for each $(x_i, y_i) \in D$ and that a polynomial is nonnegative if $f(x) \geq 0$ for all $x \in \mathbb{R}$. For a fixed-degree bound $d$, finding a nonnegative interpolating polynomial is a cone program over coefficient vectors. To see this, let $f_u(x)$ denote the univariate polynomial with coefficient vector $u \in \mathbb{R}^{d+1}$, i.e., let

$$f_u(x) := u_0 + u_1 x + u_2 x^2 + \cdots + u_d x^d.$$

Finally, let $\mathcal{K}_d \subseteq \mathbb{R}^{d+1}$ and $\mathcal{A}_d \subseteq \mathbb{R}^{d+1}$ denote the polynomials satisfying the nonnegativity and interpolation constraints, respectively:

$$\mathcal{K}_d := \{u \in \mathbb{R}^{d+1} : f_u(x) \geq 0 \;\; \forall x \in \mathbb{R}\}, \;\; \mathcal{A}_d := \left\{u \in \mathbb{R}^{d+1} : f_u(x_i) = y_i \;\; \forall (x_i, y_i) \in D\right\}.$$

Clearly $f_u(x)$ is a nonnegative polynomial that interpolates $D$ if and only if $u \in \mathcal{A}_d \cap \mathcal{K}_d$. Hence, we obtain a cone program in variable $u \in \mathbb{R}^{d+1}$

$$\begin{array}{ll} \text{minimize} & c^T u \\ \text{subject to} & u \in \mathcal{A}_d \cap \mathcal{K}_d, \end{array} \tag{2}$$

where the objective $c^T u$ is any linear function of the coefficient vector $u$, e.g., the value of $f_u(x)$ or one of its derivatives at some distinguished point (say, $x = 0$).

**Exploiting zeros (facial reduction)** Suppose the data set $D$ contains a real zero, i.e., suppose $(a, 0) \in D$ for some $a \in \mathbb{R}$. Then the subspace (indeed, hyperplane)

$$\mathcal{S} = \left\{u \in \mathbb{R}^{d+1} : f_u(a) = 0\right\}$$

contains the set of interpolating polynomials $\mathcal{A}_d$. Further, any nonnegative polynomial in $\mathcal{S}$ necessarily factors as $(x - a)^2 h(x)$ since the real roots of nonnegative polynomials are repeated. In addition, the factor $h(x)$ must be nonnegative by continuity. It follows that

$$\mathcal{K}_d \cap \mathcal{S} = \left\{u \in \mathbb{R}^{d+1} : f_u(x) = (x - a)^2 f_v(x), \text{ for } v \in \mathcal{K}_{d-2}\right\},$$

i.e., $\mathcal{K}_d \cap \mathcal{S}$ is the cone $\mathcal{K}_{d-2}$ transformed by multiplication with the fixed polynomial $(x-a)^2$. Note that this is a linear transformation of coefficients. For $d = 4$, we have explicitly that

$$\mathcal{K}_4 \cap \mathcal{S} = \{\Phi v : v \in \mathcal{K}_2\}, \qquad \Phi = \begin{bmatrix} a^2 & 0 & 0 \\ -2a & a^2 & 0 \\ 1 & -2a & a^2 \\ 0 & 1 & -2a \\ 0 & 0 & 1 \end{bmatrix}, \tag{3}$$

where $(1, x, x^2)$ labels the columns of $\Phi$ and $(1, x, x^2, x^3, x^4)$ labels the rows.

**Exploiting symmetry (symmetry reduction)** Suppose that the data $D$ is symmetric about the $y$-axis, i.e., suppose that $(-x_i, y_i) \in D$ whenever $(x_i, y_i) \in D$. Under this assumption, existence of nonnegative interpolating polynomials implies existence of such polynomials that are also even functions. In other words, if $\mathcal{A}_d \cap \mathcal{K}_d$ is nonempty, so is $\mathcal{A}_d \cap \mathcal{K}_d \cap \mathcal{S}$, where

$$\mathcal{S} := \{u \in \mathbb{R}^{d+1} : f_u(x) = f_u(-x)\}.$$

To see this, note that $f_u(x)$ interpolates $D$ if and only if its reflection $f_u(-x)$ interpolates $D$. Further, any convex combination of these polynomials interpolates $D$, including the even polynomial

$$f_u^{even}(x) := \frac{1}{2}\left(f_u(-x) + f_u(x)\right).$$

Since $f_u^{even}(x) \geq 0$ when $f_u(x) \geq 0$, the claim follows. As an aside, note the linear map from $f_u$ to $f_u^{even}$ is idempotent, i.e., $(f_u^{even})^{even} = f_u^{even}$; hence, it is a projection map.

Under an additional assumption on the cost vector $c$, the subspace $\mathcal{S}$ will intersect the set of optimal solutions. Specifically, let $u_{even}$ denote the coefficients of $f_u^{even}$ and suppose that $c^T u_{even} = c^T u$ for all $u$. Then clearly

$$\inf\{c^T u : u \in \mathcal{A}_d \cap \mathcal{K}_d\} = \inf\{c^T u : u \in \mathcal{A}_d \cap \mathcal{K}_d \cap \mathcal{S}\}$$

since for all $u \in \mathcal{A}_d \cap \mathcal{K}_d$, the point $u_{even} \in \mathcal{A}_d \cap \mathcal{K}_d \cap \mathcal{S}$ has equal cost.

Finally, $\mathcal{K}_d \cap \mathcal{S}$ is a linear transformation of $\hat{\mathcal{K}}_{d/2} := \{u \in \mathbb{R}^{d/2+1} : f_u(x) \geq 0 \ \ \forall x \geq 0\}$—the cone of degree $d/2$ polynomials nonnegative on the nonnegative real line. Precisely,

$$\mathcal{K}_d \cap \mathcal{S} = \{u \in \mathbb{R}^{d+1} : f_u(x) = f_v(x^2), \text{ for some } v \in \hat{\mathcal{K}}_{d/2}\},$$

a linear transformation of $\hat{\mathcal{K}}_{d/2}$ induced by $(1, x, x^2, \ldots, x^{d/2}) \mapsto (1, x^2, x^4, \ldots, x^d)$. For

$d = 4$, we have explicitly that

$$\mathcal{K}_4 \cap \mathcal{S} = \left\{ \Phi v : v \in \hat{\mathcal{K}}_2 \right\}, \qquad \Phi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

where $(1, x, x^2)$ labels the columns of $\Phi$ and $(1, x, x^2, x^3, x^4)$ labels the rows.

## Benefits of reduction

**Problem size reduction** The obvious benefit of facial and symmetry reduction is the transformation of a given problem into a smaller one. This can reduce total solve time and memory requirements of solution algorithms. Problem size reduction is also theoretically useful. For instance, it may provide analytic solutions, illuminate the asymptotic behavior of a problem family (e.g., [117]), or clarify the relationship between different formulations (e.g., [123]).

**Improved conditioning** Facial and symmetry reduction also improve accuracy. One reason for this is simple: smaller problems can lead to more accurate floating point computation. Another reason is less obvious: both facial and symmetry reduction can improve the intrinsic conditioning of a cone program. Specifically, they can lower its *singularity degree* [130], a quantity that bounds the difference between forward error (distance to solutions) and backward error (constraint violation). Note this latter error is what a solver can easily compute, whereas the former—which may be hard or impossible to compute—is the actual measure of solution quality.

**Pathology removal (facial reduction)** In general, cone programs can suffer a pathology: certain sufficient conditions for optimality, unboundedness, and infeasibility can be simultaneously unsatisfiable. Almost all solvers directly attempt to satisfy one of these conditions and will fail in this case. Facial reduction can remove this pathology and hence this source of failure. Interestingly, symmetry reduction doesn't remove pathologies for reasons explained in Chapter 1.5.5.

## Challenges of reduction

**Cost** Finding a subspace $\mathcal{S} \subseteq \mathbb{R}^n$ that intersects the set of solutions may carry some computational cost, as may finding a linear transformation $\Phi \in \mathbb{R}^{n \times m}$ and cone $\mathcal{C} \subseteq \mathbb{R}^m$ satisfying

$$\mathcal{K} \cap \mathcal{S} = \{\Phi z : z \in \mathcal{C}\}.$$

Hence, one must trade off these computational costs with the aforementioned benefits they afford. If this trade-off is not managed appropriately, facial and symmetry reduction can actually *increase* total solve time.

**Sparsity** Reformulating a cone program using facial or symmetry reduction can destroy problem sparsity. Indeed, if $\mathcal{K} \cap \mathcal{S} = \{\Phi z : z \in \mathcal{C}\}$, the reformulation takes the form

$$\begin{array}{ll} \text{minimize} & (\Phi^T c)^T z \\ \text{subject to} & A\Phi z = b, \\ & z \in \mathcal{C}, \end{array}$$

where $(A, b, c)$ is the original problem data. Note that if the matrix $A$ and vector $c$ are sparse, the compositions $A\Phi$ and $(\Phi^T c)$ can be dense depending on $\Phi$, increasing the total number of nonzeros in the problem data. In practice, this becomes an issue when facial and symmetry reduction achieve only a moderate decrease in dimension (i.e., the matrix $\Phi$ is nearly square). Consider, for instance, the matrix $\Phi$ from (3) in the polynomial interpolation example:

$$\Phi = \begin{bmatrix} a^2 & 0 & 0 \\ -2a & a^2 & 0 \\ 1 & -2a & a^2 \\ 0 & 1 & -2a \\ 0 & 0 & 1 \end{bmatrix}.$$

Composing the following sparse matrix $A$ with $\Phi$ yields a dense matrix $A\Phi$

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix}, \qquad A\Phi = \begin{bmatrix} 1-2a & a^2 - 2\,a & a^2 \\ 1 & 1-2a & a^2-2a \end{bmatrix}.$$

While this example is too small to be compelling, it illustrates the issue, which can be significant in practice.

**Sensitivity** Due to floating-point error, the identified subspace $\mathcal{S}$ may only *approximately* contain solutions. As a consequence, solving a reformulation over this subspace may not solve the given cone program. This issue is particularly salient when the given problem is *ill posed*, meaning its *optimal value* is infinitely sensitive to perturbations of its problem data. Unfortunately, facial reduction finds reformulations *only if* the problem is ill posed.

**Dual recovery (facial reduction)** A solution $x$ of the given cone program is easily obtained from a solution $z$ of the reformulation: we simply take $x = \Phi z$. If one uses facial reduction to construct the reformulation, the same is *not* true for the dual problem. Unfortunately, one cannot ignore this issue and incorporate facial reduction into a primal-dual solver. Such solvers must return solutions to a given cone program and its dual; indeed, the dual may be the problem that is of actual interest to a user. Interestingly, this is not an issue for symmetry reduction for reasons discussed in Chapter 1.5.5.

## Algorithmic limitations

In additional to the aforementioned challenges (which are mostly practical), symmetry reduction and facial reduction also have certain algorithmic limitations.

### Symmetry reduction

There actually is no known algorithm for finding a subspace $\mathcal{S}$ of minimum dimension within the symmetry reduction framework. Existing algorithms may find larger subspaces or are completely tailored to specific problem families. For semidefinite programs, the same is true for the related **-algebra* reduction framework (e.g., [37]).

### Facial reduction

Unlike symmetry reduction, there *are* algorithms [20, 102, 138] for finding a subspace of minimum dimension within the facial reduction framework. (This subspace is the linear span of the so-called *minimal face*.) Nevertheless, facial reduction has an algorithmic limitation of a different flavor. To explain, note that facial reduction serves two main and distinct purposes: to reduce the dimension of a problem *and* to remove pathologies if they exist. There is no algorithm, however, that does facial reduction *only if* a given program is pathological. Such an algorithm would avoid the challenges of facial reduction unless the given problem is otherwise unsolvable.

## Outline

This thesis addresses the aforementioned challenges and limitations of symmetry and facial reduction. In addition to the following outline, we also include tables summarizing key results. Table 1 indicates the challenges addressed by each chapter. Tables 2-3 summarize the features of a new facial reduction algorithm (developed in Chapter 4) and a new reduction methodology (called Jordan reduction) that address the aforementioned algorithmic limitations.

A detailed outline now follows. The opening chapter provides technical background, and the remaining chapters present original research in three parts. Part I concerns facial reduction, Part II concerns Jordan reduction, and Part III specializes techniques to polynomial optimization.

| **method** | **cost** | **sparsity** | **accuracy** | **dual soln. recovery** |
|---|---|---|---|---|
| Facial red. | Ch. 2, 4 | Ch. 2 | Ch. 2 | Ch. 3, 4 |
| Jordan red. | Ch. 7 | Ch. 6, 7 | Ch. 7 | N/A |

**Table 1:** Challenges of reduction considered by indicated chapter.

| **Facial red. algorithm** | **termination criterion** | **output if no reduction** | **dual solution recovery** |
|---|---|---|---|
| [20, 102, 138] | Slater's condition holds | nothing | sometimes possible |
| via self-dual embedding (Ch. 4) | instance not pathological | optimal primal-dual solutions | always possible if duality gap is zero |

**Table 2:** Comparison of facial reduction algorithms when applied to feasible problems.

| **method** | **type of subspace $\mathcal{S}$** | **alg. for minimal $\mathcal{S}$** | **removes pathologies** | **dual recovery** |
|---|---|---|---|---|
| Facial reduction | hyperplane exposing face | yes | yes | sometimes possible |
| Symmetry reduction | range of projection from groups | no | no | always possible |
| Jordan red. (Ch. 5) | range of projection | yes | no | always possible |

**Table 3:** Type of subspace found by method; existence of an algorithm for finding a minimal subspace within framework of method; whether or not pathological instances remain pathological if method is used; if recovery of solutions to original dual program is possible.

## Chapter 1: Background

This chapter provides background material for the ensuing chapters. It overviews cone programming, facial reduction, and symmetry reduction in full detail. It also reviews Euclidean Jordan algebra theory, which provides the machinery needed for Part II.

## Part I: Facial reduction approaches

Part I addresses the challenges encountered in facial reduction which, as mentioned, relate to cost, sparsity, sensitivity, and dual solution recovery. We also give a new facial reduction algorithm tailored only to pathological instances, correcting the mentioned limitation of current facial reduction algorithms.

## Chapter 2: Partial facial reduction

We show that introducing a user-specified approximation of the cone $\mathcal{K}$ overcomes the cost, accuracy, and sparsity issues of facial reduction. This allows one to trade-off the cost of facial reduction with its benefits. For polyhedral approximations, it allows one to find faces in exact arithmetic. For a particular type of approximation, it allows one to find provably sparse reformulations. While the use of approximations can decrease the power of facial reduction, several examples illustrate practical effectiveness of the technique. Results of this chapter appear in [105].

## Chapter 3: Dual solution recovery

Suppose one solves a cone program with a primal-dual solver after facial reduction. In this chapter, we give a simple algorithm for recovering solutions to the original dual. Note that simple recovery is sometimes impossible, e.g., when facial reduction changes the dual optimal value. Hence, this algorithm necessarily fails in some cases. We give sufficient conditions for successful recovery. Results of this chapter appear in [105].

## Chapter 4: Self-dual embeddings and facial reduction

In this chapter, we develop a procedure that does facial reduction only if a given instance is pathological. This represents an extreme of the cost-benefit trade-off: the costs of facial reduction are only paid when removing pathologies is actually needed to solve the given problem. We build this procedure upon the *self-dual embedding*, the basis of widely-used solvers. As shown, extremely minor changes to solvers like SeDuMi [128] or MOSEK [94] implement the procedure and only modify solver execution on pathological instances. Nevertheless, provable correctness of such implementations relies on an assumption rarely met in practice due to floating-point error and asymptotic convergence of numerical algorithms. (Specifically, we assume that the solver tracks the central path precisely to its limit point.) Numerical experiments illustrate that failure of this assumption can severely degrade performance. Results from this chapter appear in [109] and are partially duplicated in the thesis of co-author Henrik Friberg [57].

## Part II: Jordan reduction

The next part of this thesis develops Jordan reduction, a generalization of symmetry reduction that is applicable to any optimization problem formulated over a *symmetric cone*. Problems of this type include linear programs, second-order cone programs and semidefinite programs.

### Chapter 5: Minimal subspaces in Jordan reduction

This chapter contains methods for finding subspaces within the Jordan reduction framework. Specifically, we show if an orthogonal projection satisfies key invariance properties, one can reformulate a given cone program over its range without changing primal or dual optimal values. We also show the range intersected with the cone is isomorphic to a lower-dimensional symmetric cone—namely, the cone-of-squares of a Euclidean Jordan algebra. We then give a simple algorithm for minimizing the rank of this projection (in polynomial time) and hence the dimension of the identified subspace $\mathcal{S}$. Finally, we prove minimizing the dimension of $\mathcal{S}$ also optimizes a decomposition of $\mathcal{K} \cap \mathcal{S}$ into irreducible symmetric cones. Some results of this chapter appear in [108].

### Chapter 6: Constructing isomorphisms between Euclidean Jordan algebras

This chapter provides an algorithm for writing $\mathcal{K} \cap \mathcal{S}$ as a linear transformation of a “simpler” cone $\mathcal{C}$ within the Jordan reduction framework, i.e., when $\mathcal{S}$ is a subspace found using techniques from Chapter 5 (and also Chapter 7). Indeed, this chapter solves a more general problem: constructing isomorphisms between two Euclidean Jordan algebras given a basis for each algebra. Under natural conditions, the constructed isomorphism leads to sparse reformulations over $\mathcal{C}$.

### Chapter 7: Combinatorial variations and computational results

This chapter gives combinatorial versions of Chapter 5 techniques that only find subspaces with appealing sparsity properties. Specifically, we restrict to subspaces that have bases with efficient representations based on transitive relations and partitions. This leads to algorithms based on partition refinement and transitive closures. As we show, the structure of these bases also allows the Chapter 6 algorithm to find a sparse transformation $\Phi$ between $\mathcal{K}\cap\mathcal{S}$ and the cone $\mathcal{C}$. Results of this chapter appear in [108].

## Part III: Applications to polynomial optimization

### Chapter 8: Reduction of sum-of-squares programs

An important class of semidefinite programs (SDPs) solve so-called sum-of-squares relaxations of polynomial optimization problems [15]. In this final chapter, we show methods from Part I and Part II generalize and improve existing algorithms for simplifying these SDPs. Some results of this chapter appear in [107, 106].

Chapter 1

# Background

We provide background material shared by multiple chapters, splitting it into six major sections. The first section gives preliminaries results. The second overviews cone programming. The third reviews facial reduction. The fourth reviews symmetry reduction and the fifth related *-algebra techniques. Finally, the sixth overviews Euclidean Jordan algebra theory. Some chapters only depend on a subset of this background material as indicated in Table 1.1.

| Chapters 2-4 | Sections 1.1-1.3 |
|---|---|
| Chapters 5-7 | Sections 1.1-1.2, 1.4-1.6 |

**Table 1.1:** Background material for indicated chapters

## ■ 1.1 Preliminaries

### ■ 1.1.1 Inner product spaces

We let $\mathcal{V}$ denote a finite-dimensional vector space over $\mathbb{R}$ equipped with an inner product $\langle\cdot,\cdot\rangle : \mathcal{V}\times\mathcal{V}\to\mathbb{R}$. For instance, $\mathcal{V}$ could denote $\mathbb{R}^n$ equipped with the dot product $x^T y$. It could also denote $\mathbb{S}^n$, the vector space of $n\times n$ symmetric matrices equipped with the trace inner product $\operatorname{Tr} XY$. Note that every linear functional $\ell : \mathcal{V}\to\mathbb{R}$ equals $x\mapsto\langle s_\ell, x\rangle$ for some fixed $s_\ell\in\mathcal{V}$. Hence, the vector space of all linear functionals, i.e., the *dual space* $\mathcal{V}^*$, can be identified with $\mathcal{V}$.

### ■ 1.1.2 Affine sets

A set $\mathcal{A}\subseteq\mathcal{V}$ is *affine* if it contains each line that passes between two points in $\mathcal{A}$, i.e., if for all $x, y\in\mathcal{A}$,

$$\mathcal{A}\supseteq\{tx+(1-t)y : t\in\mathbb{R}\}.$$

Any affine set equals the solution set of linear equations. Further, any nonempty affine set equals a linear subspace shifted by a point, i.e., if $\mathcal{A}$ is nonempty and affine, then

$\mathcal{A} = \{x_0 + z : z \in \mathcal{L}\}$ for some $x_0 \in \mathcal{V}$ and linear subspace $\mathcal{L} \subseteq \mathcal{V}$. Conversely, solution sets of linear equations and shifted linear subspaces $x_0 + \mathcal{L}$ are always affine.

## 1.1.3 Convex cones

A set $\mathcal{C} \subseteq \mathcal{V}$ is *convex* if it contains each line segment that connects two points in $\mathcal{C}$, i.e., if for all $x, y \in \mathcal{C}$,

$$\mathcal{C} \supseteq \{tx + (1-t)y : \quad t \in [0,1]\}\,.$$

A set $\mathcal{K} \subseteq \mathcal{V}$ is a *convex cone* if it is convex and closed under positive scaling, i.e., if for all $x \in \mathcal{K}$,

$$\mathcal{K} \supseteq \{\lambda x : \lambda > 0\}\,.$$

Examples of convex cones include the nonnegative orthant $\mathbb{R}^n_+$, the Lorentz cone $\mathcal{Q}^{n+1}$, and the cone of $n \times n$ positive semidefinite (psd) matrices $\mathbb{S}^n_+$:

- $\mathbb{R}^n_+ := \{x \in \mathbb{R}^n : x_i \geq 0\}$,
- $\mathcal{Q}^{n+1} := \{(x_0, x) \in \mathbb{R} \times \mathbb{R}^n : x_0 \geq \|x\|_2\}$,
- $\mathbb{S}^n_+ := \left\{X \in \mathbb{S}^n : z^T X z \geq 0 \ \ \forall z \in \mathbb{R}^n\right\}$.

The *dual cone* $\mathcal{K}^*$ of a convex cone $\mathcal{K}$ is the set $\{s \in \mathcal{V} : \langle s, x\rangle \geq 0, \quad \forall x \in \mathcal{K}\}$. As the name suggests, $\mathcal{K}^*$ is a convex cone. If $\mathcal{K}$ is closed, then $(\mathcal{K}^*)^* = \mathcal{K}$. Any linear subspace $\mathcal{L}$ is also a convex cone with dual cone equal to its *orthogonal complement* $\mathcal{L}^\perp := \{s \in \mathcal{V} : \langle s, x\rangle = 0, \ \ \forall x \in \mathcal{L}\}$. Finally, a cone is *self dual* if $\mathcal{K} = \mathcal{K}^*$. The cones $\mathbb{R}^n_+$, $\mathcal{Q}^{n+1}$ and $\mathbb{S}^n_+$ are all self dual.

## 1.1.4 Symmetric cones

A convex cone $\mathcal{K}$ is *homogeneous* if for all pairs $(x, y)$ in the interior of $\mathcal{K} \times \mathcal{K}$ there exists an invertible linear map $T : \mathcal{V} \to \mathcal{V}$, depending on $(x, y)$, with the following two properties:

$$Tx = y, \qquad \{Tx : x \in \mathcal{K}\} = \mathcal{K}.$$

A cone is *symmetric* if it is self dual and homogeneous. Note that the Cartesian product $\mathcal{K}_1 \times \mathcal{K}_2$ of two symmetric cones $\mathcal{K}_1$ and $\mathcal{K}_2$ is symmetric. If a symmetric cone doesn't equal a Cartesian product of two symmetric cones (up-to invertible linear transformation), it is *irreducible*.

Though the class of symmetric cones may seem esoteric, it includes the nonnegative orthant $\mathbb{R}^n_+$, the Lorentz cone $\mathcal{Q}^{n+1}$, and the psd cone $\mathbb{S}^n_+$. Both the Lorentz cone $\mathcal{Q}^{n+1}$ and the psd cone $\mathbb{S}^n_+$ are irreducible. For $n \geq 2$, the nonnegative orthant $\mathbb{R}^n_+$ is not irreducible since it equals the Cartesian product $\mathbb{R}^1_+ \times \cdots \times \mathbb{R}^1_+$. Symmetric cones have appealing computational properties. For instance, one can check membership in any symmetric cone using linear algebra. One can also solve optimization problems formulated over symmetric cones efficiently. Indeed, there is a unified theory for solving these optimization problems that is expressed in the language of *Euclidean Jordan algebras* [122, 54, 3]. We review these algebras in Section 1.6.

## 1.2 Cone programs

A cone program is the problem of minimizing a linear function over the intersection of an affine set with a convex cone $\mathcal{K} \subseteq \mathcal{V}$. If we let $x_0 + \mathcal{L}$ denote the affine set $\{x_0 + z : z \in \mathcal{L}\}$, where $\mathcal{L} \subseteq \mathcal{V}$ is a linear subspace and $x_0 \in \mathcal{V}$, a cone program takes the form

$$\mathbf{P}: \qquad \begin{array}{ll} \text{minimize} & \langle c, x\rangle \\ \text{subject to} & x \in (x_0 + \mathcal{L}) \cap \mathcal{K}, \end{array} \tag{1.1}$$

where $c \in \mathcal{V}$ is a fixed *cost vector* and $x \in \mathcal{V}$ is the *decision variable*. A point $x \in \mathcal{V}$ is *feasible* if $x \in (x_0 + \mathcal{L}) \cap \mathcal{K}$. A feasible point $x$ is *optimal* if it attains the *optimal value* $\theta$, i.e., if $\langle c, x\rangle = \theta$, where $\theta \in \mathbb{R}^n \cup \{\pm\infty\}$ denotes the infimum

$$\theta := \inf \{\langle c, x\rangle : x \in (x_0 + \mathcal{L}) \cap \mathcal{K}\}.$$

If the feasible set $(x_0 + \mathcal{L}) \cap \mathcal{K}$ is empty ($\theta = +\infty$), then one calls $\mathbf{P}$ *infeasible*. Similarly, if $\theta = -\infty$, then one calls $\mathbf{P}$ *unbounded*. Note that for some cone programs, no point attains the optimal value $\theta$ even when it is finite.

### 1.2.1 Sufficient conditions for optimality, infeasibility, and unboundedness

How does one show that a feasible point is optimal? Similarly, how does one show that a cone program is infeasible or unbounded? For cone programs, succinct *sufficient* conditions for optimality, infeasibility, and unboundedness exist. They are also necessary if $\mathcal{K}$ is polyhedral, or if certain regularity conditions hold (see, e.g., Section 1.3.3). We state them next in terms of the orthogonal complement $\mathcal{L}^\perp \subseteq \mathcal{V}$ and the dual cone $\mathcal{K}^* \subseteq \mathcal{V}$.

#### Optimality

A point $x \in \mathcal{V}$ is optimal if for some *dual variable* $s \in \mathcal{V}$

$$x \in (x_0 + \mathcal{L}) \cap \mathcal{K}, \qquad s \in (c + \mathcal{L}^\perp) \cap \mathcal{K}^*, \qquad \langle s, x\rangle = 0. \tag{1.2}$$

One calls the first condition *primal feasibility*, the second condition *dual feasibility* and the third condition *complementary slackness*.

#### Infeasibility

The cone program **P** is infeasible if there is a hyperplane $\{x \in \mathcal{V} : \langle s, x\rangle = 0\}$ strictly separating $\mathcal{K}$ from $x_0 + \mathcal{L}$, or, equivalently, if there is $s \in \mathcal{V}$ satisfying

$$\langle x_0, s\rangle < 0, \qquad s \in \mathcal{L}^{\perp}, \qquad s \in \mathcal{K}^*. \tag{1.3}$$

By the first two conditions, $\langle s, x\rangle < 0$ for all $x \in x_0 + \mathcal{L}$ and by the third $\langle s, x\rangle \geq 0$ for all $x \in \mathcal{K}$, proving $(x_0 + \mathcal{L}) \cap \mathcal{K}$ is empty.

#### Unboundedness

The cone program **P** is unbounded if it is feasible and has an *improving ray* $x_r \in \mathcal{V}$, i.e., a point $x_r$ satisfying

$$\langle c, x_r\rangle < 0, \qquad x_r \in \mathcal{L}, \qquad x_r \in \mathcal{K}. \tag{1.4}$$

Indeed, for any feasible $\hat{x}$ and $\alpha \geq 0$, the point $\hat{x}+\alpha x_r$ is feasible and has cost $\langle c, x+\alpha x_r\rangle$ that tends to $-\infty$ as $\alpha$ tends to $\infty$.

## ■ 1.2.2 The dual problem

By symmetry, the optimality conditions (1.2) for the cone program **P** are also optimality conditions for the cone program

$$\mathbf{D}: \begin{array}{ll} \text{minimize} & \langle x_0, s\rangle \\ \text{subject to} & s \in (c + \mathcal{L}^{\perp}) \cap \mathcal{K}^*. \end{array} \tag{1.5}$$

Together, one calls **P** and **D** a *primal-dual* pair and refers to **P** as the *primal problem* and **D** as the *dual problem* of **P**. From (1.3)-(1.4), we see that strictly separating hyperplanes for **P** are improving rays for **D** and vice versa.

## ■ 1.2.3 Alternative forms

#### Representations of the primal

The affine set $x_0 + \mathcal{L}$ of the cone program **P** is usually given using one of two representations. The first is an *implicit* representation

$$x_0 + \mathcal{L} = \{x \in \mathcal{V} : Ax = b\},$$

where $A : \mathcal{V} \to \mathbb{R}^m$ is a linear map with *kernel* equal to $\mathcal{L}$ and $b \in \mathbb{R}^m$ satisfies $b = Ax_0$. The other is a *parametric* representation in parameter $z \in \mathbb{R}^n$, given by

$$x_0 + \mathcal{L} = \{x_0 + Bz : z \in \mathbb{R}^n\},$$

where $B : \mathbb{R}^n \to \mathcal{V}$ is a linear map with *range* equal to $\mathcal{L}$. Replacing $x_0 + \mathcal{L}$ with each of these representations gives implicit and parametric forms of $\mathbf{P}$:

$$\begin{array}{lll} \text{minimize} & \langle c, x\rangle & \\ \text{subject to} & Ax = b & \text{(implicit)}, \\ & x \in \mathcal{K}, & \end{array} \qquad \begin{array}{lll} \text{minimize} & \langle c, x\rangle & \\ \text{subject to} & x = x_0 + Bz & \text{(parametric)}, \\ & (x, z) \in \mathcal{K} \times \mathbb{R}^n. & \end{array}$$

**Induced representation of dual**

The kernel of any linear map $T$ and the range of its adjoint $T^*$ are complementary orthogonal subspaces, i.e., $\operatorname{range} T^* = (\operatorname{null} T)^\perp$. For this reason, a parametric (resp. implicit) representation of $x_0 + \mathcal{L}$ induces an implicit (resp. parametric) representation of $c + \mathcal{L}^\perp$—the affine set of the dual problem $\mathbf{D}$. Two forms of $\mathbf{D}$ corresponding to these induced representations are

$$\begin{array}{lll} \text{minimize} & \langle x_0, s\rangle & \\ \text{subject to} & s = c + A^*u, & \text{(parametric)} \\ & (s, u) \in \mathcal{K}^* \times \mathbb{R}^m, & \end{array} \qquad \begin{array}{lll} \text{minimize} & \langle x_0, s\rangle & \\ \text{subject to} & B^*s = B^*c & \text{(implicit)}, \\ & s \in \mathcal{K}^*, & \end{array}$$

where $B^* : \mathcal{V} \to \mathbb{R}$ and $A^* : \mathcal{V} \to \mathbb{R}$ are the adjoint maps of $A$ and $B$.

**Min-max representation and weak duality**

One can also express $\mathbf{P}$ and $\mathbf{D}$ as a pair of minimization and maximization problems, where the dual optimal value lower bounds the primal optimal value—an inequality called *weak duality*. One obtains such a primal-dual pair from the implicit form of $\mathbf{P}$ and the parametric form of $\mathbf{D}$ by modifying the latter's objective function. To see this, note that for any feasible point $s \in \mathcal{V}$, the objective $\langle x_0, s\rangle$ of $\mathbf{D}$ satisfies

$$\langle x_0, s\rangle = \langle x_0, c + A^*u\rangle = \langle x_0, c\rangle + \langle Ax_0, u\rangle = \langle x_0, c\rangle + \langle b, u\rangle.$$

Removing the constant term $\langle x_0, c\rangle$ makes the objective linear without changing the feasible or optimal set. Further, substituting $u$ with $-y$ converts $\mathbf{D}$ into a maximization problem. Together, these changes yield a primal-dual pair with the aforementioned properties:

$$\begin{array}{ll} \text{minimize} & \langle c, x\rangle \\ \text{subject to} & Ax = b, \\ & x \in \mathcal{K} \end{array} \qquad \begin{array}{ll} \text{maximize} & \langle b, y\rangle \\ \text{subject to} & s = c - A^*y \\ & (s, y) \in \mathcal{K}^* \times \mathbb{R}^m. \end{array}$$

Indeed, for this pair, the weak duality inequality $\langle b, y\rangle \leq \langle c, x\rangle$ holds for feasible points, i.e.,

$$0 \leq \langle s, x\rangle = \langle c - A^*y, x\rangle = \langle c, x\rangle - \langle y, Ax\rangle = \langle c, x\rangle - \langle y, b\rangle,$$

and it is tight, i.e., $\langle b, y\rangle = \langle c, x\rangle$, if and only if $\langle s, x\rangle = 0$.

## ■ 1.2.4 Pathological instances

We say a cone program is *pathological* if none of the following three objects exist: a solution to the optimality conditions, a feasible point paired with an improving ray, or a strictly separating hyperplane, i.e., a dual improving ray. As we saw in Section 1.2.1, these three objects demonstrate optimality of a point, or unboundedness/infeasibility of the cone program. Linear programs are never pathological: exactly one of these objects always exists. The same is not true for general cone programs.

#### Failure of the optimality conditions: duality gaps and unattainment

Suppose the optimal value of a cone program is finite. Then the cone program is pathological if and only if the optimality conditions fail to have a solution. This failure occurs for three (nonexclusive) reasons. The first is that the cone program *itself* has an unattained optimal value. The next is that the *dual* has an unattained optimal value. The third is that the primal and dual optimal values are different, i.e., there is positive *duality gap*. The next examples illustrate both unattainment and duality gaps using semidefinite programs.

**Example 1.2.1** (Unattained optimal values)**.** *Consider the semidefinite program in decision variable $x = (x_1, x_2)$:*

$$\begin{array}{ll} \text{minimize} & x_1 \\ \text{subject to} & \begin{bmatrix} x_1 & 1 \\ 1 & x_2 \end{bmatrix} \in \mathbb{S}^2_+. \end{array} \tag{1.6}$$

*A point $x$ is feasible if and only if the one-by-one principal minors $x_1$ and $x_2$ and the determinant $x_1x_2 - 1$ are all nonnegative. Hence, the optimal value of this SDP is*

$$\theta = \inf \{x_1 : x_1 \geq 0, x_2 \geq 0, x_1x_2 \geq 1\},$$

*which is finite ($\theta = 0$) and unattained.*

**Example 1.2.2** (Positive dual gap)**.** *Consider the primal-dual pair of semidefinite* *programs*

$$\begin{array}{llll} \text{minimize} & \operatorname{Tr} CX & \qquad \text{maximize} & b^T y \\ \text{subject to} & \operatorname{Tr} A_i X = b_i \;\; \forall i \in \{1,2\} & \qquad \text{subject to} & S = C - \sum_{i=1}^2 y_i A_i, \\ & X \in \mathbb{S}^3_+, & & (S, y) \in \mathbb{S}^3_+ \times \mathbb{R}^2, \end{array}$$

*where*

$$C := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad A_1 := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \qquad A_2 := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad b := \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

*The duality gap of this primal-dual pair equals one. Indeed, $X$ is primal feasible only if $\operatorname{Tr} CX = 1$, and $(y, S)$ is dual feasible only if $b^T y = 0$. Further, $X = C$ and $y = (0,0)^T$ are primal-dual feasible points.*

#### Failure of the infeasibility condition: weak infeasibility

Suppose a cone program has feasible set $\mathcal{A} \cap \mathcal{K}$, where $\mathcal{A} \subseteq \mathcal{V}$ is affine and $\mathcal{K} \subseteq \mathcal{V}$ is a convex cone. When $\mathcal{A} \cap \mathcal{K}$ is empty, the cone program is pathological if and only if no hyperplane strictly separates $\mathcal{A}$ from $\mathcal{K}$. In this situation, we say the cone program is *weakly infeasible.* The next example illustrates weak infeasibility of an SDP.

**Example 1.2.3.** *Suppose a cone program has feasible set $\mathcal{A} \cap \mathcal{K}$, where $\mathcal{K}$ is the psd cone $\mathbb{S}^2_+$ and $\mathcal{A}$ is the affine set*

$$\mathcal{A} = \left\{ \begin{bmatrix} t & 1 \\ 1 & 0 \end{bmatrix} : t \in \mathbb{R} \right\}.$$

*The intersection $\mathcal{A} \cap \mathcal{K}$ is empty but no hyperplane strictly separates $\mathcal{A}$ from $\mathcal{K}$, i.e., no $S \in \mathbb{S}^n$ satisfies*

$$\operatorname{Tr} SX < 0 \quad \forall X \in \mathcal{A}, \qquad \operatorname{Tr} SX \geq 0 \quad \forall X \in \mathbb{S}^2_+.$$

*To see this, note that these separation conditions are equivalent to*

$$S_{11} = 0, \;\; S_{12} < 0, \;\; S \in \mathbb{S}^2_+,$$

*which are unsatisfiable since $S_{11} = 0$ and $S \in \mathbb{S}^2_+$ imply that $S_{12} = 0$.*

#### Failure of the unboundedness condition: weak infeasibility of the dual

Suppose now that a cone program is unbounded. In this case it is pathological if and only if it has no improving ray since a feasible point exists by the unboundedness assumption. Equivalently, it is pathological if and only if its dual is weakly infeasible.

## 1.2.5 Projected reformulations

Consider a primal-dual pair formulated over a linear transformation of a closed, convex cone $\mathcal{C} \subseteq \mathcal{W}$:

$$\mathbf{P}_\Phi : \begin{array}{ll} \text{minimize} & \langle c, x\rangle \\ \text{subject to} & x \in x_0 + \mathcal{L}, \\ & x \in \Phi \cdot \mathcal{C}, \end{array} \qquad \mathbf{D}_\Phi : \begin{array}{ll} \text{minimize} & \langle x_0, s\rangle \\ \text{subject to} & s \in c + \mathcal{L}^\perp, \\ & s \in (\Phi \cdot \mathcal{C})^*. \end{array}$$

Here, $\Phi : \mathcal{W} \to \mathcal{V}$ is an injective linear map and $\Phi \cdot \mathcal{C} := \{\Phi z : z \in \mathcal{C}\}$ denotes the image of $\mathcal{C}$ under $\Phi$. Reformulating this primal-dual pair using only variables from $\mathcal{W}$ is a goal of later chapters. We call reformulations over $\mathcal{W}$ *projected reformulations.*

We can always find a projected reformulation of the primal $\mathbf{P}_\Phi$. Finding one for the dual $\mathbf{D}_\Phi$ requires an additional assumption.

**Assumption 1.2.1.** *The range of $\Phi$ contains $x_0 \in \mathcal{V}$, i.e., $x_0 = \Phi t_0$ for some $t_0 \in \mathcal{W}$.*

Replacing $x_0$ by any point in $(x_0 + \mathcal{L}) \cap \operatorname{range} \Phi$ does not change the primal-dual solution sets. Hence, with this replacement, this assumption is satisfiable whenever $(x_0 + \mathcal{L}) \cap \operatorname{range} \Phi$ is nonempty.

**Projected primal** Substituting $x = \Phi z$ for $z \in \mathcal{C}$ and replacing the affine set $x_0 + \mathcal{L}$ with its inverse image under $\Phi$ yields the desired reformulation of $\mathbf{P}_\Phi$:

$$\begin{array}{ll} \text{minimize} & \langle \Phi^* c, z\rangle \\ \text{subject to} & z \in \Phi^{-1}(x_0 + \mathcal{L}), \\ & z \in \mathcal{C}, \end{array}$$

where $\Phi^{-1}(x_0 + \mathcal{L}) := \{z : \Phi z \in x_0 + \mathcal{L}\}$. If $z$ solves this problem, then $\Phi z$ solves $\mathbf{P}_\Phi$, i.e., one recovers a solution to $\mathbf{P}_\Phi$ simply by evaluating $\Phi$ at $z$. We also have formulas for $\Phi^{-1}(x_0 + \mathcal{L})$ under Assumption 1.2.1.

**Lemma 1.2.1.** *Suppose that $x_0 = \Phi t_0$ for $t_0 \in \mathcal{W}$ (Assumption 1.2.1). Then,*

$$\Phi^{-1}(x_0 + \mathcal{L}) = t_0 + \Phi^{-1}(\mathcal{L}),$$

*where, in addition, $\Phi^{-1}(\mathcal{L}) = \Phi^+ \cdot (\mathcal{L} \cap \operatorname{range} \Phi)$.*

*Proof.* The inclusion $\supseteq$ is obvious. Suppose that $z$ is in the inverse image, i.e., suppose that $\Phi z = \Phi t_0 + r$ for $r \in \mathcal{L}$. Then, $r \in \operatorname{range} \Phi$. Further,

$$z = \Phi^+ \Phi z = \Phi^+(\Phi t_0 + r) = t_0 + \Phi^+ r.$$

Since $r \in \operatorname{range} \Phi$, it holds that $\Phi \Phi^+ r = r$. Since $r \in \mathcal{L}$, we conclude $\Phi^+ r \in \Phi^{-1}(\mathcal{L})$, showing that $z \in t_0 + \Phi^{-1}(\mathcal{L})$. □

Under Assumption 1.2.1, the projected primal simplifies to

$$\begin{array}{ll} \text{minimize} & \langle \Phi^* c, z\rangle \\ \text{subject to} & z \in t_0 + \Phi^{-1}(\mathcal{L}), \\ & z \in \mathcal{C}. \end{array}$$

**Partially projected dual** To reformulate the dual, we let $\Phi^+ : \mathcal{V} \to \mathcal{W}$ denote the *pseudo-inverse* of $\Phi$, given by

$$\Phi^+ = (\Phi^*\Phi)^{-1}\Phi^*,$$

where $\Phi^*\Phi$ is invertible since $\Phi$ is injective. Note that the orthogonal projection map onto the range of $\Phi$ equals $\Phi\Phi^+$ and the range of $\Phi$ equals that of $(\Phi^+)^*$. Consider now the direct-sum decomposition of the dual variable $s \in \mathcal{V}$ induced by

$$\mathcal{V} = (\operatorname{range}\Phi) \oplus (\operatorname{range}\Phi)^\perp.$$

Using the equalities $\operatorname{range}(\Phi^+)^* = \operatorname{range}\Phi$ and $(\operatorname{range}\Phi)^\perp = \operatorname{null}\Phi^*$, one can write this as

$$s = (\Phi^+)^* y + w, \qquad \Phi^* w = 0, \quad (w, y) \in \mathcal{V} \times \mathcal{W}. \tag{1.7}$$

Since $(\Phi \cdot \mathcal{C})^* = \{s \in \mathcal{V} : \Phi^* s \in \mathcal{C}^*\}$ and $\Phi^*(\Phi^+)^* y = (\Phi^+\Phi)^* y = (\Phi^+\Phi) y = y$, it follows that

$$s \in (\Phi \cdot \mathcal{C})^* \Leftrightarrow \Phi^* s \in \mathcal{C}^* \Leftrightarrow y \in \mathcal{C}^*.$$

This gives a *partially* projected dual reformulation that still includes a variable $w \in \mathcal{V}$:

$$\begin{array}{ll} \text{minimize} & \langle x_0, (\Phi^+)^* y + w\rangle \\ \text{subject to} & (\Phi^+)^* y + w \in c + \mathcal{L}^\perp, \\ & y \in \mathcal{C}^*, \Phi^* w = 0. \end{array}$$

Note that if $(y, w)$ solves this reformulation, then $s = (\Phi^+)^* y + w$ solves $\mathbf{D}_\Phi$ by (1.7).

**Fully projected dual** Under the assumption $x_0 = \Phi t_0$ (Assumption 1.2.1), we can remove the variable $w$ from the partially projected dual. Indeed, if this assumption holds, the objective function satisfies

$$\langle x_0, (\Phi^+)^* y + w\rangle = \langle \Phi t_0, (\Phi^+)^* y + w\rangle = \langle t_0, \Phi^*(\Phi^+)^* y + \Phi^* w\rangle = \langle t_0, y\rangle$$

since $\Phi^* w = 0$ and $\Phi^*(\Phi^+)^* y = y$. We can eliminate $w \in (\operatorname{range}\Phi)^\perp$ from the affine constraint by replacing $c + \mathcal{L}^\perp$ with its orthogonal projection onto the range of $\Phi$. Indeed, $y$ satisfies the affine constraint for some $w$ if and only if

$$(\Phi^+)^* y \in \Phi\Phi^+ \cdot (c + \mathcal{L}^\perp).$$

We now claim this projected affine constraint holds if and only if $y \in \Phi^* c + \Phi^* \cdot \mathcal{L}^\perp$. Sufficiency follows given that $(\Phi^+)^*\Phi^* = \Phi\Phi^+$. For necessity, suppose that $(\Phi^+)^* y \in \Phi\Phi^+(c + \mathcal{L}^\perp)$. Multiplying both sides by $\Phi^*$ and using the identities

$$\Phi^*(\Phi^+)^* y = y, \qquad \Phi^*\Phi\Phi^+ = \Phi^*\Phi(\Phi^*\Phi)^{-1}\Phi^* = \Phi^*,$$

shows that $y \in \Phi^*(c+\mathcal{L}^\perp)$ as desired. This gives a fully projected reformulation of $\mathbf{D}_\Phi$:

$$\begin{array}{ll} \text{minimize} & \langle t_0, y\rangle \\ \text{subject to} & y \in \Phi^* c + \Phi^* \cdot \mathcal{L}^\perp, \\ & y \in \mathcal{C}^*. \end{array}$$

For any feasible $y \in \mathcal{V}$, one can solve linear equations to find $w$ satisfying $\Phi^* w = 0$ and $(\Phi^+)^* y + w \in (c + \mathcal{L}^\perp)$. This in turn gives a feasible point $s = (\Phi^+)^* y + w$ of $\mathbf{D}_\Phi$.

**Primal-dual pair interpretation** Under the assumption $x_0 = \Phi t_0$ (Assumption 1.2.1), we have just individually reformulated the primal and dual as

$$\begin{array}{ll} \text{minimize} & \langle \Phi^* c, z\rangle \\ \text{subject to} & z \in t_0 + \Phi^{-1}(\mathcal{L}), \\ & z \in \mathcal{C}, \end{array} \qquad \begin{array}{ll} \text{minimize} & \langle t_0, y\rangle \\ \text{subject to} & y \in \Phi^* c + \Phi^* \cdot \mathcal{L}^\perp, \\ & y \in \mathcal{C}^*. \end{array}$$

These reformulations are a primal-dual pair in the sense of Section 1.2.2. Specifically, the orthogonal complement of $\Phi^{-1}(\mathcal{L})$ equals $\Phi^* \cdot \mathcal{L}^\perp$.

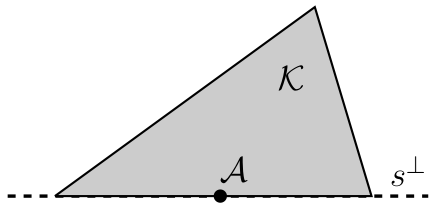
## ■ 1.3 Facial reduction

This section provides background on facial reduction [46, 20, 102]. The high level idea behind this technique is fairly simple (Figure 1.1). Given the cone program

$$\begin{array}{ll} \text{minimize} & \langle c, x\rangle \\ \text{subject to} & x \in \mathcal{A} \cap \mathcal{K}, \end{array}$$

where $\mathcal{A} \subseteq \mathcal{V}$ is affine and $\mathcal{K} \subseteq \mathcal{V}$ is a convex cone, one finds a hyperplane $s^\perp := \{x \in \mathcal{V} : \langle s, x\rangle = 0\}$ that contains $\mathcal{A}$ and hence the feasible set $\mathcal{A} \cap \mathcal{K}$. One also imposes the condition that $s \in \mathcal{K}^*$, which implies that $\mathcal{K} \cap s^\perp$ is a *face* of $\mathcal{K}$ *exposed* by $s$.

The set of $s \in \mathcal{K}^*$ satisfying $s^\perp \supseteq \mathcal{A}$ is the feasible set of an *auxiliary problem*. This problem has nontrivial solutions only if the intersection of $\mathcal{A}$ with the *relative interior* of $\mathcal{K}$ is empty. Hence, finding hyperplanes requires a failure of *Slater's condition*. To give more details, we first review some basic concepts from convex analysis.




**(a)** Geometry

| | |
|---|---|
| Find | $s \in \mathcal{K}^* \setminus (\operatorname{span} \mathcal{K})^\perp$ |
| subject to | $s^\perp \supseteq \mathcal{A}$ |

**(b)** Auxiliary problem

**Figure 1.1:** Facial reduction finds a hyperplane $s^\perp$ containing the affine set $\mathcal{A}$ that exposes a face of $\mathcal{K}$. This hyperplane is the solution of an auxiliary problem.

## ■ 1.3.1 Relative interior, faces, and exposed faces

Important sets related to a convex cone $\mathcal{K}$ are its relative interior, its faces, and its exposed faces, which we define next. Note that compatible (and more complicated) definitions can be given for general convex sets.

### Relative interior

The relative interior of a convex cone $\mathcal{K}$ is its interior relative to the smallest linear subspace containing it. Specifically,

$$\operatorname{relint} \mathcal{K} := \{x \in \mathcal{K} : B(x, r) \cap \operatorname{span} \mathcal{K} \subseteq \mathcal{K} \text{ for some } r > 0\},$$

where $B(x, r) = \{y \in \mathcal{V} : \|x - y\| \leq r\}$ is a ball of radius $r$ (in any norm) and

$$\operatorname{span} \mathcal{K} := \{t_1 x + t_2 y : \quad x, y \in \mathcal{K}, \quad t_1, t_2 \in \mathbb{R}\}.$$

Note that if $\operatorname{span} \mathcal{K}$ equals $\mathcal{V}$ then the relative interior is simply called the interior. The interior of the nonnegative orthant $\mathbb{R}^n_+$ is the subset of vectors with strictly positive entries. The interior of the Lorentz cone $\mathcal{Q}^{n+1} \subseteq \mathbb{R} \times \mathbb{R}^n$ is the subset of $(t, x) \in \mathbb{R} \times \mathbb{R}^n$ satisfying $t > \|x\|_2$. Finally, the interior of the psd cone $\mathbb{S}^n_+$ is the subset of symmetric matrices $X$ satisfying $v^T X v > 0$ for all nonzero $v \in \mathbb{R}^n$ or, equivalently, the subset of symmetric matrices with strictly positive eigenvalues (Table 1.2(a)).

The following proposition lists important properties of the relative interior.

**Proposition 1.3.1.** *Let $\mathcal{K} \subseteq \mathcal{V}$ be a nonempty, convex cone. The following statements hold for all $x \in \operatorname{relint} \mathcal{K}$.*

- *There exists an $\alpha > 0$ for which $x \pm \alpha y \in \mathcal{K}$ for all $y \in \mathcal{K}$.*
- *If $s \in \mathcal{K}^*$ and $\langle s, x \rangle = 0$, then $\langle s, z \rangle = 0$ for all $z \in \mathcal{K}$.*

*Proof.* By definition, there exists $r > 0$ for which $x \pm rz \in \mathcal{K}$ for all $z \in \mathcal{K}$ with $\|z\| = 1$. Taking $\alpha = \frac{r}{\|y\|}$ shows $x \pm \alpha y \in \mathcal{K}$, as desired. For the second statement, pick any

$z \in \mathcal{K}$. Then, $x \pm \alpha z \in \mathcal{K}$ for some $\alpha > 0$. Hence, $\langle s, x \pm \alpha z\rangle \geq 0$ since $s \in \mathcal{K}^*$. If $\langle s, x\rangle = 0$, it follows that $\pm\langle s, \alpha z\rangle \geq 0$, showing $\frac{1}{\alpha}\langle s, z\rangle = 0$. $\square$

The first statement is a convenient restatement of the definition of relative interior. The second statement, among other things, allows us to conveniently describe the subset of $\mathcal{K}^*$ not contained in $(\operatorname{span}\mathcal{K})^\perp$. Specifically, for any $x \in \operatorname{relint}\mathcal{K}$,

$$\mathcal{K}^* \setminus (\operatorname{span}\mathcal{K})^\perp = \{s \in \mathcal{K}^* : \langle s, x\rangle > 0\}.$$

As we will see, this description of $\mathcal{K}^* \setminus (\operatorname{span}\mathcal{K})^\perp$ is critical for facial reduction algorithms.

#### Faces

A *face* of a convex cone $\mathcal{K}$ is a convex cone $\mathcal{F} \subseteq \mathcal{K}$ for which $a, b \in \mathcal{K}$ and $a + b \in \mathcal{F}$ imply $a, b \in \mathcal{F}$. A face is *proper* if it is nonempty and not equal to $\mathcal{K}$. When $\mathcal{K}$ is closed, so are its faces. The arbitrary intersection of faces is a face, and faces of faces are faces. This and other properties follow.

**Proposition 1.3.2** (Properties of faces)**.** *For any face $\mathcal{F} \subseteq \mathcal{K}$ of a convex cone $\mathcal{K} \subseteq \mathcal{V}$, the following statements hold.*

- *If $\mathcal{F} \cap \operatorname{relint}\mathcal{K} \neq \emptyset$, then $\mathcal{F} = \mathcal{K}$.*
- *Any face of $\mathcal{F}$ is a face of $\mathcal{K}$.*
- *$\mathcal{F} = \mathcal{K} \cap \operatorname{span}\mathcal{F}$.*
- *If $\mathcal{G}$ is a face of $\mathcal{K}$, then $\mathcal{F} \cap \mathcal{G}$ is a face of $\mathcal{K}$.*

*Further, if $T$ is an arbitrary set of faces, then the intersection $\bigcap_{\mathcal{G}\in T}\mathcal{G}$ is a face.*

*Proof.* Suppose $x \in \mathcal{F} \cap \operatorname{relint}\mathcal{K}$. Then for arbitrary $y \in \mathcal{K}$, there exists $\alpha > 0$ for which $x - \alpha y \in \mathcal{K}$. Since $x = \alpha y + x - \alpha y$, the face $\mathcal{F}$ must contain both $\alpha y \in \mathcal{K}$ and $x - \alpha y \in \mathcal{K}$ and therefore equal $\mathcal{K}$.

Suppose $\mathcal{G}$ is a face of $\mathcal{F}$ and that $a, b \in \mathcal{K}$ satisfy $a + b \in \mathcal{G}$. Since $\mathcal{G} \subseteq \mathcal{F}$, we conclude $a, b \in \mathcal{F}$ since $\mathcal{F}$ is a face of $\mathcal{K}$. Hence, $a, b \in \mathcal{G}$ since $\mathcal{G}$ is a face of $\mathcal{F}$

The inclusion $\mathcal{F} \subseteq \mathcal{K} \cap \operatorname{span}\mathcal{F}$ is obvious. For the reverse, we can write any $z \in \mathcal{K} \cap \operatorname{span}\mathcal{F}$ as $z = x - y$ for $x, y \in \mathcal{F}$. Since $x = z + y$, we conclude $z \in \mathcal{F}$.

Suppose $\mathcal{G}$ is another face of $\mathcal{K}$. If $a, b \in \mathcal{K}$ satisfy $a + b \in \mathcal{F} \cap \mathcal{G}$. Then, $a, b \in \mathcal{F}$ and $a, b \in \mathcal{G}$ since these sets are both faces. Hence, $a, b \in \mathcal{F} \cap \mathcal{G}$.

Since the intersection of two faces is a face, the intersection of finitely-many faces is face. For an arbitrary set of faces $T$,

$$\bigcap_{\mathcal{G}\in T}\mathcal{G} = \bigcap_{\mathcal{G}\in T}\mathcal{K} \cap \operatorname{span}\mathcal{G} = \mathcal{K} \cap \left(\bigcap_{\mathcal{G}\in T}\operatorname{span}\mathcal{G}\right) = \mathcal{K} \cap \left(\bigcap_{\mathcal{G}\in S}\operatorname{span}\mathcal{G}\right)$$

for some finite subset $S \subseteq T$ given that $\mathcal{V}$ has finite dimension. Hence, an arbitrary intersection $\bigcap_{\mathcal{G}\in T}\mathcal{G}$ equals a finite intersection $\bigcap_{\mathcal{G}\in S}\mathcal{G}$ and is hence a face.

□

By this proposition, the intersection of all faces containing a subset $\mathcal{C} \subseteq \mathcal{K}$ is a face. One calls this intersection the *minimal face* of $\mathcal{C}$.

**Example 1.3.1.** The nonzero, proper faces of the nonnegative orthant $\mathbb{R}^n_+$, Lorentz cone $\mathcal{Q}^{n+1}$, and psd cone $\mathbb{S}^n_+$ are easy to describe. Subsets of $[n] := \{1, \dots, n\}$ parametrize faces of $\mathbb{R}^n_+$; the face $\mathcal{F}_I$ corresponding to $I \subseteq [n]$ is

$$\mathcal{F}_I = \{x \in \mathbb{R}^n_+ : x_I = 0\}.$$

Vectors in $\mathbb{R}^n$ parametrize faces of the Lorentz cone $\mathcal{Q}^{n+1} \subseteq \mathbb{R}$; the face $\mathcal{F}_u$ corresponding to $u \in \mathbb{R}^n$ is the one-dimensional cone generated by $(\|u\|_2, u)$:

$$\mathcal{F}_u = \{\lambda(\|u\|_2, u) : \lambda \geq 0\}.$$

Note that $\mathcal{F}_u = \mathcal{F}_v$ if $(\|u\|_2, u)$ and $(\|v\|_2, v)$ are collinear. Finally, subspaces of $\mathbb{R}^n$ parametrize faces of the psd cone; the face $\mathcal{F}_\mathcal{L}$ corresponding to the subspace $\mathcal{L} \subseteq \mathbb{R}^n$ is the set

$$\mathcal{F}_\mathcal{L} = \{X \in \mathbb{S}^n_+ : \operatorname{range} X \subseteq \mathcal{L}\},$$

a result originally due to [10]. Note that if $U \in \mathbb{R}^{n\times d}$ has full column rank and range equal to $\mathcal{L}$, then

$$\mathcal{F}_\mathcal{L} = \{U\hat{X}U^T : \hat{X} \in \mathbb{S}^d_+\}.$$

Hence, we can also parametrize faces of $\mathbb{S}^n_+$ by full rank $n \times d$ matrices. Table 1.2(b) summarizes these parametrization; see also [101]. Note that the faces $\mathcal{F}_I$, $\mathcal{F}_u$, and $\mathcal{F}_\mathcal{L}$ are injective linear transformations of $\mathbb{R}^{n-|I|}_+$, $\mathbb{R}_+$ and $\mathbb{S}^d_+$, respectively; in other words, each of these faces is isomorphic to a nonnegative orthant or a psd cone.

#### Exposed faces

For $s \in \mathcal{V}$, let $s^\perp$ denote the hyperplane normal to $s$ passing through the origin, i.e.,

$$s^\perp := \{s \in \mathcal{V} : \langle s, x\rangle = 0\}.$$

If $s \in \mathcal{K}^*$, then $\mathcal{K} \cap s^\perp$ is a face. To see this, note that if $x + y \in \mathcal{K} \cap s^\perp$ and $x, y \in \mathcal{K}$, then

$$0 = \langle s, x + y\rangle = \langle s, x\rangle + \langle s, y\rangle,$$

| $\mathcal{K}$ | $\mathcal{V}$ | **interior** | **faces** |
|---|---|---|---|
| $\mathbb{R}^n_+$ | $x \in \mathbb{R}^n$ | $x_i > 0 \;\; \forall i \in [n]$ | $x_i = 0 \;\; \forall i \in I$ |
| $\mathcal{Q}^{n+1}$ | $(x_0, x) \in \mathbb{R} \times \mathbb{R}^n$ | $x_0 > \|x\|_2$ | $x_0 = \|x\|_2$ |
| $\mathbb{S}^n_+$ | $\mathbb{S}^n$ | $z^T X z > 0 \;\; \forall z \in \mathbb{R}^n$ | $z^T X z = 0 \;\; \forall z \in W$ |

**(a)** Ambient space $\mathcal{V}$ and extra constraints satisfied by interior and proper faces of cone $\mathcal{K}$, where $I \subseteq \{1, \ldots, n\}$ and $W$ is a subspace of $\mathbb{R}^n$.

| $\mathcal{K}$ | $\mathcal{F}$ | $\mathcal{F}^*$ |
|---|---|---|
| $\mathbb{R}^n_+$ | $\left\{x \in \mathbb{R}^n : x_I = 0, x_{[n]\setminus I} \in \mathbb{R}^d_+\right\}$ | $\{s \in \mathbb{R}^n : s_i \geq 0 \;\; \forall i \in [n] \setminus I\}$ |
| $\mathcal{Q}^{n+1}$ | $\{\lambda(\|u\|_2, u) : \lambda \in \mathbb{R}_+\}$ | $\left\{(s_0, s) : s_0\|u\|_2 + s^T u \geq 0\right\}$ |
| $\mathbb{S}^n_+$ | $\left\{UXU^T : X \in \mathbb{S}^d_+\right\}$ | $\left\{S \in \mathbb{S}^n : U^T S U \in \mathbb{S}^d_+\right\}$ |

**(b)** Representations of faces as linear transformations of $\mathbb{R}^d_+$, $\mathbb{R}_+$ and $\mathbb{S}^d_+$ and the induced representation of their dual cones, where $I \subseteq \{1, \ldots, n\}$, $u \in \mathbb{R}^n$ is nonzero, and $U \in \mathbb{R}^{n \times d}$ has full column rank.

| $\mathcal{K}, s$ | **Parameters of** $\mathcal{K} \cap s^\perp$ | $\mathcal{K} \cap s^\perp$ |
|---|---|---|
| $\mathbb{R}^n_+, s$ | $I = \{i \in [n] : s_i > 0\}$ | $\left\{x \in \mathbb{R}^n : x_I = 0, x_{[n]\setminus I} \in \mathbb{R}^d_+\right\}$ |
| $\mathcal{Q}^{n+1}, (\|s\|_2, s)$ | $u = -s$ | $\{\lambda(\|u\|_2, u) : \lambda \in \mathbb{R}_+\}$ |
| $\mathbb{S}^n_+, S$ | $U \in \mathbb{R}^{n \times d}$, $U^T S = 0.$ | $\left\{UXU^T : X \in \mathbb{S}^d_+\right\}$ |
| $U\mathbb{S}^d_+ U^T, S$ | $B \in \mathbb{R}^{d \times r}$, $B^T(U^T S U) = 0$ | $\left\{(UB)X(UB)^T : X \in \mathbb{S}^r_+\right\}$ |

**(c)** Relationship between exposing vector $s \in \mathcal{K}^*$ and parameters $I, u, U, B$ of face $\mathcal{K} \cap s^\perp$.

**Table 1.2:** Properties of nonnegative orthant $\mathbb{R}^n_+$, Lorentz cone $\mathcal{Q}^{n+1}$, and psd cone $\mathbb{S}^n_+$ and their faces.

which shows that both $\langle s, x\rangle$ and $\langle s, y\rangle$ equal zero since they are both nonnegative. Hence, $\mathcal{K} \cap s^\perp$ contains both $x \in \mathcal{K}$ and $y \in \mathcal{K}$ if it contains their sum $x + y$. Faces of the form $\mathcal{K} \cap s^\perp$ are called *exposed faces* and $s$ the *exposing vector*. A cone is called *facially exposed* if all of its proper faces are exposed faces.

**Example 1.3.1** (Continued)**.** The nonnegative orthant $\mathbb{R}^n_+$, Lorentz cone $\mathcal{Q}^{n+1}$, and psd cone $\mathbb{S}^n_+$ are all facially exposed. Further, for these cones, simple parameterizations of $\mathcal{K} \cap s^\perp$ are obtained from $s$; see Table 1.2(c).

## ■ 1.3.2 Separating hyperplanes

Two subsets $\mathcal{C}_1$ and $\mathcal{C}_2$ of an inner product space $\mathcal{V}$ are disjoint (i.e., their intersection is empty) if there exists a hyperplane that *strictly separates* them, i.e., if there exists $s \in \mathcal{V}$ and $\alpha \in \mathbb{R}$ for which

$$\mathcal{C}_1 \subseteq \{x \in \mathcal{V} : \langle s, x \rangle \leq \alpha\}, \qquad \mathcal{C}_2 \subseteq \{x \in \mathcal{V} : \langle s, x \rangle > \alpha\}.$$

If $\mathcal{C}_1$ and $\mathcal{C}_2$ are both convex, then partial converses hold that vary based on the specific form of $\mathcal{C}_1$ and $\mathcal{C}_2$. The following provides such a converse assuming one set is the relative interior of a convex cone and the other set is affine. As indicated, under this assumption, there exists a separating hyperplane that actually contains the affine set.

**Proposition 1.3.3.** *Let $\mathcal{A} = x_0 + \mathcal{L}$ denote the affine set defined by $x_0 \in \mathcal{V}$ and a linear subspace $\mathcal{L} \subseteq \mathcal{V}$. Let $\mathcal{K} \subseteq \mathcal{V}$ be a nonempty, convex cone. The following are equivalent.*

1. *$\mathcal{A} \cap \operatorname{relint} \mathcal{K}$ is empty.*
2. *There exists $s \in \mathcal{K}^* \cap \mathcal{L}^\perp$ and $\alpha \leq 0$ with the following properties.*
   - *The hyperplane $\{x \in \mathcal{V} : \langle s, x \rangle = \alpha\}$ contains $\mathcal{A}$.*
   - *$s \notin (\operatorname{span} x_0)^\perp \cap (\operatorname{span} \mathcal{K})^\perp$.*

*Proof.* Suppose the hyperplane exists. If $\alpha < 0$, then $\mathcal{A} \cap \mathcal{K} = \emptyset$ by definition of $\mathcal{K}^*$. Suppose $\alpha = 0$. Then $s \in (\operatorname{span} x_0)^\perp$ and is therefore not contained in $(\operatorname{span} \mathcal{K})^\perp$. If $x \in \mathcal{A} \cap \operatorname{relint} \mathcal{K}$ exists, then $\langle s, x \rangle = 0$, which implies $s \in (\operatorname{span} \mathcal{K})^\perp$, a contradiction.

For the other direction ($1 \Rightarrow 2$), the main separation theorem [121, Theorem 11.3] states that a hyperplane exists *properly* separating these sets. Using Theorem 11.7 of [121], we can additionally assume this hyperplane passes through the origin since $\mathcal{K}$ is a cone. In other words, there exists $s \in \mathcal{K}^*$ and $z \in \mathcal{A} \cup \mathcal{K}$ satisfying

$$\begin{aligned} &\langle s, x \rangle \leq 0, \quad \forall x \in \mathcal{A}, \\ &\langle s, z \rangle \neq 0 \end{aligned}$$

Let $\alpha = \langle s, x_0 \rangle$. The result follows by showing $s \in \mathcal{L}^\perp$. Suppose then that $w \in \mathcal{L}$ satisfies $\langle s, w \rangle \neq 0$. Then $\langle s, x_0 + \beta w \rangle > 0$ for some $\beta$ with $|\beta|$ large enough, a contradiction.

□

#### Farkas lemma

Note Proposition 1.3.3 does not apply to the affine set $\{x \in \mathcal{V} : Ax = b\}$ unless the linear equations $Ax = b$ have a solution. In other words, it does not apply to affine sets that are empty. Nevertheless, we can easily extend Proposition 1.3.3 to handle this case. This extension is the following conic version of Farkas Lemma.

**Corollary 1.3.1** (Farkas Lemma)**.** *Let $\hat{x}$ be any point in the relative interior of a convex cone $\mathcal{K} \subseteq \mathcal{V}$. For a linear map $A : \mathcal{V} \to \mathbb{R}^m$ and $b \in \mathbb{R}^m$, the following statements are equivalent.*

1. *There is no $x \in \operatorname{relint} \mathcal{K}$ satisfying $Ax = b$.*
2. *There exists $y \in \mathbb{R}^m$ for which*
   - *$A^*y \in \mathcal{K}^*$ and $\langle b, y\rangle \leq 0$,*
   - *$\langle \hat{x}, A^*y\rangle > 0$ or $\langle b, y\rangle < 0$,*

   *where $A^* : \mathbb{R}^m \to \mathcal{V}$ denotes the adjoint map.*

*Proof.* Suppose first that $Ax = b$ has a solution $x_0$. Then, directly applying Proposition 1.3.3, we conclude existence of $A^*y \in \mathcal{K}^*$ where $\langle x_0, A^*y\rangle = \langle b, y\rangle \leq 0$. If $\langle b, y\rangle = 0$, then $A^*y \notin (\operatorname{span} \mathcal{K})^\perp$, implying $\langle \hat{x}, A^*y\rangle > 0$ for all $\hat{x} \in \operatorname{relint} \mathcal{K}$. Suppose next that there is no solution to $Ax = b$. Then, we can separate the range of $A$ (a cone) from the affine set $\{b\}$. In this case, the dual cone is $(\operatorname{range} A)^\perp$, which equals the null space of the adjoint $A^* : \mathbb{R}^m \to \mathcal{V}$. Hence, we can find $y \in \mathbb{R}^m$ for which $A^*y = 0$ and $\langle b, y\rangle < 0$ by Proposition 1.3.3. □

#### Weak intersection and containment in faces

Suppose that an affine set $\mathcal{A}$ intersects a convex cone $\mathcal{K}$ but not its relative interior. (When $\mathcal{A} \cap \mathcal{K}$ is the feasible set of a cone program, this situation is sometimes called *weak feasibility* [87].) The next corollary of Proposition 1.3.3 asserts existence of a hyperplane $s^\perp$ that both contains $\mathcal{A}$ and exposes a face of $\mathcal{K}$.

**Corollary 1.3.2** (Weak intersection)**.** *Let $\mathcal{A} \subseteq \mathcal{V}$ be an affine set and $\mathcal{K} \subseteq \mathcal{V}$ a convex cone for which $\mathcal{A} \cap \mathcal{K}$ is non-empty. The following are equivalent.*

1. *$\mathcal{A} \cap \operatorname{relint} \mathcal{K}$ is empty.*
2. *There exists $s \in \mathcal{K}^* \setminus (\operatorname{span} \mathcal{K})^\perp$ for which the hyperplane $s^\perp$ contains $\mathcal{A}$.*

*Proof.* We will show the second statement is equivalent to the statement statement of Proposition 1.3.3, which, in turn, is equivalent to emptiness of $\mathcal{A} \cap \operatorname{relint} \mathcal{K}$. Clearly the second statement implies that of Proposition 1.3.3. Now suppose $s \in \mathcal{K}^*$ and $\alpha$ of Proposition 1.3.3 exists. Then, $\alpha = 0$ given that $\mathcal{A} \cap \mathcal{K}$ is non-empty. It follows $s \in (\mathcal{A})^\perp$, implying $s \in \mathcal{K}^* \setminus \mathcal{K}^\perp$. □

This corollary is the theoretical unpinning of facial reduction algorithms; see, e.g., [102, Lemma 1], [20, Theorem 7.1], [31, Lemma 12.6], and [138, Lemma 3.2] for related statements.

## ■ 1.3.3 Slater's condition

When the feasible set intersects the relative interior of the cone, certain pathologies cannot occur. To explain, recall the primal-dual pair of cone programs in the min-max form:

$$\mathbf{P}: \begin{array}{ll} \text{minimize} & \langle c, x\rangle \\ \text{subject to} & Ax = b, \\ & x \in \mathcal{K}, \end{array} \qquad \mathbf{D}: \begin{array}{ll} \text{maximize} & \langle b, y\rangle \\ \text{subject to} & s = c - A^*y, \\ & (s,y) \in \mathcal{K}^* \times \mathbb{R}^m. \end{array}$$

In general, finite optimal values of the primal and/or dual can be unattained. In addition, a *duality gap* can exist, i.e., the primal and dual optimal values can be different. In these situations, the optimality conditions for the primal-dual pair have no solution. These situations are ruled out by *Slater's condition.*

**Definition 1.3.1** (Slater's condition)**.** *We say that* Slater's condition holds for $\mathbf{P}$ *if there exists* $x \in \operatorname{relint} \mathcal{K}$ *satisfying* $Ax = b$*. Similarly, we say that* Slater's condition holds for $\mathbf{D}$ *if there exists* $y \in \mathbb{R}^m$ *satisfying* $c - A^*y \in \operatorname{relint} \mathcal{K}^*$

**Proposition 1.3.4.** *Let* $\theta_p \in \mathbb{R} \cup \{\pm\infty\}$ *and* $\theta_d \in \mathbb{R} \cup \{\pm\infty\}$ *denote the optimal values of the primal-dual pair* $\mathbf{P}$ *and* $\mathbf{D}$*, i.e.,*

$$\theta_p := \inf \{\langle c, x\rangle : x \in \{u \in \mathcal{K} : Au = b\}\}, \quad \theta_d = \sup \{\langle b, y\rangle : y \in \mathbb{R}^m,\ c - A^*y \in \mathcal{K}^*\}.$$

*The following statements are true.*

- *If Slater's condition holds for* $\mathbf{P}$*, then* $\theta_p = \theta_d$ *and* $\theta_d$ *is attained when* $\theta_p$ *is finite.*
- *If Slater's condition holds for* $\mathbf{D}$*, then* $\theta_p = \theta_d$ *and* $\theta_p$ *is attained when* $\theta_d$ *is finite.*

*Proof.* To ease notation, let $\theta = \theta_p$. Let $x_s \in \operatorname{relint} \mathcal{K}$ satisfy $Ax_s = b$. We only show the first statement, since $\mathbf{D}$ can be converted into a problem of the form $\mathbf{P}$ (and vice versa) without changing optimal values, attainment or satisfiability of Slater's condition (Chapter 1.2.3). First, assume no point in the relative interior of $\mathcal{K}$ attains the optimal value $\theta$. Then, $\langle c, x_s\rangle > \theta$. Further, the affine set $\{x \in \mathcal{V} : Ax = b, \langle c, x\rangle = \theta\}$ is disjoint from the relative interior of $\mathcal{K}$. By Farkas Lemma (Corollary 1.3.1), there exists $(\lambda, y)$ satisfying

$$\begin{aligned} A^*y + c\lambda \in \mathcal{K}^* \\ \langle b, y\rangle + \lambda\theta \leq 0 \\ \langle b, y\rangle + \lambda\theta < 0 \text{ or } \langle b, y\rangle + \lambda\langle c, x_s\rangle > 0. \end{aligned} \tag{1.8}$$

If $\lambda = 0$, these conditions imply $\langle b, y\rangle < 0$ and $A^*y \in \mathcal{K}^*$, which contradicts

$$0 \leq \langle x_s, A^*y\rangle = \langle Ax_s, y\rangle = \langle b, y\rangle.$$

Hence, $\lambda \neq 0$. Combining $\langle x_s, A^*y + c\lambda\rangle = \langle b, y\rangle + \lambda\langle c, x_s\rangle \geq 0$ with (1.8) shows

$$\lambda\left(\langle c, x_s\rangle - \theta\right) \geq 0.$$

Using the fact $\langle c, x_s\rangle - \theta > 0$ combined with $\lambda \neq 0$, we conclude $\lambda > 0$ and

$$A^*\frac{y}{\lambda} + c \in \mathcal{K}^*,$$

i.e., $-\frac{1}{\lambda}y$ is dual feasible. Weak duality $-\frac{1}{\lambda}\langle b, y\rangle \leq \theta$ and (1.8) imply $\frac{1}{\lambda}\langle b, y\rangle = \theta$.

We now assume a point in the relative interior of $\mathcal{K}$ is optimal, i.e., that it attains $\theta$. To construct a dual feasible point with objective $\theta$ we first let $\mathcal{L}$ denote the kernel of $A$. If $\hat{x} \in \operatorname{relint}\mathcal{K}$ is optimal, then for any $d \in \mathcal{L} \cap \operatorname{span}\mathcal{K}$ there exists $\alpha > 0$ for which $\hat{x} \pm \alpha d$ is feasible, implying $\langle c, d\rangle = 0$ by optimality of $\hat{x}$. Hence, $\mathcal{L} \cap (\operatorname{span}\mathcal{K}) \subseteq (\operatorname{span}\{c\})^\perp$. But this shows that $\operatorname{span}\{c\} \subseteq (\operatorname{span}\mathcal{K})^\perp + \mathcal{L}^\perp$. It follows that $c + \mathcal{L}^\perp$ contains a point $\hat{s} \in (\operatorname{span}\mathcal{K})^\perp$. Letting $\hat{s} = c + A^*\hat{y}$ for $\hat{y} \in \mathbb{R}^m$, we have that $0 = \langle \hat{x}, \hat{s}\rangle = \theta - b^T\hat{y}$.

□

This motivates the following question: can one transform a (feasible) cone program such that Slater's condition holds without changing the optimal value? Facial reduction provides a positive answer.

## ■ 1.3.4 The facial reduction algorithm

For any nonempty, convex cone $\mathcal{K} \subseteq \mathcal{V}$ and affine set $\mathcal{A} \subseteq \mathcal{V}$, the main separation theorem (Proposition 1.3.3) implies that at least one of the following statements holds:

(a) $\mathcal{A} \cap \operatorname{relint}\mathcal{K}$ is nonempty.

(b) A hyperplane exists that strictly separates $\mathcal{A}$ and $\mathcal{K}$.

(c) A hyperplane $s^\perp := \{x \in \mathcal{V} : \langle s, x\rangle = 0\}$ exists that contains $\mathcal{A}$ and exposes a proper face of $\mathcal{K}$.

Further, under the assumption $\mathcal{K}$ is closed, the face $\mathcal{K} \cap s^\perp$ is nonempty. (For instance, it contains the zero vector.) Under this assumption, one can always replace $\mathcal{K}$ with a face containing $\mathcal{A} \cap \mathcal{K}$ such that statement (a) or statement (b) holds. This replacement is called *facial reduction*.

To see this, note that statement (c) is equivalent to feasibility of the following *auxiliary problem*:

$$\begin{array}{ll} \text{Find} & s \in \mathcal{K}^* \setminus (\operatorname{span}\mathcal{K})^\perp \\ \text{subject to} & s^\perp \supseteq \mathcal{A}. \end{array}$$

Hence, if no solution $s$ exists, either statement (a) or (b) holds. On the other hand, given a solution $s$, we can replace $\mathcal{K}$ with $\mathcal{K}\cap s^\perp$ and then resolve. If we do this replace-and-resolve process a sufficient number of times (upper bounded by the dimension of $\mathcal{V}$), the auxiliary problem must become infeasible, implying that statement (a) or statement (b) holds as desired. These iterations are performed by the *facial reduction algorithm* (Algorithm 1.1), which has the following properties.

**Proposition 1.3.5.** *Let $\mathcal{F}$ be the output of the facial reduction algorithm (Algorithm 1.1) given a closed, convex cone $\mathcal{K} \subseteq \mathcal{V}$ and nonempty, affine set $\mathcal{A} \subseteq \mathcal{V}$ as input. Exactly one of the following statements holds.*

- *There exists a hyperplane strictly separating $\mathcal{F}$ from $\mathcal{A}$.*
- *$\mathcal{A} \cap \operatorname{relint} \mathcal{F}$ is non-empty.*

*Further, $\mathcal{A}\cap\mathcal{K} = \mathcal{A}\cap\mathcal{F}$. Finally, if $\mathcal{A}\cap\mathcal{K}$ is nonempty, then $\mathcal{F}$ is the minimal face of $\mathcal{A}\cap\mathcal{K}$, i.e., it is the intersection of all faces of $\mathcal{K}$ containing $\mathcal{A}\cap\mathcal{K}$.*

*Proof.* The listed properties follow by the paragraph preceding this proposition. That $\mathcal{A}\cap\mathcal{K} = \mathcal{A}\cap\mathcal{F}$ is obvious. Let $\mathcal{G}$ denote the minimal face. By definition, $\mathcal{F}$ contains $\mathcal{G}$ and $\mathcal{G}$ contains $\mathcal{A}\cap\mathcal{F}$. But this means $\mathcal{G}$ contains $\mathcal{A}\cap\operatorname{relint}\mathcal{F}$, which is nonempty. Hence, $\mathcal{G}$ must equal $\mathcal{F}$ by properties of faces (Proposition 1.3.2). □

**Algorithm 1.1:** Facial reduction

**Inputs:** closed, convex cone $\mathcal{K} \subseteq \mathcal{V}$ and affine set $\mathcal{A} \subseteq \mathcal{V}$
**Output:** a face $\mathcal{F}$ of $\mathcal{K}$ containing $\mathcal{A}\cap\mathcal{K}$
**begin**
  $\mathcal{F} \leftarrow \mathcal{K}$
  **repeat**
    1. Find hyperplane $s^\perp$ containing $\mathcal{A}$ that exposes a face of $\mathcal{F}$, i.e.,

$$\begin{array}{ll} \text{Find} & s \in \mathcal{F}^* \setminus (\operatorname{span}\mathcal{F})^\perp \\ \text{subject to} & s^\perp \supseteq \mathcal{A} \end{array} \quad (\star)$$

    2. Set $\mathcal{F} = \mathcal{F}\cap s^\perp$ if $(\star)$ is feasible
  **until** $(\star)$ *is infeasible*
**end**

| constraint on $s \in \mathcal{V}$ | set of $s$ satisfying constraint |
|---|---|
| $s \in \mathcal{F}^* \setminus \mathcal{F}^\perp$ | $\{s \in \mathcal{F}^* : \langle s, \hat{x} \rangle > 0\}$ for any $\hat{x} \in \operatorname{relint} \mathcal{F}$ |
| $s^\perp \supseteq \mathcal{A}$ where $\mathcal{A} = \{x \in \mathcal{V} : Ax = b\}$ | $\left\{A^* y : y \in \mathbb{R}^m,\ \ b^T y = 0\right\}$ |
| $s^\perp \supseteq \mathcal{A}$ where $\mathcal{A} = \{x_0 + Bz : z \in \mathbb{R}^n\}$ | $\{s \in \mathcal{V} : \langle x_0, s \rangle = 0, B^* s = 0\}$ |

**Table 1.3:** The sets satisfying individual constraints of the facial reduction algorithm (Algorithm 1.1) auxiliary problem. Lines 2-3 correspond to two possible representations of the affine set $\mathcal{A}$.

#### Solving auxiliary problems

The auxiliary problem ($\star$) is actually a cone program in disguise. That is, the feasible set is a subspace intersected with a convex cone. Table 1.3 makes this clear by giving the solution set of each individual constraint. As indicated, $\mathcal{F}^* \setminus (\operatorname{span} \mathcal{F})^\perp$ is the cone of elements that have nonzero inner product with any fixed point in the relative interior of $\mathcal{F}$ (Proposition 1.3.1). The set of $s \in \mathcal{V}$ satisfying $s^\perp \supseteq \mathcal{A}$ is the subspace $(\operatorname{span} \mathcal{A})^\perp$; see Table 1.3 for explicit descriptions of $(\operatorname{span} \mathcal{A})^\perp$ for parametric and implicit representations of $\mathcal{A}$. When $\mathcal{F}$ is a face of the nonnegative orthant, the Lorentz cone, or the psd cone, linear and semidefinite constraints describe $\mathcal{F}^*$; see Table 1.2(b) and Example 1.3.1. There are also formulas for $\mathcal{F} \cap s^\perp$ in these cases given by Table 1.2(c).

#### Singularity degree and iteration bounds

The number of iterations Algorithm 1.1 executes depends on the specific hyperplanes obtained at each iteration. The minimum number of possible iterations is called the *singularity degree* of $(\mathcal{A}, \mathcal{K})$, a parameter used in error analysis [130] and convergence analysis [48] of cone programs with feasible set $\mathcal{A} \cap \mathcal{K}$. Bounds on singularity degree are given in [86]. If $\mathcal{A} = \{x \in \mathcal{V} : Ax = b\}$, then the singularity degree is at most one if the image of $\mathcal{K}$ under $A : \mathcal{V} \to \mathbb{R}^m$ is facially exposed [47]. This image need not be facially exposed when $\mathcal{K}$ is facially exposed, e.g., when $\mathcal{K} = \mathbb{S}^n_+$. Indeed, when $\mathcal{K} = \mathbb{S}^n_+$, the singularity degree can be as large as $n - 1$; see, e.g., [132, Section 2.6].

### ■ 1.3.5 Benefits

For a linear map $A : \mathcal{V} \to \mathbb{R}^m$, $b \in \mathbb{R}^m$ and $c \in \mathcal{V}$, consider the primal-dual pair parametrized by a closed, convex cone $\mathcal{C} \subseteq \mathcal{V}$:

$$\mathbf{P}(\mathcal{C}) : \begin{array}{ll} \text{minimize} & \langle c, x \rangle \\ \text{subject to} & Ax = b, \\ & x \in \mathcal{C} \end{array} \qquad \mathbf{D}(\mathcal{C}) : \begin{array}{ll} \text{maximize} & \langle b, y \rangle \\ \text{subject to} & s = c - A^* y \\ & (s, y) \in \mathcal{C}^* \times \mathbb{R}^m. \end{array} \tag{1.9}$$

That is, for another convex cone $\mathcal{K} \subseteq \mathcal{V}$, the problems $\mathbf{P}(\mathcal{K})$ and $\mathbf{D}(\mathcal{K})$ denote (1.9) with the cones $\mathcal{C}$ and $\mathcal{C}^*$ replaced by $\mathcal{K}$ and $\mathcal{K}^*$. Finally, let $\mathcal{A}$ denote the solutions to $Ax = b$.

If we execute the facial reduction algorithm using $(\mathcal{A}, \mathcal{K})$ as input, it returns a face $\mathcal{F}$ of $\mathcal{K}$ containing $\mathcal{A} \cap \mathcal{K}$; hence, $\mathbf{P}(\mathcal{K})$ and $\mathbf{P}(\mathcal{F})$ have equal optimal values, i.e.,

$$\inf\{\langle c, x\rangle : x \in \mathcal{A} \cap \mathcal{K}\} = \inf\{\langle c, x\rangle : x \in \mathcal{A} \cap \mathcal{F}\}.$$

Further, any solution $x$ of $\mathbf{P}(\mathcal{F})$ also solves $\mathbf{P}(\mathcal{K})$. We now overview the benefits of solving $\mathbf{P}(\mathcal{F})$ as a means of solving $\mathbf{P}(\mathcal{K})$, which include dimension reduction, pathology removal, and improved conditioning.

#### Dimension reduction

The obvious benefit of facial reduction is the ability to solve the lower dimensional problem $\mathbf{P}(\mathcal{F})$ instead of $\mathbf{P}(\mathcal{K})$. Indeed, if $\mathcal{F} \neq \mathcal{K}$, then a proper subspace of span$(\mathcal{K})$ contains $\mathcal{F}$ (Proposition 1.3.2). Further, when $\mathcal{K}$ is the psd cone $\mathbb{S}^n_+$, the Lorentz cone $\mathcal{Q}^{n+1}$, or the nonnegative orthant $\mathbb{R}^n_+$, any proper face $\mathcal{F} \subset \mathcal{K}$ is isomorphic to $\mathbb{S}^d_+$ (with $d < n$), the nonnegative real line $\mathbb{R}_+$, or the nonnegative orthant $\mathbb{R}^d_+$ (with $d < n$), respectively. This in turn allows us to solve $\mathbf{P}(\mathcal{F})$ by solving a projected reformulation over the isomorphic cone (Section 1.2.5). The next example illustrates the facial reduction procedure and this reformulation.

**Example 1.3.2** (Dimension reduction)**.** Consider the semidefinite program $\mathbf{P}(\mathbb{S}^3_+)$:

$$\mathbf{P}(\mathbb{S}^3_+): \begin{array}{ll} \text{minimize} & \operatorname{Tr} CX \\ \text{subject to} & \operatorname{Tr} A_1 X = b_1, \\ & \operatorname{Tr} A_2 X = b_2, \\ & X \in \mathbb{S}^3_+, \end{array}$$

where

$$C := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}, \qquad A_1 := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad A_2 := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \qquad b := \begin{bmatrix} 4 \\ 0 \end{bmatrix}.$$

This problem is feasible, but Slater's condition fails. Indeed, the equation $\operatorname{Tr} A_2 X = 0$ is equivalent to $v^T X v = 0$ for $v = (0, 1, 1)^T$. Hence, we can reformulate this problem over a lower dimensional face using facial reduction.

By Corollary 1.3.2, a hyperplane $\hat{S}^\perp$ exists that contains all solutions of $\operatorname{Tr} A_i X_i = b_i$; further, $\hat{S}$ is a nonzero point in $\mathbb{S}^3_+$, implying the hyperplane $\hat{S}^\perp$ exposes a proper face of $\mathbb{S}^n_+$. Indeed, taking $\hat{S} = A_2$ yields a hyperplane with these properties. The face

$\mathcal{F} = \mathbb{S}^3_+ \cap \hat{S}^\perp$ exposed by $\hat{S}^\perp$ satisfies

$$\mathcal{F} = U\mathbb{S}^2_+ U^T \text{ for } U = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\sqrt{2}} \\ 0 & \frac{-1}{\sqrt{2}}. \end{bmatrix},$$

where $U \in \mathbb{R}^{3\times 2}$ was picked to satisfy $\hat{S}U = 0$. A projected reformulation of $\mathbf{P}(\mathcal{F})$ is

$$\bar{\mathbf{P}}(\mathcal{F}): \begin{array}{ll} \text{minimize} & \operatorname{Tr} \bar{C}X \\ \text{subject to} & \operatorname{Tr} \bar{A}_1 X = 4, \\ & \operatorname{Tr} \bar{A}_2 X = 0, \\ & X \in \mathbb{S}^2_+, \end{array}$$

where $\bar{C} = U^T C U$ and $\bar{A}_i = U^T A_i U$, i.e.,

$$\bar{C} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \bar{A}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \bar{A}_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

An optimal solution of this reformulation is

$$\bar{X} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \qquad \bar{y} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \bar{S} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

The point $U\bar{X}U^T$ solves $\mathbf{P}(\mathcal{F})$.

#### Pathology removal

Facial reduction can also *remove* pathologies (Section 1.2.4) if they exist. To see this, recall that either the cone program $\mathbf{P}(\mathcal{F})$ satisfies Slater's condition or a hyperplane exists that strictly separates the affine set $\mathcal{A}$ from $\mathcal{F}$ (Proposition 1.3.5). This rules out three sources of pathologies: duality gaps, unattained dual optimal values, and weak infeasibility of $\mathbf{P}(\mathcal{F})$. Indeed, if these pathologies occur for $\mathbf{P}(\mathcal{K})$, replacing $\mathcal{K}$ with $\mathcal{F}$ removes them.

**Example 1.3.3** (Removing weak infeasibility)**.** Consider the primal-dual pair of semidefinite programs

$$\mathbf{P}(\mathbb{S}^2_+): \begin{array}{ll} \text{minimize} & \operatorname{Tr} CX \\ \text{subject to} & \operatorname{Tr} A_1 X = b_1, \\ & \operatorname{Tr} A_2 X = b_2, \\ & X \in \mathbb{S}^2_+, \end{array} \qquad \mathbf{D}(\mathbb{S}^2_+): \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & S = C - \sum_{i=1}^2 y_i A_i, \\ & (S, y) \in \mathbb{S}^2_+ \times \mathbb{R}^2, \end{array}$$

where

$$C := \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \qquad A_1 := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad A_2 := \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \qquad b := \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

As demonstrated earlier in Example 1.2.3, the primal problem $\mathbf{P}(\mathbb{S}^2_+)$ is weakly infeasible, i.e., it is infeasible but $\mathbf{D}(\mathbb{S}^2_+)$ has no improving ray. We will remove this pathology using facial reduction.

Taking $\hat{S} = A_2$ yields a hyperplane $\hat{S}^\perp$ that contains the solutions to $\operatorname{Tr} A_i X = b_i$, where in addition $\hat{S} \in \mathbb{S}^2_+$. The face $\mathcal{F} = \mathbb{S}^2_+ \cap \hat{S}^\perp$ exposed by $\hat{S}^\perp$ satisfies

$$\mathcal{F} = \left\{ UXU^T : X \in \mathbb{S}^1_+ \right\} \text{ for } U = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

This yields the primal-dual pair

$$\mathbf{P}(\mathcal{F}) : \begin{array}{ll} \text{minimize} & \operatorname{Tr} CX \\ \text{subject to} & \operatorname{Tr} A_1 UXU^T = b_1, \\ & \operatorname{Tr} A_2 UXU^T = b_2, \\ & X \in \mathbb{S}^1_+, \end{array} \qquad \mathbf{D}(\mathcal{F}) : \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & S = C - \sum_{i=1}^2 y_i A_i, \\ & (U^T S U, y) \in \mathbb{S}^1_+ \times \mathbb{R}^2, \end{array}$$

where $\mathbf{P}(\mathcal{F})$ and $\mathbf{P}(\mathcal{K})$ have equal optimal values. Further, $S = -A_1$ and $y = (1, 0)$ form an improving ray for $\mathbf{D}(\mathcal{F})$, proving infeasibility of $\mathbf{P}(\mathcal{F})$ and hence of $\mathbf{P}(\mathcal{K})$.

**Example 1.3.4** (Removing duality gaps)**.** Consider the primal-dual pair of semidefinite programs

$$\mathbf{P}(\mathbb{S}^3_+) : \begin{array}{ll} \text{minimize} & \operatorname{Tr} CX \\ \text{subject to} & \operatorname{Tr} A_i X = b_i \ \ \forall i \in \{1, 2\}, \\ & X \in \mathbb{S}^3_+ \end{array} \qquad \mathbf{D}(\mathbb{S}^3_+) : \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & S = C - \sum_{i=1}^2 y_i A_i, \\ & (S, y) \in \mathbb{S}^3_+ \times \mathbb{R}^2, \end{array}$$

where

$$C := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad A_1 := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \qquad A_2 := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad b := \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

As shown by Example 1.2.2, the duality gap of this primal-dual pair is one. Both problems are feasible. Further, $X$ is primal feasible only if $\operatorname{Tr} CX = 1$, and $(y, S)$ is dual feasible only if $b^T y = 0$.

For $\hat{S} = A_2$, the hyperplane $\hat{S}^\perp$ contains $\{X \in \mathbb{S}^3 : \operatorname{Tr} XA_1 = b_1, \operatorname{Tr} A_2 X = b_2\}$.

The face $\mathcal{F} = \mathbb{S}^3_+ \cap \hat{S}^\perp$ exposed by $\hat{S}^\perp$ satisfies

$$\mathcal{F} = U\mathbb{S}^2_+ U^T \text{ for } U = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1. \end{bmatrix}.$$

The primal-dual pair reformulated over $\mathcal{F}$ and $\mathcal{F}^*$ is

$$\mathbf{P}(\mathcal{F}): \begin{array}{ll} \text{minimize} & \operatorname{Tr} CX \\ \text{subject to} & \operatorname{Tr} A_1 UXU^T = b_1, \\ & \operatorname{Tr} A_2 UXU^T = b_2, \\ & X \in \mathbb{S}^2_+, \end{array} \qquad \mathbf{D}(\mathcal{F}): \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & S = C - \sum_{i=1}^2 y_i A_i, \\ & (U^T S U, y) \in \mathbb{S}^2_+ \times \mathbb{R}^2, \end{array}$$

where a primal-dual feasible point $(X_\star, S_\star, y_\star)$ satisfying $\operatorname{Tr} CX_\star = b^T y_\star$ is

$$X_\star = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad S_\star = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}, \qquad y_\star = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

#### Improved conditioning

Facial reduction can also improve accuracy. One reason for this is simple: smaller problems can lead to more accurate floating point computation. Another reason is less obvious: an iteration of facial reduction improves the intrinsic conditioning of a cone program. Specifically, it lowers the singularity degree—improving bounds on forward error (distance to solutions) obtained from backward error (constraint violation) [130]. Note the latter error is what a solver can easily compute, whereas the former—which may be hard or impossible to compute—is the actual measure of solution quality. We illustrate these concepts with an example.

**Example 1.3.5** ([130, Example 2])**.** The following SDP has singularity degree equal to $(n-1)$—the worst case for SDP. It has decision variable $X \in \mathbb{S}^n_+$ and constraints

$$\begin{aligned} [X]_{1,1} &= 1, \\ [X]_{2,2} &= 0, \\ [X]_{k+1,k+1} &= [X]_{1,k} \qquad \forall k \in \{2, 3, \ldots, n-1\}. \end{aligned}$$

For $n = 4$, these constraints are equivalent to $\mathrm{Tr}\, A_i X = b_i$ for $b = (1, 0, 0, 0)^T$ and

$$A_1 := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad A_3 := \begin{bmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$A_2 := \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad A_4 := \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

Note that the feasible set contains a single point: the rank one matrix $e_1 e_1^T$.

Letting $\| \cdot \|_{fro}$ denote the Frobenius norm yields expressions for the forward and backward error:

$$e_{forward}(X) := \|X - e_1 e_1^T\|_{fro}, \qquad e_{backward}(X) := \min\left(0, \lambda_{min}(X)\right) + \|r(X)\|_2,$$

where $r(X) := \|(b_1 - \mathrm{Tr}\, A_1 X, \ldots, b_4 - \mathrm{Tr}\, A_4 X)\|$ and $\lambda_{min}(X)$ denotes the minimum eigenvalue of $X$. For $\epsilon = 10^{-4}$, the approximate solution $X(\epsilon)$

$$X(\epsilon) = \begin{bmatrix} 1 & -\epsilon & -\epsilon & 0 \\ -\epsilon & 0 & 0 & 0 \\ -\epsilon & 0 & \epsilon & 0 \\ 0 & 0 & 0 & \epsilon \end{bmatrix}$$

has forward error ($\sqrt{6} \cdot 10^{-4}$) that is roughly four orders of magnitude larger than its backward error ($\approx 10^{-8}$). On the other hand, using $n - 1$ iterations of facial reduction yields a reformulation

$$\text{Find } \lambda \geq 0 \text{ subject to } \lambda e_1 e_1^T = e_1 e_1^T$$

whose backward error equals its forward error for $\lambda \geq 0$:

$$e_{forward}(\lambda) = |1 - \lambda|, \qquad e_{backward}(\lambda) = |1 - \lambda| + \min(0, \lambda).$$

Hence, after facial reduction, backward error becomes a better indicator of forward error.

## ■ 1.3.6 Challenges

### Cost

The costs of solving auxiliary problems (Section 1.3.4) may overwhelm the benefit of dimension reduction. Indeed, if the given cone program is an SDP, the first auxiliary problem is also an SDP of the same order.

Fortunately, the auxiliary problem can be extremely well structured in practice. For instance, it may have solutions in a subset of the dual cone that is easy to describe (e.g., the subset of diagonal matrices when the cone is $\mathbb{S}^n_+$). Exploiting this structure will be the topic of Chapter 2.

### Dual solution recovery

Facial reduction replaces the cone $\mathcal{K}$ with a proper subset $\mathcal{F} \subset \mathcal{K}$ and the dual cone $\mathcal{K}^*$ with a superset $\mathcal{F}^* \supset \mathcal{K}^*$—i.e., it *relaxes* the dual. For this reason, solutions of the reformulated dual $\mathbf{D}(\mathcal{F})$ are not necessarily even feasible points of the original dual $\mathbf{D}(\mathcal{K})$, as the next example illustrates.

**Example 1.3.6** (Relaxation of the dual)**.** *Recall the primal-dual pair of Example 1.3.2*

$$\mathbf{P}(\mathbb{S}^3_+): \begin{array}{ll} \text{minimize} & \operatorname{Tr} CX \\ \text{subject to} & \operatorname{Tr} A_1 X = b_1, \\ & \operatorname{Tr} A_2 X = b_2, \\ & X \in \mathbb{S}^3_+, \end{array} \qquad \mathbf{D}(\mathbb{S}^3_+): \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & S = C - \sum_{i=1}^2 y_i A_i, \\ & (S, y) \in \mathbb{S}^3_+ \times \mathbb{R}^2, \end{array}$$

*with problem data*

$$C := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}, \qquad A_1 := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad A_2 := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \qquad b := \begin{bmatrix} 4 \\ 0 \end{bmatrix}.$$

*Reformulating over $\mathcal{F} = \{U\bar{X}U^T : \bar{X} \in \mathbb{S}^2_+\}$ and $\mathcal{F}^* = \{S \in \mathbb{S}^3 : U^T S U \in \mathbb{S}^2_+\}$ yields*

$$\mathbf{P}(\mathcal{F}): \begin{array}{ll} \text{minimize} & \operatorname{Tr} C\bar{X} \\ \text{subject to} & \operatorname{Tr} A_1 U\bar{X}U^T = b_1, \\ & \operatorname{Tr} A_2 U\bar{X}U^T = b_2, \\ & \bar{X} \in \mathbb{S}^2_+, \end{array} \qquad \mathbf{D}(\mathcal{F}): \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & S = C - \sum_{i=1}^2 y_i A_i, \\ & (U^T S U, y) \in \mathbb{S}^2_+ \times \mathbb{R}^2, \end{array}$$

*where*

$$U = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\sqrt{2}} \\ 0 & \frac{-1}{\sqrt{2}}. \end{bmatrix}.$$

*If $\bar{X} = 2I$, then $U\bar{X}U^T$ solves $\mathbf{P}(\mathcal{F})$ and $\mathbf{P}(\mathcal{K})$. For all $\alpha \in \mathbb{R}$, the point $y = (1, \alpha)$ solves $\mathbf{D}(\mathcal{F})$. However, this $y$ solves $\mathbf{D}(\mathbb{S}^3_+)$ only for $\alpha \leq -1$; indeed, it is infeasible otherwise.*

Unfortunately, one cannot ignore this issue and incorporate facial reduction into a primal-dual solver. Such solvers must return solutions to a given cone program and its dual; indeed, the dual may be the problem that is of actual interest to a user. Hence, these solvers need post-processing methods for dual solution recovery. For this previous example, recovery means taking a solution $(y, \alpha)$ of $\mathbf{D}(\mathcal{F})$ and decreasing $\alpha$ until it is feasible for $\mathbf{D}(\mathbb{S}^n_+)$. In other words, recovery amounts to a line search on $\alpha$. We will formalize this line search idea in Chapter 3 and study when it succeeds and when it fails. Note recovery will necessarily fail when facial reduction removes the following pathologies: duality gaps, weak infeasibility, and unattainment of the dual optimal value.

### Sparsity

Constructing maximally-sparse projected reformulations is another challenge. Indeed, for semidefinite programs, we always have a degree-of-freedom in how we parametrize the face. For instance, the faces $\mathcal{F} \subseteq \mathbb{S}^n_+$ and $\mathcal{G} \subseteq \mathbb{S}^n_+$, given by

$$\mathcal{F} = \{UXU^T : X \in \mathbb{S}^d_+\}, \qquad \mathcal{G} = \{VXV^T : X \in \mathbb{S}^d_+\},$$

are equal if $U \in \mathbb{R}^{n\times d}$ and $V \in \mathbb{R}^{n\times d}$ have the same range. To give an example, we reproduce $U$ of Example 1.3.2 and give a dense matrix $V$ with the same range below:

$$U = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\sqrt{2}} \\ 0 & \frac{-1}{\sqrt{2}}. \end{bmatrix}, \qquad V = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ -2 & -1 \end{bmatrix}.$$

If we use $V$ instead of $U$ to construct the projected reformulation of Example 1.3.2, we obtain data matrices $V^T A_i V$ and $V^T CV$ with twice as many nonzeros as $U^T A_i U$ and $U^T CU$. Picking a parameterization of $\mathcal{F}$ to minimize nonzeros is a challenge addressed in Chapter 2.

### Sensitivity

If the auxiliary problem is solved numerically, the obtained hyperplane $s^\perp$ may only contain the affine set $\mathcal{A}$ after an epsilon perturbation, i.e., the inclusion

$$(s + \epsilon)^\perp \supseteq \mathcal{A}$$

may only hold for some nonzero $\epsilon$ whose magnitude depends on the accuracy to which $s$ satisfies the constraints of the auxiliary problem. As a consequence, replacing $\mathcal{K}$ with $\mathcal{K} \cap s^\perp$ can change the optimal value of the given cone program. Given that $\mathcal{A}$ has empty intersection with the relative interior of $\mathcal{K}$, the change in optimal value can also be infinite—meaning replacing $\mathcal{K}$ with $\mathcal{K} \cap s^\perp$ can make a feasible problem infeasible. Example 4.4.2 of Chapter 4 illustrates this unfortunate phenomenon on concrete example.

### 1.3.7 Limitations

As we have seen, one benefit of facial reduction is pathology removal. Nevertheless, facial reduction iterations will occur (i.e, the auxiliary problem will have a solution) even if no pathology is present (e.g., Example 1.3.2). The basic reason for this is the following: the facial reduction algorithm terminates when a necessary condition for pathologies (failure of Slater's condition) no longer holds, as opposed to a necessary and sufficient condition.

From the dimension reduction point of view, this is a good thing: facial reduction applies to a wider family of problems. Nevertheless, it still raises the question: can one do facial reduction *only if* the given instance is pathological? This would be useful if one only wanted to confront the challenges of facial reduction (e.g., the costs of solving auxiliary problems) for pathological instances, which, without facial reduction, are perhaps unsolvable. In Chapter 4, we show the answer to this question is yes and provide an algorithm.

Note also that during its last iteration, the facial reduction algorithm obtains no information from the auxiliary problem other than an indicator to terminate. Further, if only one iteration is performed, facial reduction pays the cost of solving an auxiliary problem but makes no changes to the given instance. The Chapter 4 algorithm does not have this defect. Indeed, during its last iteration, it not only terminates, but automatically provides solutions to the reformulation.

## 1.4 Symmetry reduction

Symmetry reduction, like facial reduction, also identifies a subset of the cone $\mathcal{K}$ that contains solutions. This subset is the intersection of $\mathcal{K}$ with the range of a special projection map $P : \mathcal{V} \to \mathcal{V}$. We state the key properties of this projection in terms of the following primal-dual pair in decision variables $x \in \mathcal{V}$ and $s \in \mathcal{V}$:

$$
\begin{array}{lllll}
\text{minimize} & \langle c, x\rangle & \qquad & \text{minimize} & \langle x_0, s\rangle \\
\text{subject to} & x \in x_0 + \mathcal{L}, & & \text{subject to} & s \in c + \mathcal{L}^\perp, \\
 & x \in \mathcal{K}, & & & s \in \mathcal{K}^*,
\end{array}
$$

where, as in previous sections, $x_0 \in \mathcal{V}$ and $c \in \mathcal{V}$ are fixed and $\mathcal{L} \subseteq \mathcal{V}$ is a linear subspace with orthogonal complement $\mathcal{L}^\perp \subseteq \mathcal{V}$. Given this primal-dual pair, symmetry reduction finds a projection that satisfies the *Constraint Set Invariance Conditions*.

**Definition 1.4.1.** *A projection $P : \mathcal{V} \to \mathcal{V}$ satisfies the* Constraint Set Invariance Conditions *for $(\mathcal{K}, x_0 + \mathcal{L}, c)$ if* *(a)* $P \cdot \mathcal{K} \subseteq \mathcal{K}$ *(or, equivalently,* $P^* \cdot \mathcal{K}^* \subseteq \mathcal{K}^*$*),*

*(b)* $P \cdot (x_0 + \mathcal{L}) \subseteq x_0 + \mathcal{L}$,

*(c)* $P^* \cdot (c + \mathcal{L}^\perp) \subseteq c + \mathcal{L}^\perp$,

*where* $P^* : \mathcal{V} \to \mathcal{V}$ *is the adjoint of* $P$ *and* $P \cdot \mathcal{C} := \{Px : x \in \mathcal{C}\}$ *for any set* $\mathcal{C} \subseteq \mathcal{V}$*.*

The next proposition shows the range of $P$ contains primal solutions (if they exist) under the Constraint Set Invariance Conditions. By symmetry of these conditions, the range of $P^*$ also contains dual solutions.

**Proposition 1.4.1.** *Suppose a projection* $P : \mathcal{V} \to \mathcal{V}$ *satisfies the Constraint Set Invariance Conditions for* $(\mathcal{K}, x_0 + \mathcal{L}, c)$*. The following statements hold*

- *If* $x \in (x_0 + \mathcal{L}) \cap \mathcal{K}$*, then* $Px \in (x_0 + \mathcal{L}) \cap \mathcal{K}$*. Further,* $\langle c, x \rangle = \langle c, Px \rangle$*.*
- *If* $s \in (c + \mathcal{L}^\perp) \cap \mathcal{K}^*$*, then* $P^*s \in (c + \mathcal{L}^\perp) \cap \mathcal{K}^*$*. Further,* $\langle x_0, s \rangle = \langle x_0, P^*s \rangle$*.*

*Proof.* The statements have identical proofs. Further, the only part of the first statement not immediate is that $\langle c, x \rangle = \langle c, Px \rangle$. To see this holds, consider $x \in (x_0 + \mathcal{L})$. The conditions (b) and (c) state both $x$ and $Px$ are in $x_0 + \mathcal{L}$ and both $c$ and $P^*c$ are in $c + \mathcal{L}^\perp$; hence, (b) and (c) imply

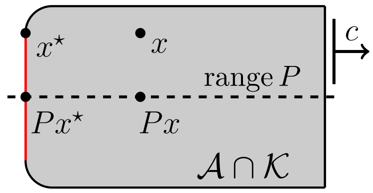
$$x - Px \in \mathcal{L}, \qquad c - P^*c \in \mathcal{L}^\perp, \tag{1.10}$$

showing $x - Px$ and $c - P^*c$ are contained in orthogonal subspaces. It follows that

$$\begin{aligned}
\langle c, x \rangle &= \langle c - P^*c + P^*c, x - Px + Px \rangle \\
&= \langle c - P^*c, Px \rangle + \langle P^*c, x \rangle && \text{Equation (1.10)} \\
&= \langle P^*c - P^*P^*c, x \rangle + \langle c, Px \rangle \\
&= \langle P^*c - P^*c, x \rangle + \langle c, Px \rangle && P^*P^* = P^* \\
&= \langle c, Px \rangle.
\end{aligned}$$

□

We now explain how symmetry reduction obtains a projection $P$ that satisfies the Constraint Set Invariance Conditions. The main idea is to find a special *group* of *automorphisms* of the primal and dual feasible sets. Note that groups are mathematical objects used to formalize notions of symmetry in engineering and physics—hence, the name symmetry reduction.




(a) $P \cdot \mathcal{K} \subseteq \mathcal{K}$

(b) $P \cdot (x_0 + \mathcal{L}) \subseteq x_0 + \mathcal{L}$,

(c) $P \cdot (c + \mathcal{L}^\perp) \subseteq c + \mathcal{L}^\perp$

**Figure 1.2:** In symmetry reduction, the image of feasible $x$ (resp., optimal $x^\star$) is feasible (resp., optimal) under a special projection map $P : \mathcal{V} \to \mathcal{V}$. This projection satisfies the conditions (a)-(c), where $\mathcal{A} := x_0 + \mathcal{L}$.

## ■ 1.4.1 Groups of linear maps

A group $\mathcal{G}$ is a set equipped with a binary operation $a \cdot b$ satisfying the following properties.

- The set $\mathcal{G}$ contains $a \cdot b$ for all $a, b \in \mathcal{G}$.
- The associative law holds, i.e., $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for all $a, b, c \in \mathcal{G}$.
- There is an identity, i.e., there exists $e \in \mathcal{G}$ for which $e \cdot a = a \cdot e = a$ for all $a \in \mathcal{G}$.
- Each $a \in \mathcal{G}$ has an inverse, i.e., there exists $a^{-1} \in \mathcal{G}$ such that $a^{-1} \cdot a = a \cdot a^{-1} = e$.

A *subgroup* $\mathcal{H} \subseteq \mathcal{G}$ is a subset that is also a group with the same operation $a \cdot b$. In other words, $\mathcal{H}$ is closed under the group operation, contains an identity, and contains inverses for all of its elements.

For an inner product space $\mathcal{V}$, the *general linear group* $GL(\mathcal{V})$ is the set of all invertible linear maps $T : \mathcal{V} \to \mathcal{V}$ equipped with composition as a group operation. Important subgroups of $GL(\mathcal{V})$ include the *orthogonal group* $O(\mathcal{V})$ and the *automorphism groups* induced by subsets of $\mathcal{V}$.

#### Automorphism groups

Let $\mathcal{X} \subseteq \mathcal{V}$ be a subset of $\mathcal{V}$. The *automorphism group of* $\mathcal{X}$, denoted $\text{aut}(\mathcal{X})$, is the subgroup

$$\text{aut}(\mathcal{X}) := \{T \in GL(\mathcal{V}) : T \cdot \mathcal{X} = \mathcal{X}\},$$

where $T \cdot \mathcal{X}$ denotes the image of $\mathcal{X}$ under $T$, i.e., $T \cdot \mathcal{X} := \{Tx : x \in \mathcal{X}\}$. It is easy to verify $\text{aut}(\mathcal{X})$ is a subgroup: it contains the identity map and it is closed under both composition and taking inverses.

#### The orthogonal group

The *orthogonal group* $O(\mathcal{V}) \subset GL(\mathcal{V})$ is the subgroup of maps whose adjoints equal their inverses, i.e.,

$$O(\mathcal{V}) := \left\{ T \in GL(\mathcal{V}) : T^{-1} = T^* \right\}.$$

It is also easy to verify $O(\mathcal{V})$ is a subgroup of $GL(\mathcal{V})$. If $T \in O(\mathcal{V})$, then so is the adjoint map given that $(T^*)^{-1} = (T^{-1})^* = (T^*)^*$. Hence all elements of $O(\mathcal{V})$ have inverses in $O(\mathcal{V})$. Further, if $T_1, T_2 \in O(\mathcal{V})$, then so is their composition $T_1T_2$ given that

$$(T_1T_2)^{-1} = T_2^{-1}T_1^{-1} = T_2^*T_1^* = (T_1T_2)^*.$$

Finally, the identity map is contained in $O(\mathcal{V})$ since it equals its own inverse and adjoint. Note that the set of *orthogonal matrices*, i.e., the set of matrices $U$ satisfying $U^TU = I$, is the orthogonal group $O(\mathbb{R}^n)$ of $\mathbb{R}^n$ under the dot product $x^Ty$.

### ■ 1.4.2 Reynolds operators and fixed-point subspaces

Let $\mathcal{G}$ be a finite subgroup of $GL(\mathcal{V})$. Associated with $\mathcal{G}$ is a linear map called the *Reynolds operator* $R_\mathcal{G} : \mathcal{V} \to \mathcal{V}$ and a set called the *fixed-point subspace* $\mathcal{V}_\mathcal{G}$:

$$R_\mathcal{G} := \frac{1}{|\mathcal{G}|} \sum_{T \in \mathcal{G}} T, \qquad \mathcal{V}_\mathcal{G} := \{x \in \mathcal{V} : Tx = x \text{ for all } T \in \mathcal{G}\}.$$

One expects $R_\mathcal{G}$ and $\mathcal{V}_\mathcal{G}$ to have special structure given their construction. It turns out that $R_\mathcal{G}$ is a projection map onto $\mathcal{V}_\mathcal{G}$, i.e., $R_\mathcal{G}$ is idempotent and has range equal to $\mathcal{V}_\mathcal{G}$:

$$R_\mathcal{G}R_\mathcal{G} = R_\mathcal{G} \quad \text{(idempotent)}, \qquad \operatorname{range} R_\mathcal{G} = \mathcal{V}_\mathcal{G}.$$

Further, if the orthogonal subgroup $O(\mathcal{V})$ contains $\mathcal{G}$, then $R_\mathcal{G}$ is *self-adjoint* ($R_\mathcal{G}^* = R_\mathcal{G}$) and hence equals the orthogonal projection map onto $\mathcal{V}_\mathcal{G}$. The next lemma restates these properties with proof. As the proof indicates, they are straightforward consequences of the group-theoretic identity

$$\mathcal{G} = \{UT : T \in \mathcal{G}\} \quad \forall U \in \mathcal{G},$$

which states that the *left coset* of $\mathcal{G}$ with respect to any $U \in \mathcal{G}$ equals $\mathcal{G}$.

**Lemma 1.4.1.** *Let $R_\mathcal{G}$ be the Reynolds operator of a finite subgroup $\mathcal{G}$ of $GL(\mathcal{V})$. Then, $R_\mathcal{G}$ is idempotent ($R_\mathcal{G}R_\mathcal{G} = R_\mathcal{G}$) and has range equal to the fixed-point subspace $\mathcal{V}_\mathcal{G}$. Further, if $\mathcal{G} \subseteq O(\mathcal{V})$, then $R_\mathcal{G}$ is self-adjoint.*

*Proof.* The map $R_{\mathcal{G}}$ is linear. Hence, it is a projection map if $R_{\mathcal{G}}R_{\mathcal{G}} = R_{\mathcal{G}}$. Indeed,

$$R_{\mathcal{G}}R_{\mathcal{G}} = \frac{1}{|\mathcal{G}|^2}\left(\sum_{U\in\mathcal{G}} U \sum_{T\in\mathcal{G}} T\right) = \frac{1}{|\mathcal{G}|^2}\left(\sum_{U\in\mathcal{G}}\sum_{T\in\mathcal{G}} UT\right) \stackrel{a}{=} \frac{1}{|\mathcal{G}|^2}\left(|\mathcal{G}|\sum_{T\in\mathcal{G}} T\right) = R_{\mathcal{G}},$$

where the equality $\stackrel{a}{=}$ follows from the coset identity $\mathcal{G} = \{UT : T \in \mathcal{G}\}$. To show that the range of $R_{\mathcal{G}}$ equals $\mathcal{V}_{\mathcal{G}}$, consider $x \in \mathcal{V}_{\mathcal{G}}$. Then,

$$R_{\mathcal{G}}x = \frac{1}{|\mathcal{G}|}\sum_{T\in\mathcal{G}} Tx = \frac{1}{|\mathcal{G}|}\sum_{T\in\mathcal{G}} x = \frac{1}{|\mathcal{G}|}|\mathcal{G}|x = x,$$

showing that the range contains $\mathcal{V}_{\mathcal{G}}$. Now consider an arbitrary point $R_{\mathcal{G}}z$ in the range. For each $U \in \mathcal{G}$ and $z \in \mathcal{V}$,

$$UR_{\mathcal{G}}z = \sum_{T\in\mathcal{G}} UTz \stackrel{b}{=} \sum_{T\in\mathcal{G}} Tz = R_{\mathcal{G}}z,$$

where $\stackrel{b}{=}$ uses the coset identity $\mathcal{G} = \{UT : T \in \mathcal{G}\}$. This shows that $R_{\mathcal{G}}z \in \mathcal{V}_{\mathcal{G}}$.

Now suppose that $\mathcal{G} \subseteq O(\mathcal{V})$. Clearly $R^*_{\mathcal{G}} = \frac{1}{|\mathcal{G}|}\sum_{T\in\mathcal{G}} T^*$. Since $\mathcal{G}$ is closed under inverses, inverses are unique, and $T^* = T^{-1}$, we must have $R^*_{\mathcal{G}} = R_{\mathcal{G}}$.

□

Consider a nonempty set $\mathcal{X} \subseteq \mathcal{V}$, a subgroup $\mathcal{G} \subseteq \text{aut}(\mathcal{X})$ of automorphisms, and the image $R_{\mathcal{G}} \cdot \mathcal{X} = \{R_{\mathcal{G}}x : x \in \mathcal{X}\}$ of $\mathcal{X}$ under $R_{\mathcal{G}}$. When does $R_{\mathcal{G}} \cdot \mathcal{X} \subseteq \mathcal{X}$ hold, i.e., when does $R_{\mathcal{G}}$ map $\mathcal{X}$ into $\mathcal{X}$? It turns out convexity of $\mathcal{X}$ implies this inclusion. Consider the following.

**Proposition 1.4.2** (Invariance Lemma)**.** *Let $\mathcal{X} \subseteq \mathcal{V}$ be nonempty, convex set and let $\mathcal{G}$ be a finite subgroup of* $\text{aut}(\mathcal{X})$*. Then, the Reynolds operator $R_{\mathcal{G}}$ leaves $\mathcal{X}$ invariant, i.e.,*

$$R_{\mathcal{G}} \cdot \mathcal{X} \subseteq \mathcal{X}.$$

*Further, $\mathcal{X} \cap \mathcal{V}_{\mathcal{G}}$ is nonempty.*

*Proof.* For all $x \in \mathcal{X}$, it holds that $R_{\mathcal{G}}x = \frac{1}{|\mathcal{G}|}\sum_{T\in\mathcal{G}} Tx$ is a convex combination of points in $\mathcal{X}$ since $T \in \text{aut}(\mathcal{X})$. But $\mathcal{X}$ is convex and hence contains any convex combination of its points. Hence, $R_{\mathcal{G}}x \in \mathcal{X}$. In addition, $\mathcal{V}_{\mathcal{G}}$ must intersect $\mathcal{X}$ since it equals the range of $R_{\mathcal{G}}$ by Lemma 1.4.1. □

The next example shows the proposition fails without the convexity assumption.

**Example 1.4.1** (Counterexample for nonconvex sets)**.** *Let $\mathcal{X}$ denote the unit sphere and let $\mathcal{G}$ denote the subgroup of* $\text{aut}(\mathcal{X})$ *consisting of the identity map and the reflection* *operator $x \mapsto -x$, i.e., let*

$$\mathcal{X} = \{x \in \mathcal{V} : \langle x, x\rangle = 1\}, \qquad \mathcal{G} = \{x \mapsto x, x \mapsto -x\}.$$

*Then $R_G x = \frac{1}{2}(x - x) = 0$ for all $x \in \mathcal{X}$ (and all $x \in \mathcal{V}$). Hence, $\mathcal{V}_\mathcal{G} = \{0\}$ and doesn't intersect $\mathcal{X}$. Note, however, that $\mathcal{V}_\mathcal{G}$ does intersect the convex hull of $\mathcal{X}$ (the unit ball).*

## ■ 1.4.3 Finding groups

As the next proposition shows, a suitable group of automorphisms $\mathcal{G} \subseteq GL(\mathcal{V})$ induces a projection $R_\mathcal{G} : \mathcal{V} \to \mathcal{V}$ that satisfies the Constraint Set Invariance Conditions.

**Proposition 1.4.3.** *Let $\mathcal{G}$ be a finite subgroup of $GL(\mathcal{V})$ with the following properties:*

- $\mathcal{G} \subseteq \operatorname{aut}(\mathcal{K})$
- $\mathcal{G} \subseteq \operatorname{aut}(x_0 + \mathcal{L})$
- $\mathcal{G} \subseteq \operatorname{aut}(\{x \in \mathcal{V} : \langle c, x\rangle \leq \alpha\})$ *for all $\alpha \in \mathbb{R}$.*

*Then, the Reynolds operator $R_\mathcal{G} : \mathcal{V} \to \mathcal{V}$ satisfies the Constraint Set Invariance Conditions (Definition 1.4.1).*

*Proof.* That $R_\mathcal{G} \cdot \mathcal{K} \subseteq \mathcal{K}$ and $R_\mathcal{G} \cdot (x_0 + \mathcal{L}) \subseteq x_0 + \mathcal{L}$ is immediate from Proposition 1.4.2. We prove the remaining inclusion $R_\mathcal{G}^* \cdot (c + \mathcal{L}^\perp) \subseteq c + \mathcal{L}^\perp$ at the end of this section; see Lemmas 1.4.2 and 1.4.3. □

Methods that construct groups satisfying these conditions exist for different families of semidefinite programs; see, e.g., [82, Lemma 2], [23, Theorem 2.1], or [39, Theorem 4]. In the next example, we illustrate construction of a group for the MAXCUT SDP relaxation of Goemans and Williamson.

**Example 1.4.2** (MAXCUT SDP relaxation)**.** Consider an undirected graph $G = ([n], E)$ with node set $[n] := \{1, \ldots, n\}$ and edge set $E \subseteq \binom{[n]}{2}$. A *cut* is a partition of nodes into two disjoint subsets $S_1$ and $S_2$. Finding a cut that maximizes the number of edges connecting $S_1$ and $S_2$ is the MAXCUT problem. Formally, we want to find $S_1 \subseteq [n]$ and $S_2 \subseteq [n]$ that maximizes the number of edges $\{i, j\} \in E$ satisfying $i \in S_1$ and $j \in S_2$. This problem is NP-hard. Nevertheless, the size of a maximum cut is famously upper bounded by the optimal value of the SDP

$$\begin{array}{ll} \text{maximize} & \frac{1}{4}\operatorname{Tr} LX \\ \text{subject to} & X \in \mathcal{A} \cap \mathbb{S}^n_+, \end{array} \tag{1.11}$$

where $\mathcal{A} := \{X \in \mathbb{S}^n : X_{ii} = 1 \text{ for all } i \in [n]\}$ and $L \in \mathbb{S}^n$ is the Laplacian matrix [62, Section 5]. (Recall the Laplacian matrix equals $D - A$, where $D$ is the diagonal matrix with $[D]_{ii}$ equal to the number of edges incident to node $i$ and $A$ is the adjacency matrix.) Further, if this SDP has a rank one solution $X$, then $X = vv^T$ for $v \in \{-1, 1\}^n$, where the sign pattern of $v$ induces a partition of $[n]$ equal to a maximum cut.

**Finding automorphisms** We can construct a group of automorphisms directly from symmetries of the graph $G$, i.e., the permutation matrices $P$ that satisfy $PLP^T = L$. Indeed, if we let $\mathcal{U}$ denote this set

$$\mathcal{U} = \left\{P \text{ a permutation matrix} : PLP^T = L\right\},$$

and define the following subgroup of $O(\mathbb{S}^n)$

$$\mathcal{G} = \left\{X \mapsto PXP^T : P \in \mathcal{U}\right\},$$

then the following inclusions hold:

$$\mathcal{G} \subseteq \operatorname{aut}(\mathbb{S}^n_+), \quad \mathcal{G} \subseteq \operatorname{aut}(\mathcal{A}), \qquad \mathcal{G} \subseteq \operatorname{aut}(\{X \in \mathcal{V} : \operatorname{Tr} LX \leq \alpha\}) \text{ for all } \alpha \in \mathbb{R}. \quad (1.12)$$

Hence, by Proposition 1.4.3, the Reynolds operator $R_{\mathcal{G}}$ satisfies the Constraint Set Invariance Conditions, implying the fixed-point subspace $\mathcal{V}_{\mathcal{G}}$ intersects the set of solutions to the SDP.

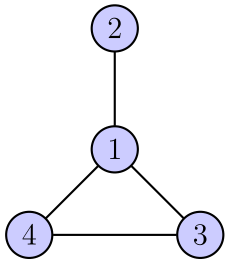
For completeness, we verify the inclusions (1.12). That $\mathcal{G} \subseteq \operatorname{aut}(\mathbb{S}^n_+)$ follows because $PXP^T \in \mathbb{S}^n_+$ when $X \in \mathbb{S}^n_+$. To see that $\mathcal{G} \subseteq \operatorname{aut}(\mathcal{A})$, first note that $\mathcal{A}$ is the set of symmetric matrices whose diagonal entries all equal one. Hence, $\mathcal{A}$ is invariant under simultaneous permutation of rows and columns; in other words, if $X \in \mathcal{A}$ then $PXP^T \in \mathcal{A}$ for any permutation matrix $P$. Finally, that $\mathcal{G} \subseteq \operatorname{aut}(\{X \in \mathcal{V} : \operatorname{Tr} LX \leq \alpha\}$ follows essentially by definition of $\mathcal{U}$: for all $P \in \mathcal{U}$,

$$\operatorname{Tr} LPXP^T = \operatorname{Tr} P^T LPX = \operatorname{Tr} LX,$$

where the first equality uses the cyclic property of trace $\operatorname{Tr} ABC = \operatorname{Tr} CAB$ and the second equality uses the fact that $P^T \in \mathcal{U}$ when $P \in \mathcal{U}$.

**Concrete instance** Figure 1.3 gives a concrete example of a graph, a maximum cut, and the group $\mathcal{U} \subset O(\mathbb{R}^4)$. For the group $\mathcal{G} \subset O(\mathbb{S}^4)$ induced by $\mathcal{U}$, the following SDP

$$\begin{array}{ll} \text{maximize} & \frac{1}{4} \operatorname{Tr} LX \\ \text{subject to} & X \in \mathcal{A} \cap \mathbb{S}^4_+ \cap \mathcal{V}_{\mathcal{G}}, \end{array}$$




$S_1 = \{1\}, S_2 = \{2, 3, 4\}$

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix},$$

$$\mathcal{U} = \left\{ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \right\}.$$

**Figure 1.3:** A graph, a maximum cut $\{S_1, S_2\}$, the adjacency matrix $A$, the degree matrix $D$, and the permutation matrices $\mathcal{U}$ commuting with the Laplacian $L = D - A$.

has the same optimal value as the Goemans and Williamson SDP (1.11), where the fixed-point subspace $\mathcal{V}_{\mathcal{G}}$ satisfies

$$\mathcal{V}_{\mathcal{G}} = \left\{ \begin{pmatrix} t_1 & t_5 & t_6 & t_6 \\ t_5 & t_2 & t_4 & t_4 \\ t_6 & t_4 & t_3 & t_7 \\ t_6 & t_4 & t_7 & t_3 \end{pmatrix} : t \in \mathbb{R}^7 \right\}.$$

Section 1.4.4 shows that $\mathcal{V}_{\mathcal{G}}$ also has a canonical direct-sum decomposition. Using this decomposition allows one to write $\mathbb{S}^n_+ \cap \mathcal{V}_{\mathcal{G}}$ as a product of smaller cones; see Example 1.4.3.

#### Proof of Proposition 1.4.3

We only need to show that the Reynolds operator satisfies $R^*_{\mathcal{G}} \cdot (c + \mathcal{L}^\perp) \subseteq c + \mathcal{L}^\perp$. For this, we first show that $R^*_{\mathcal{G}} c = c$, which is immediate from the following lemma.

**Lemma 1.4.2.** *Let $T : \mathcal{V} \to \mathcal{V}$ be a linear map with adjoint $T^* : \mathcal{V} \to \mathcal{V}$. For $c \in \mathcal{V}$ and $\alpha \in \mathbb{R}$, let $\Omega_\alpha = \{x : \langle c, x \rangle \leq \alpha\}$. The following are equivalent*

- $T^* c = c$
- $T \cdot \Omega_\alpha \subseteq \Omega_\alpha$ *for all* $\alpha \in \mathbb{R}$

*Proof.* That the first statement implies the second is immediate since

$$\langle c, Tx \rangle = \langle T^* c, x \rangle = \langle c, x \rangle.$$

Suppose the second statement holds. Then, $\langle c, y \rangle \geq \langle c, Ty \rangle$ and $\langle c, -y \rangle \geq \langle c, -Ty \rangle$ for all $y \in \mathcal{V}$. Hence, $\langle c, y \rangle = \langle c, Ty \rangle$ for all $y \in \mathcal{V}$. The claim follows by applying this to the identity

$$\langle c - T^*c, c - T^*c\rangle = \langle c, c\rangle + \langle T^*c, T^*c\rangle - 2\langle c, T^*c\rangle.$$

Indeed, $\langle T^*c, T^*c\rangle = \langle c, TT^*c\rangle = \langle c, T^*c\rangle$ and $\langle c, c\rangle = \langle c, Tc\rangle = \langle T^*c, c\rangle$. Hence,

$$\langle c - T^*c, c - T^*c\rangle = 0.$$

Hence, $c = T^*c$. □

Next we show that $R_{\mathcal{G}}^* \cdot \mathcal{L}^\perp \subseteq \mathcal{L}^\perp$:

**Lemma 1.4.3.** *Let $P : \mathcal{V} \to \mathcal{V}$ be a projection map and $x_0 + \mathcal{L}$ an affine set. If $P \cdot (x_0 + \mathcal{L}) \subseteq x_0 + \mathcal{L}$, then $P \cdot \mathcal{L} \subseteq \mathcal{L}$ and $P^* \cdot (\mathcal{L}^\perp) \subseteq \mathcal{L}^\perp$.*

*Proof.* If $P \cdot (x_0 + \mathcal{L}) \subseteq x_0 + \mathcal{L}$, then $x_0 + \mathcal{L} = Px_0 + \mathcal{L}$. Further, for all $x_L \in \mathcal{L}$, it holds that

$$Px_0 + x_L - (PPx_0 + Px_L) \in \mathcal{L}.$$

Since $PP = P$, this shows that $x_L - Px_L \in \mathcal{L}$, implying that $Px_L \in \mathcal{L}$. Hence, $P \cdot \mathcal{L} \subseteq \mathcal{L}$, which in turn implies that $P^* \cdot (\mathcal{L}^\perp) \subseteq \mathcal{L}^\perp$. □

Putting everything together, we conclude that

$$R_{\mathcal{G}}^* \cdot (c + \mathcal{L}^\perp) = c + R_{\mathcal{G}}^* \cdot \mathcal{L}^\perp \subseteq c + \mathcal{L}^\perp,$$

as desired.

### 1.4.4 Structure of fixed-point subspaces

We now study the fixed-point subspace $\mathcal{V}_{\mathcal{G}} \subseteq \mathcal{V}$ in more detail, shifting focus specifically to semidefinite programming. Going forward, we make the following assumption.

**Assumption 1.4.1.** *$\mathcal{V} = \mathbb{S}^n$ and, for a group of orthogonal matrices $\mathcal{U} \subseteq O(\mathbb{R}^n)$,*

$$\mathcal{G} = \{X \mapsto UXU^T : U \in \mathcal{U}\}.$$

This assumption implies existence of an injective linear map $\Phi$ for which

$$\mathbb{S}_+^n \cap \mathcal{V}_{\mathcal{G}} = \Phi \cdot (\mathcal{C}_1 \times \cdots \times \mathcal{C}_r), \tag{1.13}$$

where $\mathcal{C}_i$ is the cone of Hermitian psd matrices of order $n_i$ with real, complex, or quaternion entries, and $\sum_{i=1}^r n_i \leq n$. This allows one to construct a projected reformulation over $\mathcal{C}_1 \times \cdots \times \mathcal{C}_r$ if one can find $\Phi$ and $\mathcal{C}_i$.

**Example 1.4.3** (Continuation of the MAXCUT example)**.** *The fixed-point subspace $\mathcal{V}_{\mathcal{G}}$ of Example 1.4.2 satisfies*

$$\mathbb{S}^4_+ \cap \mathcal{V}_{\mathcal{G}} = \left\{ Q_1 X_1 Q_1^T + Q_2 X_2 Q_2^T : X_1 \in \mathbb{S}^3_+, X_2 \in \mathbb{S}^1_+ \right\},$$

*where*

$$Q_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \qquad Q_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}.$$

*This yields a projected reformulation of the Goemans and Williamson SDP:*

$$\begin{array}{ll} \text{maximize} & \frac{1}{4}\operatorname{Tr}(Q_1^T L Q_1) X_1 + \frac{1}{4}\operatorname{Tr}(Q_2^T L Q_2) X_2 \\ \text{subject to} & \operatorname{Tr}(Q_1^T E_{ii} Q_1) X_1 + \operatorname{Tr}(Q_2^T E_{ii} Q_2) X_2 = 1 \qquad \forall i \in \{1,2,3,4\}, \\ & X_1 \in \mathbb{S}^3_+, X_2 \in \mathbb{S}^1_+. \end{array}$$

One finds $\Phi$ and the cones $\mathcal{C}_i$ by first finding a canonical direct-sum decomposition of $\mathcal{V}_{\mathcal{G}}$. This decomposition arises from the following observation.

**Proposition 1.4.4.** *Suppose Assumption 1.4.1 holds. Then, $\mathcal{V}_{\mathcal{G}}$ equals the subspace of symmetric matrices that commute with all $U \in \mathcal{U}$ under ordinary matrix multiplication, i.e.,*

$$\mathcal{V}_{\mathcal{G}} = \{X \in \mathbb{S}^n : UX = XU \ \ \forall U \in \mathcal{U}\}.$$

*Proof.* To begin, suppose $X$ commutes with each $U$. Then,

$$R_{\mathcal{G}} X = \frac{1}{|\mathcal{U}|} \sum_{U \in \mathcal{U}} U X U^T \overset{a}{=} \frac{1}{|\mathcal{U}|} \sum_{U \in \mathcal{U}} X U U^T \overset{b}{=} \frac{1}{|\mathcal{U}|} \sum_{U \in \mathcal{U}} X = X$$

where the equality $\overset{a}{=}$ uses the assumption $UX = XU$ and $\overset{b}{=}$ the fact $UU^T = I$ given that $U$ is an orthogonal matrix. Hence, $X$ is in the range of $R_{\mathcal{G}}$, which equals $\mathcal{V}_{\mathcal{G}}$ (Lemma 1.4.1). Next suppose that $X \in \mathcal{V}_{\mathcal{G}}$, i.e., that

$$X = UXU^T \text{ for all } U \in \mathcal{U}.$$

Then multiplying both sides by $U^T$ shows

$$U^T X = U^T U X U^T = X U^T,$$

showing that $X$ commutes with $U^T$. Since $U \in \mathcal{U}$ implies $U^T \in \mathcal{U}$ (given that $\mathcal{U}$ is a group), we conclude that $X$ also commutes with $U$.

□

An *invariant subspace* of $\mathcal{T} \subseteq \mathbb{R}^{n\times n}$ is a subspace of $\mathbb{R}^n$ that contains its image under $U$ for all $U \in \mathcal{T}$. An invariant subspace is *minimal* if it is nonzero and properly contains no nonzero invariant subspace. The next lemma shows the minimal invariant subspaces of $\mathcal{V}_{\mathcal{G}} \cup \mathcal{U}$ induce a direct-sum decomposition of $\mathcal{V}_{\mathcal{G}}$, owing in part to the commuting relationship established by Proposition 1.4.4.

**Lemma 1.4.4.** *Let $\mathcal{U} \subseteq \mathbb{R}^{n\times n}$ be a finite set closed under transposition and let*

$$\mathcal{U}' = \{X \in \mathbb{S}^n : UX = XU \ \ \forall U \in \mathcal{U}\}.$$

*Then, the minimal invariant subspaces $\mathcal{S}_i \subseteq \mathbb{R}^n$ of $\mathcal{U} \cup \mathcal{U}'$ form an orthogonal direct-sum decomposition of $\mathbb{R}^n$, i.e., $\mathbb{R}^n = \oplus_{i=1}^r \mathcal{S}_i$. Further, $\mathcal{U}' = \bigoplus_{i=1}^r \mathcal{U}'_i$, where*

$$\mathcal{U}'_i := \{X \in \mathcal{U}' : \operatorname{range} X \subseteq \mathcal{S}_i\}.$$

*Proof.* Let $\mathcal{W} \subseteq \mathbb{R}^n$ be an invariant subspace that is not minimal. Then, it contains an invariant subspace $\mathcal{S}$. We claim the orthogonal complement $\mathcal{S}^\perp \cap \mathcal{W}$ of $\mathcal{S}$ in $\mathcal{W}$ is invariant. Indeed, for all $x \in \mathcal{S}$, $y \in \mathcal{S}^\perp \cap \mathcal{W}$ and $U \in \mathcal{U} \cup \mathcal{U}'$, it holds that $0 = y^T U x = x^T(U^T y)$. Hence, $U^T y \in \mathcal{S}^\perp$. Further, $U^T y \in \mathcal{W}$ since $\mathcal{W}$ is invariant. Hence, $\mathcal{S}^\perp \cap \mathcal{W}$ is invariant. Taking $\mathcal{W} = \mathbb{R}^n$, it follows we can iteratively split $\mathbb{R}^n$ into orthogonal invariant subspaces, terminating when each subspace is minimal. This yields the orthogonal decomposition $\oplus_{i=1}^r \mathcal{S}_i$.

We now show $\mathcal{S}_1, \ldots, \mathcal{S}_r$ are the only minimal invariant subspaces. For this, we first establish that the orthogonal projection matrix $E$ onto an invariant subspace $\mathcal{S}$ commutes with all $U \in \mathcal{U} \cup \mathcal{U}'$. Since $\mathcal{S}^\perp$ is also invariant (as just established), we have that $EU(I-E) = 0$ and $(I-E)UE = 0$, showing that $EU = EUE$ and $UE = EUE$; hence, $EU = UE$. Note that this also shows $E \in \mathcal{U}'$. Now, let $F$ denote the projection onto a minimal invariant subspace and let $E_i$ denote the projection matrix onto $\mathcal{S}_i$. Since $E_i \in \mathcal{U}'$, we have that $FE_i = E_iF$, showing that $FE_i$ is the projection onto the intersection of $\operatorname{range} F \cap \operatorname{range} E_i$. If $F$ and $E_i$ are distinct, we have by minimality that $\operatorname{range} F \cap \operatorname{range} E_i = 0$, showing that $FE_i = 0$, which implies $\operatorname{range} F$ and $\operatorname{range} E_i$ are orthogonal subspaces. It follows that $F = E_j$ for some $j$—otherwise, $F$ is orthogonal to $I = \sum_{i=1}^r E_i$ and hence equals the zero matrix.

We now show the desired decomposition of $\mathcal{U}'$ holds. To begin, if $X \in \mathcal{U}'$, then

$$X \overset{a}{=} \sum_{i=1}^r E_i X \sum_{j=1}^r E_j = \sum_{i=1}^r \sum_{j=1}^r E_i X E_j \overset{b}{=} \sum_{i=1}^r E_i X E_i,$$

where the equality $\overset{a}{=}$ follows given that $\sum_{i=1}^r E_i = I$ and the equality $\overset{b}{=}$ follows given that $E_i X E_j = E_i E_j X$, where $E_i E_j = 0$ if $i \neq j$. Finally, $E_i X E_i \in \mathcal{U}'$ since $X, E_i \in \mathcal{U}'$. $\square$

Note that this lemma states we can put $\mathcal{V}_\mathcal{G}$ into block-diagonal form. To be precise, the inclusion

$$\{X \in \mathcal{U}' : \operatorname{range} X \subseteq \mathcal{S}_i\} \subseteq \{X \in \mathbb{S}^n : \operatorname{range} X \subseteq \mathcal{S}_i\} \tag{1.14}$$

implies the subspace $\oplus_{i=1}^r \mathbb{S}^{n_i}$ (where $n_i$ equals the dimension of $\mathcal{S}_i$) contains $\mathcal{U}'$ up-to congruence transformation by some $(Q_1, \ldots, Q_r) \in \mathbb{R}^{n \times n}$. Specifically, one can pick any $Q_i \in \mathbb{R}^{n \times n_i}$ with range equal to $\mathcal{S}_i$. Further, if the inclusion (1.14) holds with equality for each $\mathcal{S}_i$, then

$$\mathcal{V}_\mathcal{G} \cap \mathbb{S}^n_+ = \left\{ \sum_{i=1}^r Q_i X_i Q_i^T : X_i \in \mathbb{S}^{n_i}_+ \right\},$$

i.e., we obtain a decomposition of $\mathcal{V}_\mathcal{G} \cap \mathbb{S}^n_+$ as an injective transformation of $\mathbb{S}^{n_1}_+ \times \cdots \times \mathbb{S}^{n_r}_+$.

**Example 1.4.4.** *We can reinterpret the decomposition of Example 1.4.3 in terms of invariant subspaces of $\mathcal{U}$. The two minimal invariant subspaces $S_1$ and $S_2$ of $\mathcal{U}$ equal the range of $Q_1$ and $Q_2$, respectively:*

$$\mathcal{U} = \left\{ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \right\}, \qquad Q_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \qquad Q_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}.$$

*Each direct-summand of $\mathcal{V}_\mathcal{G}$ satisfies $\mathcal{V}_{\mathcal{G},i} = \{X \in \mathbb{S}^n : \operatorname{range} X \subseteq \mathcal{S}_i\}$.*

In general, the inclusion

$$\mathcal{V}_\mathcal{G} \cap \mathbb{S}^n_+ \subseteq \left\{ \sum_{i=1}^r Q_i X_i Q_i^T : X_i \in \mathbb{S}^{n_i}_+ \right\}$$

won't hold with equality unless we impose additional constraints on the cones $\mathbb{S}^{n_i}_+$. Nevertheless, these constraints are well understood: $\{Q_i^T X Q_i : X \in \mathcal{V}_\mathcal{G} \cap \mathbb{S}^n_+\}$ must be isomorphic to a cone of psd matrices with real, complex, or quaternion entries. Unfortunately, finding the isomorphism and specific cone is quite technical. We forgo further explanation here, but devote Chapter 6 to this topic; indeed, this chapter concerns a more general class of isomorphisms and cones. Authors [89, 42] have also addressed this topic using the theory of matrix *-algebras, which we review next.

## 1.5 Reductions via *-algebras and completely positive projections

This section continues Section 1.4, giving an alternative way of satisfying the Constraint Set Invariance Conditions without using group theory. Recall under Assumption 1.4.1 the group $\mathcal{G}$ was of the form $\{X \mapsto UXU^T : U \in \mathcal{U}\}$ for some finite group $\mathcal{U} \subset O(\mathbb{R}^n)$ of orthogonal matrices. Under this assumption, the Reynolds operator was, by definition, a sum of congruence transformations $R_{\mathcal{G}}(X) = \frac{1}{|\mathcal{U}|}\sum_{U\in\mathcal{U}} UXU^T$. For an arbitrary finite subset $\mathcal{U} \subset \mathbb{R}^{n\times n}$, one calls the map $\sum_{U\in\mathcal{U}} UXU^T$ *completely positive*. This section reviews techniques that find completely positive projections satisfying the Constraint Set Invariance Conditions. To do this, we first establish the correspondence between completely positive projections and *associative *-algebras*.

Going forward, it is convenient to view $X \mapsto \sum_{U\in\mathcal{U}} UXU^T$ not as a map on $\mathbb{S}^n$ but a map on $\mathbb{R}^{n\times n}$. We also equip $\mathbb{R}^{n\times n}$ with the trace inner product $\langle X, Y\rangle := \operatorname{Tr} X^T Y$ and view it as an associative *-algebra with product given by matrix multiplication and *-involution given by transposition. Finally, since we have restricted to a vector space $\mathbb{R}^{n\times n}$ of square matrices, we will use $\Psi$ to denote linear maps instead of capital letters to avoid confusion with matrices.

## ■ 1.5.1 Completely positive projections, commutants, and *-subalgebras

This section proves the following characterization of completely positive projections, establishing their connection to *-algebras.

**Proposition 1.5.1.** *Let $\Psi : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ be an orthogonal projection map. Consider the following statements.*

1. *The map $\Psi$ is completely positive, i.e., there exists a finite set $\mathcal{U} \subset \mathbb{R}^{n\times n}$ for which $\Psi(X) = \sum_{U\in\mathcal{U}} U^T XU$ for all $X \in \mathbb{R}^{n\times n}$.*

2. *The range of $\Psi$ is a commutant, i.e., there exists a finite set $\mathcal{U} \subset \mathbb{R}^{n\times n}$ for which*
$$\operatorname{range}\Psi = \left\{X \in \mathbb{R}^{n\times n} : XU = UX,\ \ XU^T = U^T X\ \ \forall U \in \mathcal{U}\right\}.$$

3. *The range of $\Psi$ is a *-subalgebra of $\mathbb{R}^{n\times n}$, i.e., it is closed under matrix multiplication and transposition.*

*Then,* $(2 \Rightarrow 3)$ *and* $(3 \Rightarrow 1)$*. If $\Phi(I) = I$, these statements are equivalent.*

The equivalence of statements (2) and (3) (under the assumption $\Psi(I) = I$) is called the *bicommutant theorem*; see, e.g., [7, Theorem 1.2.1 ] for a direct proof. The implication $(3 \Rightarrow 1)$, i.e., that the projection onto a *-subalgebra is completely positive is a well known result for *complex* *-algebras; see, e.g., [126, Theorems 2.2.6 and 2.2.2] and [143, Section 6].

To prove this proposition, we first state needed results on completely positive maps.

**Completely-positive maps and Choi matrices**

Completely positivity of a map $\Psi : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ relates to its *Choi matrix* $C_\Psi \in \mathbb{R}^{n^2\times n^2}$, the block matrix defined via

$$C_\Psi := \sum_{i,j=1}^{n} E_{ij} \otimes \Psi(E_{ij}),$$

where $\otimes$ denotes the Kronecker product and $E_{ij} \in \mathbb{R}^{n\times n}$ is the 0/1 matrix nonzero only at the $(i,j)^{th}$ entry. Note that the $(i,j)^{th}$ block of $C_\Psi$ is just the evaluation of $\Psi$ at $E_{ij}$. The next theorem of Choi states $\Psi$ is completely positive if and only if $C_\Psi$ is symmetric and positive semidefinite.

**Proposition 1.5.2** (Choi's Theorem, [33]). *Let $\Psi : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ be a linear map. The following statements are equivalent.*

1. *The Choi matrix $C_\Psi := \sum_{i,j=1}^n E_{ij} \otimes \Psi(E_{ij})$ is symmetric and psd.*
2. *The map $\Psi$ is completely positive, i.e., there exists a finite set $\mathcal{U} \subset \mathbb{R}^{n\times n}$ for which $\Psi(X) = \sum_{U\in\mathcal{U}} UXU^T$ for all $X \in \mathbb{R}^{n\times n}$.*

Note that Choi proves this proposition with $\mathbb{R}^{n\times n}$ replaced by the set of $n \times n$ complex matrices and the transpose operator replaced by conjugate transposition. A simple modification of Choi's argument proves it as stated here; see also [68, p. 415]. Choi's proof is also constructive: the matrices $U$ are eigenvectors of $C_\Psi$ reshaped into $n \times n$ matrices.

To use Choi's theorem, we need a few facts about the Choi matrices of orthogonal projections.

**Lemma 1.5.1.** *Let $\Psi : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ be an orthogonal projection whose range is a *-subalgebra, i.e., whose range is closed under matrix multiplication and transposition. The following statements hold.*

- *The Choi matrix $C_\Psi := \sum_{i,j=1}^n E_{ij} \otimes \Psi(E_{ij})$ is symmetric.*
- *If $F = \sum_{i,j=1}^n E_{ij} \otimes F_{ij}$ is the orthogonal projection matrix onto an eigenspace of $C_\Psi$ not equal to the kernel, then $\Psi(F_{ij}) = F_{ij}$.*

*Proof.* The matrix $C_\Psi$ is symmetric if $\Psi(E_{ij}) = \Psi(E_{ji})$. To see this holds, let $\mathcal{B}$ be an orthonormal basis for the range. Then, for all $X \in \mathbb{R}^{n\times n}$,

$$(\Psi(X))^T = \left(\sum_{B\in\mathcal{B}} (\operatorname{Tr} B^T X)B\right)^T = \sum_{B\in\mathcal{B}} (\operatorname{Tr} B^T X)B^T = \sum_{B\in\mathcal{B}} (\operatorname{Tr} XB^T)B^T = \Psi(X^T),$$

where the last equality follows since $\{B^T : B \in \mathcal{B}\}$ is also an orthonormal basis since the range is closed under transposition. Hence, $\Psi(E_{ij}) = \Psi(E_{ji})$ as desired.

For the next statement, we'll show the following subspace is closed under $X \mapsto X^2$:

$$\left\{ \sum_{i,j=1}^{n} E_{ij} \otimes \Psi(X_{ij}) : X_{ij} \in \mathbb{R}^{n\times n} \right\} \cap \mathbb{S}^n.$$

The statement then follows by Lemma 5.2.1. Suppose that $Z = \sum_{i,j=1}^n E_{ij} \otimes \Psi(X_{ij})$. Then

$$\begin{aligned} Z^2 &= \left(\sum_{i,j=1}^{n} E_{ij} \otimes \Psi(X_{ij})\right)\left(\sum_{k,\ell=1}^{n} E_{k\ell} \otimes \Psi(X_{k\ell})\right) \\ &= \sum_{i,j=1}^{n}\sum_{k,\ell=1}^{n} E_{ij}E_{k\ell} \otimes \Psi(X_{ij})\Psi(X_{k\ell}) \\ &= \sum_{i,j=1}^{n} E_{ij} \otimes \Psi(Z_{ij}), \end{aligned}$$

for some $Z_{ij} \in \mathbb{R}^{n\times n}$ given that $E_{ij}E_{k\ell} \in \operatorname{span} E_{mn}$ and $\Psi(X_{ij})\Psi(X_{k\ell}) \in \operatorname{range} \Psi$. Further, if $Z$ is symmetric, then so is $Z^2$. □

We are now ready to prove Proposition 1.5.1.

#### Proof of Proposition 1.5.1

The implication $(2 \Rightarrow 3)$ is obvious. To show that $(3 \Rightarrow 1)$, it suffices to show that the Choi matrix $C_\Psi$ of $\Psi$ is psd. Towards this, suppose that $F = \sum_{i,j=1}^n E_{ij} \otimes F_{ij}$ is a symmetric projection matrix with range equal to an eigenspace of $C_\Psi$. Suppose also that the associated eigenvalue $\lambda_F$ is nonzero. By Lemma 1.5.1, it holds that $\Psi(F_{ij}) = F_{ij}$. Hence,

$$\lambda_F\langle F, F\rangle = \langle F, C_\Psi\rangle = \sum_{i,j=1}^{n} \langle \Psi(F_{ij}), E_{ij}\rangle = \sum_{i,j=1}^{n} \langle F_{ij}, E_{ij}\rangle = \langle F, \sum_{i,j=1}^{n} E_{ij} \times E_{ij}\rangle.$$

But $\langle F, \sum_{i,j=1}^n E_{ij} \times E_{ij}\rangle \geq 0$ since it is the trace inner product between two psd matrices. Hence, $\lambda_F\langle F, F\rangle \geq 0$, which implies $\lambda_F \geq 0$.

We now show $(1 \Rightarrow 2)$ under the assumption $\Psi(I) = I$. That the range contains the commutant is easy to check. For the reverse, we use the argument sketched in [143, Section 6]. To begin, for $X \in \mathbb{R}^{n\times n}$, define the psd matrix $T \in \mathbb{S}^n$ as follows:

$$\begin{aligned} T &:= \frac{1}{2}\left(\sum_{U\in\mathcal{U}}(UX - XU)^T(UX - XU) + (U^TX - XU^T)^T(U^TX - XU^T)\right) \\ &= X^T\Psi(I)X - X^T\Psi(X) - \Psi(X^T)X + \Psi(X^TX) \\ &= X^TX - X^T\Psi(X) - \Psi(X)^TX + \Psi(X^TX). \end{aligned}$$

The second line uses the fact that $\Psi = \sum_{U\in\mathcal{U}} U^TXU = \sum_{U\in\mathcal{U}} UXU^T$ since $\Psi$ is self-adjoint. The third line uses the fact that $\Psi(X^T) = \Psi(X)^T$, as is easily checked. For all $X \in \operatorname{range}\Psi$, it in addition holds that

$$T = \Psi(X^TX) - X^TX.$$

We will show $T = 0$ for all $X \in \operatorname{range}\Psi$, which in turn implies that $UX - XU = 0$ and $U^TX - XU^T = 0$.

To see that $T = 0$, note that $\Psi$, being completely positive, satisfies the Schwartz inequality:

$$\Psi(X^TX) - \Psi(X)^T\Psi(X) \in \mathbb{S}^n_+.$$

Hence, if $\Psi(X) = X$, then $\Psi(X^TX) - X^TX \in \mathbb{S}^n_+$, i.e., $T$ is psd. On the other hand, since $\Psi$ is idempotent,

$$\Psi(T) = \Psi\left(\Psi(X^TX) - X^TX\right) = \Psi(X^TX) - \Psi(X^TX) = 0.$$

This shows that the trace of $T$ vanishes:

$$\operatorname{Tr} T = \langle I, T\rangle = \langle \Psi(I), T\rangle = \langle I, \Psi(T)\rangle = 0.$$

Since $T$ is psd, we conclude $T = 0$.

## ■ 1.5.2 Constraint set invariance via *-algebras

The characterization of completely positive projections in terms of *-subalgebras yields approaches for satisfying the Constraint Set Invariance Conditions. Such approaches were pioneered by Schrijver [123]; see also the survey [37]. We overview two approaches.

#### Algebras generated by constraint sets

For fixed $X_0, C \in \mathbb{R}^{n\times n}$ and a linear subspace $\mathcal{L} \subseteq \mathbb{R}^{n\times n}$, consider the following semidefinite program

$$\begin{array}{ll} \text{minimize} & \operatorname{Tr} CX \\ \text{subject to} & X \in X_0 + \mathcal{L}, \\ & X \in \mathbb{S}^n_+, \end{array} \tag{1.15}$$

where the psd cone $\mathbb{S}^n_+ \subseteq \mathbb{R}^{n\times n}$ is viewed as a subset of $\mathbb{R}^{n\times n}$. Our goal is to find an orthogonal projection $\Psi : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ satisfying the Constraint Set Invariance Conditions which, for this SDP, are

$$\Psi \cdot \mathbb{S}^n_+ \subseteq \mathbb{S}^n_+, \qquad \Psi \cdot (X_0 + \mathcal{L}) \subseteq X_0 + \mathcal{L} \qquad \Psi \cdot (C + \mathcal{L}^\perp) \subseteq C + \mathcal{L}^\perp.$$

The condition $\Psi \cdot \mathbb{S}^n_+ \subseteq \mathbb{S}^n_+$ holds automatically if the range is a *-subalgebra since, in this case, $\Psi$ is completely positive. The remaining conditions hold if the range also contains $X_0 + \mathcal{L}$ or $C + \mathcal{L}^\perp$.

**Proposition 1.5.3.** *The orthogonal projections onto the following subspaces satisfy the Constraint Set Invariance Conditions for the SDP* (1.15)*:*

- *Any *-subalgebra containing $X_0$ and $\mathcal{L}$*
- *Any *-subalgebra containing $C$ and $\mathcal{L}^\perp$*

*Proof.* The statements have identical proofs. We only show the first. To begin, the range contains $X_0$ and $\mathcal{L}$. Hence $\Psi \cdot (X_0 + \mathcal{L}) = X_0 + \mathcal{L}$. Indeed, since $\Psi(X_0) = X_0$, we must have that $\Psi \cdot \mathcal{L} = \mathcal{L}$. This implies that $\Psi \cdot \mathcal{L}^\perp \subseteq \mathcal{L}^\perp$. Further, since

$$C + \mathcal{L}^\perp = C_{\mathcal{L}} + \mathcal{L}^\perp,$$

where $C_{\mathcal{L}}$ denotes the orthogonal projection of $C$ onto $\mathcal{L}$, it follows that

$$\Psi \cdot (C + \mathcal{L}^\perp) = C_{\mathcal{L}} + \Psi \cdot \mathcal{L}^\perp \subseteq C + \mathcal{L}^\perp.$$

The claim therefore follows since the projection onto any *-subalgebra is completely positive (Proposition 1.5.1). □

Note if $X_0 + \mathcal{L}$ denotes the solutions to a set of linear equations $\operatorname{Tr} A_i X_i = b_i$, then a *-subalgebra (indeed, any subspace) contains $C$ and $\mathcal{L}^\perp$ if and only if it contains the data matrices $\{C, A_1, A_2, \ldots, A_m\}$. This shows *generating* an algebra from the SDP data matrices yields a projection that satisfies the Constraint Set Invariance Conditions; see [37] for more information on this approach.

#### Application specific approaches via coherent algebras

One can also identify *-subalgebras using application specific techniques. We illustrate this for the MAXCUT relaxation by finding a *-subalgebra that is *coherent*. By definition coherent *-subalgebras contain the identity matrix, the all-ones-matrix, and have an orthogonal basis of 0/1 matrices [70]. The following *-subalgebra is coherent:

$$\mathbf{A} := \left\{ \begin{pmatrix} t_1 & t_6 & t_9 & t_9 \\ t_7 & t_2 & t_4 & t_4 \\ t_8 & t_5 & t_3 & t_{10} \\ t_8 & t_5 & t_{10} & t_3 \end{pmatrix} : t \in \mathbb{R}^{10} \right\}. \tag{1.16}$$

To find a coherent *-subalgebra for the MAXCUT relaxation, we need the following property:

**Lemma 1.5.2.** *Let $\Psi : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ be an orthogonal projection onto a coherent *-subalgebra. Then, the set of diagonal matrices and its orthogonal complement are invariant subspaces of $\Psi$.*

**Example 1.5.1** (*-algebra reductions of MAXCUT)**.** For an undirected graph $G$ with Laplacian matrix $L \in \mathbb{S}^n$, recall the MAXCUT relaxation

$$\begin{aligned} \text{maximize} \ & \frac{1}{4} \operatorname{Tr} LX \\ \text{subject to} \ & \operatorname{Tr} E_{ii} X = 1 \qquad i \in \{1, \ldots, n\}, \\ & X \in \mathbb{S}^n_+, \end{aligned}$$

where $E_{ij} \in \mathbb{R}^{n\times n}$ is the 0/1 matrix with support equal to $(i, j)$. Let $\Psi : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ be the orthogonal projection onto any coherent *-subalgebra containing $L$. Then, $\Psi$ satisfies the Constraint Set Invariance Conditions. To see this, it suffices to show that $\Psi$ leaves the solution set of $\operatorname{Tr} E_{ii} X = 1$ invariant. Suppose $\operatorname{Tr} E_{ii} X = 1$. Then, $X = I + Y$ for $Y_{ii} = 0$ and $\Psi(X) = I + \Psi(Y)$ where $[\Psi(Y)]_{ii} = 0$ by Lemma 1.5.2. Hence, $\operatorname{Tr} E_{ii} \Psi(X) = 1$.

## ■ 1.5.3 Aside: more on coherent algebras

As illustrated, the 0/1 basis of a coherent algebra induces a partition of $[n] \times [n]$, where $[n] := \{1, \ldots, n\}$. One calls this partition a *coherent configuration.* The commutant of a group $\mathcal{U} \subset \mathbb{R}^{n\times n}$ of permutation matrices is always coherent. In this case, one calls the underlying coherent configuration *Schurian.* For Schurian coherent configurations, the partition class of $(i, j) \in [n] \times [n]$ equals its *orbit* $\{(\sigma(i), \sigma(j)) : \sigma \in \mathcal{P}\}$, where $\mathcal{P}$ is the set of permutations $\sigma : [n] \to [n]$ realized by the matrices $\mathcal{U}$. Note that (1.16) is the commutant of the following set $\mathcal{U}$:

$$\mathbf{A} = \left\{X \in \mathbb{R}^{n\times n} : UX = XU \ \ \forall U \in \mathcal{U}\right\}, \ \ \mathcal{U} = \left\{ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \right\}.$$

Not all coherent algebras are the commutants of permutation groups [70]. As a consequence, coherent *-subalgebra reduction methods can be more powerful than group theoretic ones. For instance, in the MAXCUT example, the coherent *-subalgebra *generated* by the Laplacian can be properly contained in the commutant of its automorphism group. (The *Shrikhande graph* has this property [125].) Hence, reformulating over the former *-subalgebra can yield a smaller SDP than reformulating over the latter. The gap between these *-subalgebras and the implications for symmetry reduction are discussed in [135, Section 7.2].

## ■ 1.5.4 Wedderburn decomposition

Any *-subalgebra $\mathbf{A}$ of $\mathbb{R}^{n\times n}$ also has a canonical direct-sum decomposition

$$\mathbf{A} = \oplus_{i=1}^{r} \mathbf{A}_i$$

into *simple algebras* $\mathbf{A}_i$. This decomposition is called the *Wedderburn decomposition.* By the bicommutant theorem (Proposition 1.5.1), any *-subalgebra containing the identity matrix $I$ has the form

$$\mathbf{A} = \{X \in \mathbb{R}^{n\times n} : XU = UX, \ \ U \in \mathcal{U}\}$$

for some finite set $\mathcal{U} \subset \mathbb{R}^{n\times n}$ closed under transposition. Each direct-summand of $\mathbf{A}_i$ corresponds to a minimal invariant subspace $\mathcal{S}_i$ of $\mathcal{U} \cup \mathbf{A}$. (Invariant subspaces of $\mathcal{U} \cup \mathbf{A}$ are called *hyper-invariant* subspaces of $\mathcal{U}$.) Specifically, $\mathbf{A}_i$ is the subset of matrices in $\mathbf{A}$ whose ranges are contained in $\mathcal{S}_i$. For this reason, the Wedderburn decomposition is compatible with the decomposition of fixed-point subspaces $\oplus_{i=1}^{r} \mathcal{V}_{\mathcal{G},i}$ reviewed in Section 1.4.4. The next example illustrates this.

**Example 1.5.2.** *For Example 1.4.4, an algebra* $\mathbf{A}$ *satisfying* $\mathbf{A} \cap \mathbb{S}^n = \mathcal{V}_{\mathcal{G}}$ *is*

$$\mathbf{A} = \left\{ \begin{pmatrix} t_1 & t_6 & t_9 & t_9 \\ t_7 & t_2 & t_4 & t_4 \\ t_8 & t_5 & t_3 & t_{10} \\ t_8 & t_5 & t_{10} & t_3 \end{pmatrix} : t \in \mathbb{R}^{10} \right\}.$$

*It has Wedderburn decomposition* $\mathbf{A}_1 \oplus \mathbf{A}_2$*, where*

$$\mathbf{A}_1 = \left\{ \begin{bmatrix} t_1 & t_6 & t_9 & t_9 \\ t_7 & t_2 & t_4 & t_4 \\ t_8 & t_5 & t_3 & t_3 \\ t_8 & t_5 & t_3 & t_3 \end{bmatrix} : t \in \mathbb{R}^9 \right\}, \quad \mathbf{A}_2 = \left\{ \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & t & -t \\ 0 & 0 & -t & t \end{bmatrix} : t \in \mathbb{R} \right\}. \tag{1.17}$$

*Each direct-summand* $\mathcal{V}_{\mathcal{G},i}$ *given by Example 1.4.4 equals* $\mathbf{A}_i \cap \mathbb{S}^n$*.*

Numerical algorithms can find the Wedderburn decomposition of any *-subalgebra; see [89, 49]. Versions also exist for decomposing algebras over the complex numbers [42].

## ■ 1.5.5 Benefits, challenges, and facial reduction comparisons

There are common challenges and benefits of symmetry reduction, *-algebra methods, and facial reduction. Each method enables one to solve smaller problems. Each also has a common challenge of sparsity preservation; for instance, in symmetry reduction and *-algebra methods, one preserves sparsity by finding sparse bases for invariant subspaces (Example 1.4.3). Nevertheless, there are important differences related to pathology removal, dual solution recovery, and algorithms. To explain, we call any reduction technique based on the Constraint Set Invariance Conditions a *projection-based* method. Projection-based methods include symmetry reduction, *-algebra techniques, and a generalization of symmetry reduction proposed in later chapters.

#### Pathology removal and dual solution recovery

**Dual solution recovery** Like facial reduction, projection-based methods *restrict* the primal to a subspace and hence *relax* the dual. In facial reduction, recovering solutions to the original dual from an analogous relaxation was nontrivial and, indeed, impossible in some cases (Section 1.3.6). Remarkably, dual solution recovery is always possible after using a projection-based method. The fundamental reason is the primal-dual symmetry of the Constraint Set Invariance Conditions (Proposition 1.4.1). Chapter 5 elaborates on this more.

**No pathology removal** The flip side to guaranteed recovery is the inability to remove pathologies. Specifically, a projection-based method never changes the primal or dual optimal values nor their attainment. It also doesn't change existence of improving rays. As a consequence, these methods cannot remove pathologies.

**Example 1.5.3** (Failure to remove duality gaps)**.** Consider the primal-dual pair

$$\mathbf{P}(\mathbb{S}^3_+): \begin{array}{ll} \text{minimize} & \operatorname{Tr} CX \\ \text{subject to} & \operatorname{Tr} A_i X = b_i, \\ & X \in \mathbb{S}^3_+ \end{array} \qquad \mathbf{D}(\mathbb{S}^3_+): \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & S = C - \sum_{i=1}^3 y_i A_i \\ & (S, y) \in \mathbb{S}^3_+ \times \mathbb{R}^2, \end{array}$$

where $b = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T$ and

$$C := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad A_1 := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \qquad A_2 := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad A_3 := \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

(The primal SDP is from Example 1.3.4 with the additional constraint that $\operatorname{Tr} A_3 X = 0$.) The projection onto the following coordinate subspace $\mathcal{S}$

$$\begin{bmatrix} * & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{bmatrix}$$

satisfies the Constraint Set Invariance Conditions. The cone $\mathcal{C} = \mathbb{S}^3_+ \cap \mathcal{S}$ satisfies

$$\mathcal{C} = \{X \in \mathbb{S}^3_+ : X_{12} = X_{13} = 0\}, \qquad \mathcal{C}^* = \{X + W : X \in \mathbb{S}^3_+, W \in \mathcal{S}^\perp\}$$

and contains both primal and dual solutions. Nevertheless, restricting the primal to $\mathcal{C}$ and relaxing the dual to $\mathcal{C}^*$ does not remove the duality gap:

$$\inf\{\operatorname{Tr} CX : A_i \cdot X = b_i, X \in \mathcal{C}\} = 1, \qquad \sup\{b^T y : C - \sum_{i=1}^3 y_i A_i \in \mathcal{C}^*\} = 0.$$

The duality gap also persists if one relaxes the primal to $\mathcal{C}^*$ and restricts the dual to $\mathcal{C}$:

$$\inf\{\operatorname{Tr} CX : A_i \cdot X = b_i, X \in \mathcal{C}^*\} = 1, \qquad \sup\{b^T y : C - \sum_{i=1}^3 y_i A_i \in \mathcal{C}\} = 0.$$

#### Slater's condition and singularity degree

While a projection-based method will not remove pathologies, it can, surprisingly, reduce the singularity degree (Section 1.3.4) and hence improve accuracy in the same way as facial reduction (Section 1.3.5). In fact, it can even restore Slater's condition! (Note however that restoring Slater's condition implies the original instance had no pathology—in other words, Slater's condition failed in a benign way.)

**Example 1.5.4.** *Recall the SDP of Example 1.3.5, which had decision variable $X \in \mathbb{S}^n_+$ and constraints*

$$\begin{aligned} [X]_{1,1} &= 1, \\ [X]_{2,2} &= 0, \\ [X]_{k+1,k+1} &= [X]_{1,k} \qquad \forall k \in \{2, 3, \ldots, n-1\}. \end{aligned}$$

*This SDP fails Slater's condition and has singularity degree equal to $n-1$. The orthogo-*

*nal projection onto the subspace spanned by* $e_1e_1^T$ *satisfies the Constraint Set Invariance Conditions. Reformulating over this subspace yields an SDP that satisfies Slater's condition*

$$\text{Find } \lambda \geq 0 \text{ subject to } \lambda e_1 e_1^T = e_1 e_1^T.$$

*Note that* $n-1$ *iterations of facial reduction are needed to obtain the same reformulation.*

#### Algorithms

**Finding minimal reductions** In facial reduction, there is a well defined *minimal face* of the cone program

$$\begin{array}{ll} \text{minimize} & \langle c, x\rangle \\ \text{subject to} & x \in \mathcal{A} \cap \mathcal{K}, \end{array}$$

and an algorithm (Algorithm 1.1) for finding it; indeed, for feasible problems, the minimal face is the unique face of the convex cone $\mathcal{K}$ whose relative interior intersects the affine set $\mathcal{A}$.

In symmetry reduction, there is a subgroup analogous to the minimal face, namely, the largest subgroup $\mathcal{G} \subseteq O(\mathcal{V})$ that satisfies

$$\mathcal{G} \subseteq \text{aut}(\mathcal{K}), \qquad \mathcal{G} \subseteq \text{aut}(x_0 + \mathcal{L}), \qquad \mathcal{G} \subseteq \text{aut}(c + \mathcal{L}^\perp),$$

where $\mathcal{A} = x_0 + \mathcal{L}$. To our knowledge, no algorithm exists for finding $\mathcal{G}$. Nevertheless, if $\mathcal{K}$ is polyhedral, then $\text{aut}(\mathcal{K})$ is well understood and algorithms exists for finding canonical subgroups [26].

For *-algebra methods there is also an analogous object: the range of the minimum rank completely positive projection satisfying the Constraint Set Invariance Conditions. Unfortunately, no algorithm exists for finding this group or this projection, at least to our knowledge. Further, general techniques for approximating these objects—such as generating *-subalgebras from data (Section 1.5.2)—can fail badly in cases where application specific methods succeed. The following example illustrates failure of a general technique on the MAXCUT SDP relaxation. (See Examples 1.4.2 and 1.5.1 for successful application specific methods.) Specifically, it shows the *-subalgebra generated by the data equals the whole space $\mathbb{R}^{n\times n}$ when the graph is connected and hence provides no reductions, even if the graph is highly symmetric.

**Lemma 1.5.3.** *Let* $L$ *be the Laplacian of a graph* $G$ *and consider the MAXCUT relax-*

*ation*

$$\begin{aligned} \text{maximize} \ & \frac{1}{4} \operatorname{Tr} LX \\ \text{subject to} \ & \operatorname{Tr} E_{ii} X = 1 \qquad i \in \{1, \dots, n\}, \\ & X \in \mathbb{S}^n_+, \end{aligned}$$

*where $E_{ij} \in \mathbb{R}^{n \times n}$ is the zero-one matrix with support equal to $(i, j)$. If $G$ is connected, then the *-subalgebra of $\mathbb{R}^{n\times n}$ generated by $E_{11}, E_{22}, \dots, E_{nn}$ and $L$ equals $\mathbb{R}^{n\times n}$.*

*Proof.* If $G$ is connected, there exists $T \in \operatorname{span}\{I, L, L^2, \dots, L^n\}$ that is nonzero in each entry, i.e., $t_{ij} := [T]_{ij} \neq 0$. Since $E_{jj} T E_{ii} = t_{ij} E_{ij}$ is in the algebra and $t_{ij} \neq 0$, the algebra contains every standard basis matrix $E_{ij}$. Hence, it contains $\mathbb{R}^{n\times n}$. □

**Implementations for symmetric cones** The facial reduction algorithm is implementable for any symmetric cone (see, e.g., [86, Section 5]). Symmetry reduction and *-algebra methods, in contrast, are tailored to linear and semidefinite programming (e.g.,[18, 37]), a proper subset of symmetric cone optimization problems. We address this issue in Chapter 5. To do this, we develop another algebraic reduction technique based on *Euclidean Jordan Algebras*, which we review next.

## 1.6 Euclidean Jordan algebras

Chapters 5-7 contain a generalization of symmetry reduction. Specifically, it will show how find a minimum rank projection satisfying the Constraint Set Invariance conditions for any symmetric cone $\mathcal{K}$. This rank minimization approach is grounded in Euclidean Jordan algebra theory, which we now overview. We first define these algebras and then discuss fundamental topics such as the direct-sum decomposition into simple ideals and the classification of simple algebras. Note Euclidean Jordan algebras are precisely the formally-real Jordan algebras [79, Chapter VI, Theorem 12]; some cited references use this latter terminology (e.g., [67]).

### 1.6.1 Preliminaries

We begin with definitions:

**Definition 1.6.1.** *A* real Jordan algebra $\mathbf{J}$ *is an algebra over* $\mathbb{R}$ *(with product denoted $x \circ y$) satisfying the following axioms:*

- $x \circ y = y \circ x$ *(Commutative law)*
- $(x \circ y) \circ x^2 = x \circ (y \circ x^2)$, *where $x^2$ denotes $x \circ x$.* *(Jordan identity)*

*The algebra* $\mathbf{J}$ *is called* Euclidean *if, in addition, it is equipped with an inner product* $\langle \cdot, \cdot \rangle_{\mathbf{J}}$ *satisfying*

$$\langle x \circ y, z \rangle_{\mathbf{J}} = \langle y, x \circ z \rangle_{\mathbf{J}} \qquad \forall x, y, z \in \mathbf{J}.$$

A Euclidean Jordan algebra $\mathbf{J}$ always has an identity element, which we denote $e$. (See, e.g., [79, Chapter III, Theorem 9 and Chapter VI, Corollary 5].) While $\mathbf{J}$ is not in general associative, it is always *power-associative*, meaning $x \circ x \circ \cdots \circ x$ is independent of the order of multiplication—hence, $x^q$ is well-defined for all integers $q > 0$. Algebras are *isomorphic* if there is an isomorphism between them:

**Definition 1.6.2.** *An isomorphism* $\Phi : \mathbf{J}^A \to \mathbf{J}^B$ *between two Euclidean Jordan algebras* $\mathbf{J}^A$ *and* $\mathbf{J}^B$ *is an invertible linear map satisfying*

$$\Phi(x \circ y) = (\Phi x) \circ (\Phi y),$$

*where the multiplication* $x \circ y$ *is carried out in* $\mathbf{J}^A$ *and* $(\Phi x) \circ (\Phi y)$ *in* $\mathbf{J}^B$*.*

Since Jordan algebras are commutative, the identity $2a \circ b = (a+b)^2 - a^2 - b^2$ holds. It follows isomorphisms are precisely the linear maps satisfying $\Phi(x^2) = (\Phi x)^2$.

## ■ 1.6.2 Decomposition into simple ideals

An ideal $\mathbf{I} \subseteq \mathbf{J}$ is a subspace closed under multiplication by arbitrary elements of $\mathbf{J}$, i.e., $\mathbf{I}$ is an ideal if $\mathbf{I}$ contains $\{x \circ y : x \in \mathbf{J}, y \in \mathbf{I}\}$. An algebra $\mathbf{J}$ is called *simple* if the only ideals are $\mathbf{J}$ and $\{0\}$—the so-called trivial ideals. The simple algebras are fully classified up to isomorphism [77]. Further, any algebra has an orthogonal direct-sum decomposition into *simple ideals*—ideals that, when viewed as algebras, are simple [79, Chapter 3, Theorem 11]. Formally:

**Proposition 1.6.1** (Jordan, von Neumann, Wigner)**.** *Any Euclidean Jordan Algebra* $\mathbf{J}$ *equals an orthogonal direct-sum* $\mathbf{J} = \bigoplus_{k=1}^m \mathbf{J}_k$ *of ideals* $\mathbf{J}_k$*, where each ideal (viewed as an algebra) is simple and isomorphic to one of the following:*

1. *The spin-factor algebra* $\mathbb{R} \oplus \mathbb{R}^m$ *with product* $(x_0, x) \circ (y_0, y) = (x_0 y_0 + x^T y, x_0 y + y_0 x)$*.*
2. *The Hermitian matrices of order* $n$ *with real, complex, or quaternion entries, denoted* $\mathbf{H}_n(\mathbb{R})$*,* $\mathbf{H}_n(\mathbb{C})$*,* $\mathbf{H}_n(\mathbb{H})$*, respectively, with product* $X \circ Y = \frac{1}{2}(XY + YX)$*.*
3. *The Hermitian matrices of order* $3$ *with octonion entries, denoted* $\mathbf{H}_3(\mathbb{O})$*, with product* $X \circ Y = \frac{1}{2}(XY + YX)$*.*

*Further, the decomposition* $\mathbf{J} = \bigoplus_{k=1}^m \mathbf{J}_k$ *is unique up-to permutation of the direct summands.*

The third algebra on this list is called the *exceptional algebra*. Note that $\mathbf{H}_n(\mathbb{R})$ is just the vector space $\mathbb{S}^n$ of real symmetric matrices.

### 1.6.3 Cones-of-squares

For any Euclidean Jordan algebra $\mathbf{J}$, the set $\{x \circ x : x \in \mathbf{J}\}$ is called the *cone-of-squares* of $\mathbf{J}$. The cone-of-squares is a closed, convex cone; indeed, it is a symmetric cone, i.e., it is self-dual and homogeneous (Section 1.1). The cone-of-squares is an irreducible symmetric cone if and only if $\mathbf{J}$ is simple. Conversely, any (irreducible) symmetric cone is the cone-of-squares of some (simple) algebra $\mathbf{J}$. It follows there are only three types of irreducible symmetric cones up to invertible linear transformation:

1. The cone-of-squares of the spin-factor algebra, i.e., the Lorentz cone.
2. The cones-of-squares of Hermitian matrices $\mathbf{H}_n(\mathbb{R})$, $\mathbf{H}_n(\mathbb{C})$, $\mathbf{H}_n(\mathbb{H})$ of order $n$, i.e., the psd matrices of order $n$ with real, complex, or quaternion entries.
3. The cone-of-squares of the exceptional algebra.

Note also that any subalgebra $\mathcal{S}$ of $\mathbf{J}$ can be viewed as a Jordan algebra with cone-of-squares $\mathbf{J} \cap \mathcal{S}$. In other words, $\mathbf{J} \cap \mathcal{S}$ is a symmetric cone (viewing $\mathcal{S}$ as the ambient space) when $\mathcal{S}$ is a subalgebra.

### 1.6.4 Idempotents, rank, and spectral decomposition

An element of $x \in \mathbf{J}$ is called idempotent if $x \circ x = x$. An idempotent is called *primitive* if it cannot be written as the sum of two idempotents. A *Jordan frame* is a set of primitive idempotents that are pairwise orthogonal and sum to the identity. All Jordan frames have the same cardinality equal to the *rank* of the algebra, where

$$\operatorname{rank} \mathbf{J} = \max \left\{ n : \{e, x, x^2, \cdots, x^n\} \text{ is linearly independent}, x \in \mathbf{J} \right\}.$$

Jordan frames arise from the *spectral decomposition* of $x \in \mathbf{J}$ which, for symmetric matrices, is the usual eigenvalue decomposition.

**Proposition 1.6.2** (Spectral decomposition, [51, Theorem III.1.2])**.** *Let $\mathbf{J}$ be a Euclidean Jordan algebra. For every element $x \in \mathbf{J}$ there exists a Jordan frame $e_1, \ldots, e_k$ and real numbers $\lambda_1, \ldots, \lambda_k$ (not-necessarily distinct) for which*

$$x = \sum_{i=1}^{k} \lambda_i e_i.$$

*The numbers $\lambda_i$ (with their multiplicities) are uniquely determined by $x$.*

The spectral decomposition also provides a formula for metric projection onto the cone-of-squares:

**Proposition 1.6.3** (Projection onto the cone-of-squares)**.** *Let* $\mathbf{J}$ *be a Euclidean Jordan algebra with cone-of-squares* $\mathcal{K}$*. If* $x \in \mathbf{J}$ *has spectral decomposition* $x = \sum_{i=1}^{k} \lambda_i e_i$*, then*

$$\operatorname*{proj}_{\mathcal{K}}(x) = \sum_{i=1}^{k} \max(\lambda_i, 0) e_i,$$

*where* $\operatorname{proj}_{\mathcal{K}}(x) := \arg\min_{w \in \mathcal{K}} \langle x - w, x - w \rangle$*.*

*Proof.* Let $y = \operatorname{proj}_{\mathcal{K}}(x)$ and $z = \sum_{i=1}^{k} \min(\lambda_i, 0) e_i$. Then, $x = z + y$ where $\langle z, y \rangle = 0$ and $(y, -z) \in \mathcal{K} \times \mathcal{K}^*$ given that $\mathcal{K}^* = \mathcal{K}$. By the Moreau's Theorem [71, Theorem 3.2.5], the result follows. ☐

Note that when $\mathcal{K} = \mathbb{R}^n_+$, this projection operation simply sets negative components of $x \in \mathbb{R}^n$ to zero. When $\mathcal{K} = \mathbb{S}^n_+$, this projection operation sets negative eigenvalues to zero.

## ■ 1.6.5 Special Jordan algebras and representability

Given a real associative algebra $\mathbf{A}$, one obtains a real Jordan algebra (not necessarily Euclidean) by equipping $\mathbf{A}$ with product $x \circ y = \frac{1}{2}(xy + yx)$. This algebra is denoted $\mathbf{A}^+$. A Jordan algebra is called *special* if it is isomorphic to a subalgebra of $\mathbf{A}^+$ for some $\mathbf{A}$, otherwise it is called *exceptional*. The only simple exceptional algebra is $\mathbf{H}_3(\mathbb{O})$. Further, any subalgebra of $\mathbb{S}^n$ is special; indeed, a converse of this statement also holds.

**Proposition 1.6.4** (Characterization of special algebras)**.** *Let* $\mathbf{J}$ *be a Euclidean Jordan algebra. The following statements are equivalent.*

1. *The algebra* $\mathbf{J}$ *is special.*
2. *No ideal of* $\mathbf{J}$ *is isomorphic to* $\mathbf{H}_3(\mathbb{O})$*.*
3. *The algebra* $\mathbf{J}$ *is isomorphic to a subalgebra of* $\mathbb{S}^m$ *for some* $m$*.*

The only nonobvious property is that the first statement implies the third. It is easy to construct these $\mathbb{S}^m$-subalgebras for $\mathbf{H}_n(\mathbb{C})$ and $\mathbf{H}_n(\mathbb{H})$. For instance, $\mathbf{H}_n(\mathbb{C})$ is isomorphic to the subalgebra $\mathcal{S}$ of $\mathbb{S}^{2n}$ given by

$$\mathcal{S} = \left\{ \begin{bmatrix} X & S \\ S^T & X \end{bmatrix} : \;\; X \in \mathbb{S}^n, S \in \mathbb{R}^{n \times n}, \;\; S^T = -S \right\}. \tag{1.18}$$

Further, each spin-factor algebra is isomorphic to a subalgebra of $\mathbf{H}_n(\mathbb{C})$ for some $n$ [67, Section 6.2]; in light of (1.18), it follows each is also isomorphic to a subalgebra of $\mathbb{S}^m$

for some $m$. (We also note that any spin-factor algebra is isomorphic to a subalgebra of $\mathbf{A}^+$, where $\mathbf{A}$ is a real *Clifford* algebra.)

As a consequence of this proposition, we can study any special algebra using the familiar properties of real symmetric matrices. We use this fact to prove theorems in Chapter 5. However, finding an isomorphic subalgebra of $\mathbb{S}^m$ is not always useful for computation: for spin-factors, the order $n$ of the isomorphic $\mathbf{H}_n(\mathbb{C})$-subalgebra given by [67, Section 6.2] (and the order $m$ of the $\mathbb{S}^m$-subalgebra induced by (1.18)) grows exponentially in the dimension of the spin-factor.

# Part I

# Facial reduction

Chapter 2

# Partial facial reduction

## ■ 2.1 Introduction

As explained in Chapter 1.3, facial reduction finds a face containing the feasible set by solving a series of conic feasibility problems. Unfortunately, solving these problems can be expensive and reformulating over the identified face can destroy problem sparsity. Further, reformulating may not leave the optimal value truly unchanged if there is floating-point round-off error. These issues preclude its use in general purpose solvers. It is also inconsistent with a pre-processing philosophy of Andersen and Andersen, who argue the best strategy for pre-processing LPs is to find *simple* simplifications *quickly*. In this chapter, we develop a new facial reduction procedure that avoids these issues and is consistent with this philosophy.

How, then, does one find only 'simple' simplification in the context of facial reduction? One way is to look only for faces exposed by hyperplanes of a particular type. Consider the SDP

$$\begin{array}{ll} \text{Find} & y_1, y_2, y_3 \in \mathbb{R} \\ \text{subject to} & \\ A(y) = \begin{pmatrix} y_1 & 0 & 0 \\ 0 & -y_1 & y_2 \\ 0 & y_2 & y_2 + y_3 \end{pmatrix} \succeq 0. & \end{array}$$

Here, the matrix $A(y)$ is psd if and only if it is contained in the face $\mathbb{S}^3_+ \cap S^\perp$, where

$$\mathbb{S}^3_+ \cap S^\perp = \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & x \end{pmatrix} : x \geq 0 \right\}, \qquad S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Further, the hyperplane $S^\perp$ exposing $\mathbb{S}^3_+ \cap S^\perp$ is structured: the matrix $S$ is contained in the set of matrices that are nonnegative and diagonal—a simple *inner approximation* of the psd cone. This leads to the main idea of this chapter: a facial reduction methodology—which we call *partial facial reduction*—that restricts the search for hyperplanes to a user-specified approximation of the psd cone.

It may seem that ad-hoc specification of the approximation will not lead to a procedure that is particularly useful—if the approximation is not matched to a problem instance, the procedure will fail to find a proper face containing the feasible set, even if one exists. Nevertheless, we show natural approximations are effective for a wide class of examples. We also prove these approximations yield faces that always have sparse representations, allowing one to construct projected reformulations (Chapter 1.2.5) over the face without destroying sparsity. Finally, we demonstrate even *polyhedral* approximations (e.g., nonnegative diagonal matrices) are effective, allowing one to perform facial reduction in exact arithmetic.

This chapter is organized as follows. In Section 2.2, we modify the basic facial reduction algorithm to yield our technique and describe example approximations of the psd cone in Section 2.3. Section 2.4 finds maximum rank solutions to conic optimization problems formulated over these approximations (which helps us find faces of minimal dimension). Section 2.5 shows approximations can be chosen to preserve sparsity. Section 2.6 gives simple, illustrative examples. Section 2.7 describes a freely-available implementation and Section 2.8 shows effectiveness of the method on examples arising in practice.

### 2.1.1 Contributions

**Partial facial reduction** Our main contribution is a facial reduction procedure incorporating a user-specified approximation (Algorithm 2.1). When a polyhedral approximation is specified, this procedure solves only linear programs and computes nullspaces of matrices, which can be done in exact arithmetic, if desired. We demonstrate simple approximations are effective in practice.

**Maximum rank solutions** Related to finding a face of minimal dimension is finding a maximum rank matrix in a subspace intersected with a specified approximation. We show maximum rank matrices can be found by solving a single convex problem (Theorem 2.4.1). When approximations are polyhedral, this convex problem is a linear program.

**Sparse reformulations** We show approximations can also be chosen to preserve sparsity. Specifically, we prove if approximations are contained in the cone of *scaled-diagonally-dominant* [19] matrices, the identified face can always be written as $U\mathbb{S}^d_+U^T$, where the columns of $U \in \mathbb{R}^{n\times d}$ have disjoint support (Theorem 2.5.1). This guarantees an SDP can be reformulated over the identified face without increasing the number of nonzero entries of its data matrices.

**Software implementation** We provide a MATLAB implementation `frlib`, available at `www.mit.edu/~fperment`. If interfaced directly, the code takes as input SDPs in SeDuMi format [129]. It can also be interfaced via the parser YALMIP [84] or SOSTOOLS [112].

## 2.2 Partial facial reduction

We recall the basic idea behind facial reduction (Chapter 1.3.4). Given an inner product space $\mathcal{V}$, an affine set $\mathcal{A} \subseteq \mathcal{V}$ and convex cone $\mathcal{K} \subseteq \mathcal{V}$, facial reduction finds a chain of faces $\mathcal{K} = \mathcal{F}_0 \supset \mathcal{F}_1 \supset \cdots \supset \mathcal{F}_N$ containing $\mathcal{A} \cap \mathcal{K}$ via the recursion

$$\mathcal{F}_0 = \mathcal{K}, \qquad \mathcal{F}_{i+1} = \mathcal{F}_i \cap s_i^{\perp},$$

where each $s_i$ solves the following feasibility problem:

$$\begin{array}{ll} \text{Find} & s_i \in \mathcal{F}_i^* \setminus \mathcal{F}_i^{\perp} \\ \text{subject to} & s_i^{\perp} \supseteq \mathcal{A}, \end{array} \tag{2.1}$$

i.e., each $s_i$ defines a hyperplane $s_i^{\perp}$ containing the affine set $\mathcal{A}$ that exposes a proper face of $\mathcal{F}_i$ (and hence of $\mathcal{K}$).

This leads to a trade-off between the cost of finding $s_i$ (solving the feasibility problem) and the benefit of finding $\mathcal{F}_{i+1}$ (obtaining a lower-dimensional face containing $\mathcal{A} \cap \mathcal{K}$). To manage this trade-off, we propose a simple idea: inner approximate the dual cone $\mathcal{F}_i^*$ at each step of the recursion, or, equivalently, outer approximate $\mathcal{F}_i$. Specifically, we propose selecting $\mathcal{F}_{i,outer}$ that satisfies:

1. $\mathcal{F}_{i,outer} \supseteq \mathcal{F}_i$ (which implies $\mathcal{F}_{i,outer}^* \subseteq \mathcal{F}_i^*$)
2. $\operatorname{span} \mathcal{F}_{i,outer} = \operatorname{span} \mathcal{F}_i$ (i.e., $\mathcal{F}_{i,outer}^{\perp} = \mathcal{F}_i^{\perp}$)
3. $\mathcal{F}_{i,outer}^*$ has low search complexity.

Using the approximation $\mathcal{F}_{i,outer}$, one can then compute a recursion $\mathcal{F}_{i+1} = \mathcal{F}_i \cap s_i^{\perp}$, where each $s_i$ solves

$$\begin{array}{ll} \text{Find} & s_i \in \mathcal{F}_{i,outer}^* \setminus \mathcal{F}_{i,outer}^{\perp} \subseteq \mathcal{F}_i^* \setminus \mathcal{F}_i^{\perp} \\ \text{subject to} & s_i^{\perp} \supseteq \mathcal{A}. \end{array} \tag{2.2}$$

a conic optimization problem that by construction is easy to solve, and, by construction, yields a face $\mathcal{F}_{i+1}$ containing the feasible set $\mathcal{A} \cap \mathcal{K}$. In other words, this approach correctly identifies a face at a user-specified cost—the search complexity of $\mathcal{F}_{i,outer}^*$.

#### Geometric interpretation

Because we have introduced the approximation $\mathcal{F}_{i,outer}$, the feasibility problem (2.2) may not have a solution even if (2.1) does—that is, we may fail to find a proper face containing $\mathcal{A} \cap \mathcal{K}$ even if one exists. We can use Corollary 1.3.2 to interpret this geometrically. Under the assumption that $\mathcal{A} \cap \mathcal{K}$ is nonempty, this corollary implies feasibility of (2.2) is *equivalent* to emptiness of $\mathcal{A} \cap \operatorname{relint} \mathcal{F}_{i,outer}$, whereas feasibility of (2.1) is equivalent to a weaker condition: emptiness of $\mathcal{A} \cap \operatorname{relint} \mathcal{F}$. Figure 2.1 illustrates success and failure of these conditions.

### ■ 2.2.1 Partial facial reduction of SDPs

#### Approximating faces of $\mathbb{S}^n_+$

To apply this idea to SDP, we need a way of approximating faces of $\mathbb{S}^n_+$. To see how this can be done, let $\mathcal{F}$ denote the face $U\mathbb{S}^d_+U^T$ of $\mathbb{S}^n_+$ defined by fixed $U \in \mathbb{R}^{n\times d}$. An approximation $\mathcal{F}_{outer}$ of $\mathcal{F}$ is obtained from an approximation $\hat{\mathbb{S}}^d_+$ of $\mathbb{S}^d_+$. Moreover, the search complexity of $\mathcal{F}^*_{outer}$ depends on the search complexity of $\hat{\mathbb{S}}^d_+$. Consider the following (whose proof is straightforward and omitted):

**Lemma 2.2.1.** *Let $\hat{\mathbb{S}}^d_+ \subseteq \mathbb{S}^d$ be a convex cone containing $\mathbb{S}^d_+$. For $U \in \mathbb{R}^{n\times d}$, let $\mathcal{F}_{outer}$ and $\mathcal{F}$ denote the sets $U\hat{\mathbb{S}}^d_+U^T$ and $U\mathbb{S}^d_+U^T$, respectively. The following statements are true.*

1. $\mathcal{F} \subseteq \mathcal{F}_{outer}$.
2. $\operatorname{span} \mathcal{F} = \operatorname{span} \mathcal{F}_{outer}$
3. $\mathcal{F}^*_{outer} \setminus \mathcal{F}^{\perp}_{outer} = \left\{ X \in \mathbb{S}^n : U^T X U \in (\hat{\mathbb{S}}^d_+)^* \setminus \{0\} \right\}$.

This lemma leads to an SDP partial-facial-reduction procedure, which we state explicitly in Algorithm 2.1. Each iteration of this procedure relies on an approximation $\hat{\mathbb{S}}^d_+$ of the psd cone $\mathbb{S}^d_+$. Example approximations are explored in the next section.

## ■ 2.3 Approximations of $\mathbb{S}^d_+$

In this section, we explore an outer approximation $\mathcal{C}(\mathbb{W})$ of $\mathbb{S}^d_+$ parametrized by a set $\mathbb{W}$ of $d \times k$ rectangular matrices. The parametrization is chosen such that the dual cone $\mathcal{C}(\mathbb{W})^*$ equals the *Minkowski sum* of faces $W_i\mathbb{S}^k_+W_i^T$ of $\mathbb{S}^d_+$ for $W_i \in \mathbb{W}$. It is defined below:

**Lemma 2.3.1.** *For a set $\mathbb{W} := \left\{ W_1, W_2, \ldots, W_{|\mathbb{W}|} \right\}$ of $d \times k$ matrices, let $\mathcal{C}(\mathbb{W})$ denote the following convex cone:*

$$\mathcal{C}(\mathbb{W}) := \left\{ X \in \mathbb{S}^d : W_i^T X W_i \in \mathbb{S}^k_+ \quad i = 1, \ldots, |\mathbb{W}| \right\}.$$

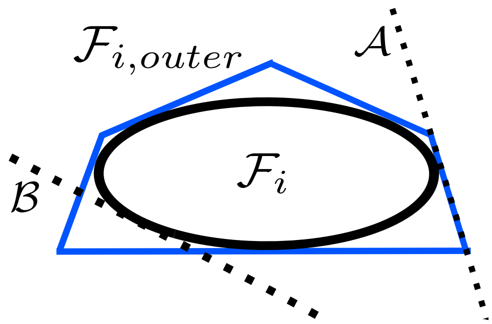



**Figure 2.1:** Illustrates when the feasibility problems (2.1) and (2.2) have solutions. For the affine set $\mathcal{A}$ shown, solutions exist for both. However, if we replace $\mathcal{A}$ with $\mathcal{B}$, solutions do not exist for (2.2) since $\mathcal{B} \cap \operatorname{relint} \mathcal{F}_{i,outer}$ is non-empty. In other words, in this latter case, facial reduction finds a face whereas partial facial reduction fails.

**Algorithm 2.1:** Partial facial reduction algorithm. Given affine set $\mathcal{A} \subseteq \mathbb{S}^n$, finds face $U\mathbb{S}^d_+U^T$ containing $\mathcal{A} \cap \mathbb{S}^n_+$.

**begin**
  Initialize: $U \leftarrow I$, $d \leftarrow n$
  **repeat**
    1. Pick outer-approximation $\hat{\mathbb{S}}^d_+ \supseteq \mathbb{S}^d_+$ and solve

$$\begin{array}{ll} \text{Find} & S \in \mathbb{S}^n \\ \text{subject to} & S^\perp \text{ contains } \mathcal{A} \\ & U^T S U \in (\hat{\mathbb{S}}^d_+)^* \setminus \{0\}. \end{array} \qquad (\star)$$

    2. Find matrix $B \in \mathbb{R}^{d \times r}$ whose columns are a basis for $\operatorname{null} U^T S U$.

    3. Intersect $U\mathbb{S}^d U^T$ with $S^\perp$, i.e., set $U \leftarrow UB$ and $d \leftarrow d - r$.
  **until** $(\star)$ is infeasible;
**end**

*The dual cone* $\mathcal{C}(\mathbb{W})^*$ *satisfies*

$$\mathcal{C}(\mathbb{W})^* = \left\{ \sum_{i=1}^{|\mathbb{W}|} W_i X_i W_i^T : X_i \in \mathbb{S}^k_+ \right\}, \tag{2.3}$$

*and the following inclusions hold:*

$$\mathcal{C}(\mathbb{W})^* \subseteq \mathbb{S}^d_+ \subseteq \mathcal{C}(\mathbb{W}).$$

*Proof.* The inclusions are obvious from the definitions of $\mathcal{C}(\mathbb{W})^*$ and $\mathcal{C}(\mathbb{W})$ (as is the fact that $\mathcal{C}(\mathbb{W})$ is a convex cone). It remains to show correctness of (2.3). To show this, let $\mathcal{T}$ denote the set on the right-hand side of (2.3). It is easy to check that $\mathcal{T}^* = \mathcal{C}(\mathbb{W})$, which implies $\mathcal{T}^{**} = \mathcal{C}(\mathbb{W})^*$. Since $\mathcal{T}$ is a convex cone (as is easily checked), $\mathcal{T}^{**}$ equals the closure of $\mathcal{T}$. The result therefore follows by showing $\mathcal{T}$ is closed. To see this, note that $\mathcal{T}$ equals the Minkowski sum of closed cones $W_i \mathbb{S}^k_+ W_i^T$. For matrices $Z_i \in W_i \mathbb{S}^k_+ W_i^T$, we have that $\sum_{i=1}^{|\mathbb{W}|} Z_i = 0$ only if $Z_i = 0$ for each $i$. This shows that $\sum_{i=1}^{|\mathbb{W}|} Z_i = 0$ only if $Z_i$ is in the lineality space of $W_i \mathbb{S}^k_+ W_i^T$. Direct application of the closedness criteria Corollary 9.1.3 of Rockafellar [121] shows $\mathcal{T}$ is closed. □

Since the modification to the SDP facial reduction algorithm (Algorithm 2.1) will involve searching over $\mathcal{C}(\mathbb{W})^*$ (as indicated by Lemma 2.2.1), we will investigate $\mathcal{C}(\mathbb{W})$ by studying the dual cone $\mathcal{C}(\mathbb{W})^*$. We first make a few comments regarding the search complexity of $\mathcal{C}(\mathbb{W})^*$ for different choices of $\mathbb{W}$. Note when $k = 1$, each $W_j$ in $\mathbb{W}$ is a vector and $\mathcal{C}(\mathbb{W})^*$ is the conic hull of a finite set of rank one matrices. In other words, $\mathcal{C}(\mathbb{W})^*$ is *polyhedral* and can be described by *linear programming*. When $k = 2$, the set $\mathcal{C}(\mathbb{W})^*$ is defined by $2 \times 2$ semidefinite constraints and can hence be described by *second-order cone programming* (SOCP). This follows since each $X_i \in \mathbb{S}^2_+$ can be expressed using scalars $a, b, c$ constrained as follows:

$$X_i = \begin{pmatrix} a+b & c \\ c & a-b \end{pmatrix} \succeq 0 \quad \Leftrightarrow \quad a \geq 0 \quad \text{and } a^2 \geq b^2 + c^2. \tag{2.4}$$

Example choices for $\mathcal{C}(\mathbb{W})^*$ are now given. As we will see, well studied approximations of $\mathbb{S}^d_+$ can be expressed as sets of the form $\mathcal{C}(\mathbb{W})^*$.

**Examples** Example choices for $\mathcal{C}(\mathbb{W})^*$ are given in Table 2.1 along with the cardinality of the set $\mathbb{W}$ that yields each entry. Included are $d \times d$ nonnegative diagonal matrices $\mathcal{D}^d$, diagonally-dominant matrices $\mathcal{DD}^d$, scaled diagonally-dominant matrices $\mathcal{SDD}^d$ as well as matrices $\mathcal{FW}^d_k$ with *factor-width* [19] bounded by $k$. These sets satisfy

$$\mathcal{D}^d = \mathcal{FW}^d_1 \subseteq \mathcal{DD}^d \subseteq \mathcal{SDD}^d = \mathcal{FW}^d_2 \subseteq \mathcal{FW}^d_3 \subseteq \cdots \subseteq \mathcal{FW}^d_d = \mathbb{S}^d_+,$$

| $\mathcal{C}(\mathbb{W})$ | $\mathcal{C}(\mathbb{W})^*$ | Search | $\lvert\mathbb{W}\rvert$ |
|---|---|---|---|
| $X_{ii} \geq 0$ | Non-negative diagonal ($\mathcal{D}^d$) | LP | $\mathcal{O}(d)$ |
| $X_{ii} \geq 0,\ X_{jj} + X_{ii} \pm 2X_{ij} \geq 0$ | Diagonally-dominant ($\mathcal{DD}^d$) | LP | $\mathcal{O}(d^2)$ |
| $2 \times 2$ principal sub-matrices psd | Scaled diagonally-dominant ($\mathcal{SDD}^d$) | SOCP | $\mathcal{O}(d^2)$ |
| $k \times k$ principal sub-matrices psd | Factor width-$k$ ($\mathcal{FW}_k^d$) | SDP | $\mathcal{O}(\binom{d}{k})$ |

**Table 2.1:** Example outer approximations of $\mathbb{S}_+^d$ and their dual cones, the search algorithm for $\mathcal{C}(\mathbb{W})^*$, and the cardinality of the set $\mathbb{W}$.

and the sets $\mathcal{D}^d$ and $\mathcal{DD}^d$ are polyhedral. We now describe the approximations in more detail.

## 2.3.1 Non-negative diagonal matrices ($\mathcal{D}^d$)

A simple choice for $\mathcal{C}(\mathbb{W})^* \subseteq \mathbb{S}_+^d$ is the set of nonnegative diagonal matrices:

$$\mathcal{D}^d := \left\{ X \in \mathbb{S}^d : X_{ii} \geq 0, \quad X_{ij} = 0 \ \ \forall i \neq j \right\}.$$

The set $\mathcal{D}^d$ contains nonnegative combinations of matrices $w_i w_i^T$, where $w_i$ is a permutation of $(1, 0, \ldots, 0, 0)^T$. In other words, the set $\mathcal{D}^d$ corresponds to the set $\mathcal{C}(\mathbb{W})^*$ if we take

$$\mathbb{W} = \left\{ (1, 0, \ldots, 0, 0)^T, (0, 1, \ldots, 0, 0)^T, \ldots, (0, 0, \ldots, 0, 1)^T \right\}.$$

## 2.3.2 Diagonally-dominant matrices ($\mathcal{DD}^d$)

Another well studied choice for $\mathcal{C}(\mathbb{W})^*$ is the cone of symmetric diagonally-dominant matrices with nonnegative diagonal entries [10]:

$$\mathcal{DD}^d := \left\{ X \in \mathbb{S}^d : X_{ii} \geq \sum_{j \neq i} |X_{ij}| \right\}.$$

This set is polyhedral. The *extreme rays* of $\mathcal{DD}^d$ are matrices of the form $w_i w_i^T$, where $w_i$ is any permutation of

$$(1, 0, 0, \ldots, 0)^T, (1, 1, 0, \ldots, 0)^T, \text{ or } (1, -1, 0, \ldots, 0)^T.$$

Taking $\mathbb{W}$ equal to the set of all such permutations gives $\mathcal{C}(\mathbb{W})^* = \mathcal{DD}^d$. This representation makes the inclusion $\mathcal{DD}^d \subseteq \mathbb{S}_+^d$ obvious. We also see that $\mathcal{DD}^d$ contains $\mathcal{D}^d$.

### 2.3.3 Scaled diagonally-dominant matrices ($\mathcal{SDD}^d$)

A non-polyhedral generalization of $\mathcal{DD}^d$ is the set of *scaled diagonally-dominant matrices* $\mathcal{SDD}^d$. This set equals all matrices obtained by pre- and post-multiplying diagonally-dominant matrices by diagonal matrices with strictly positive diagonal entries:

$$\mathcal{SDD}^d := \left\{DTD : D \in \mathcal{D}^d, D_{ii} > 0, T \in \mathcal{DD}^d\right\}.$$

The set $\mathcal{SDD}^d$ can be equivalently defined as the set of matrices that equal the sum of psd matrices nonzero only on a $2 \times 2$ principal sub-matrix (Theorem 9 of [19]). As an explicit example, we have that $\mathcal{SDD}^3$ are all matrices $X$ of the form

$$X = \underbrace{\begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{12} & a_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{X_1} + \underbrace{\begin{pmatrix} b_{11} & 0 & b_{13} \\ 0 & 0 & 0 \\ b_{13} & 0 & b_{33} \end{pmatrix}}_{X_2} + \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & c_{22} & c_{23} \\ 0 & c_{23} & c_{33} \end{pmatrix}}_{X_3},$$

where $a_{ij}$, $b_{ij}$, and $c_{ij}$ are scalars chosen such that $X_1, X_2$ and $X_3$ are psd. In general, $\mathcal{SDD}^d$ equals $\mathcal{C}(\mathbb{W})^*$ when $\mathbb{W}$ equals the set of $d \times 2$ matrices $W$ for which $W^T X W$ returns a $2 \times 2$ principal sub-matrix of $X$. For $\mathcal{SDD}^3$, we have

$$\mathcal{SDD}^3 = \mathcal{C}(\{W_1, W_2, W_3\})^* = \left\{\sum_{i=1}^{3} W_i X_i W_i^T : X_i \in \mathbb{S}^2_+\right\},$$

where

$$W_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \quad W_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \quad W_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Also note from (2.4) that $\mathcal{SDD}^d$ can be represented using second-order cone constraints. This latter fact is used in recent work of Ahmadi and Majumdar [1] to define an SOCP-based method for testing polynomial nonnegativity. (A similar LP-based method is also presented in [1] that incorporates $\mathcal{DD}^d$.)

**The kernels of $\mathcal{SDD}$ matrices** The kernel of a scaled diagonally-dominant matrix has a structured basis of vectors with *disjoint support*, where the support of a vector $u \in \mathbb{R}^n$ is the set of indices $i$ for which $u_i \neq 0$. This follows because, up-to permutation, a scaled diagonally-dominant is block-diagonal, where each block is either positive definite, equals the zero matrix, or has co-rank one (i.e., has a one dimensional kernel), as shown in [30]. In Section 2.5, we use this result to show SDPs can be reformulated over faces found via $\mathcal{SDD}$ approximations without damaging sparsity (which, since

$\mathcal{D}^d \subseteq \mathcal{DD}^d \subseteq \mathcal{SDD}^d$, also holds for diagonally-dominant or diagonal approximations). The following proposition summarizes relevant results of Chen and Toledo [30]. We include an elementary—and different—proof for completeness.

**Proposition 2.3.1.** *Suppose $X \in \mathbb{S}^d_+$ is scaled-diagonally dominant. Then, there is a permutation matrix $P \in \mathbb{R}^{d\times d}$ for which*

$$PXP^T = \begin{pmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & X_M \end{pmatrix}, \tag{2.5}$$

*where, for all $m \in \{1, \ldots, M\}$, the matrix $X_m \in \mathbb{S}^{d_m}_+$ is either positive definite, a matrix of all zeros, or has co-rank one. Moreover, when $X$ has co-rank $r$, there is a matrix $U \in \mathbb{R}^{d\times r}$ whose columns have disjoint support and span the nullspace of $X$.*

*Proof.* For $X \in \mathbb{S}^d$, let $G_X := ([d], E)$ denote the graph with node set $[d] := \{1, \ldots, d\}$, where $\{i, j\}$ is in the edge set $E$ if and only if $X_{ij} \neq 0$. Clearly there is a permutation matrix $P$ that block-diagonalizes $X$ as in (2.5), defined in the obvious way by the connected-components of $G_X$.

Now suppose $P$ in (2.5) equals this permutation. That $X_m$ has the claimed properties is immediate when $d_m \leq 2$. Now suppose $d_m > 2$ and that $X_m$ is nonzero and not positive definite. Also, define the graph $G_{X_m} = ([d_m], E_m)$, where $\{i, j\}$ is in the edge set $E_m$ if and only if $[X_m]_{ij} \neq 0$ and observe $G_{X_m}$ is connected (and, indeed, isomorphic to a connected component of $G_X$ defined above.)

We first claim all entries of $v \in \operatorname{null} X_m \setminus \{0\}$ are nonzero. To begin, pick $i \in [d_m]$ such that $v_i$ is nonzero. For arbitrary $t \in [d_m] \setminus i$, there is a path $T \subseteq E_m$ from $i$ to $t$ for which

$$X_m = \bar{X} + \sum_{\{r,s\}\in T} (e_r, e_s) X_{rs} (e_r, e_s)^T,$$

where all entries of $X_{rs} \in \mathbb{S}^2_+$ are nonzero and $\bar{X}$ is positive semidefinite. Picking the first edge $\{i, j\} \in T$, we conclude that $X_{ij}(e_i, e_j)^T v = X_{ij}(v_i, v_j)^T = 0$. For the sake of contradiction, suppose $v_j = e_j^T v = 0$. Then, $(v_i, 0)^T$ is in the kernel of $X_{ij}$, showing a diagonal entry of $X_{ij}$ is zero (since $v_i \neq 0$), contradicting the fact all entries of $X_{ij}$ are nonzero. Hence, $v_j \neq 0$. Repeating this argument using the next edge $\{j, k\}$ in $T$ shows $v_k \neq 0$. Repeating for all edges in $T$ shows $v_t \neq 0$. Since $t$ was arbitrary, all components of $v$ are nonzero.

Now pick another nonzero $w \in \operatorname{null} X_m$ and consider the consecutive edges $\{i, j\}$

and $\{j,k\}$ in the path $T$. Then, for scalars $\lambda$ and $\gamma$,

$$(e_i,e_j)^T w = \lambda(e_i,e_j)^T v, \quad (e_j,e_k)^T w = \gamma(e_j,e_k)^T v,$$

otherwise the nonzero matrices $X_{ij} \in \mathbb{S}^2_+$ and $X_{jk} \in \mathbb{S}^2_+$ have two-dimensional kernels, and are therefore the zero matrix, a contradiction. But since $v_j$ and $w_j$ are nonzero, we also have $\lambda = \gamma$. Since any $s,t \in [d_m]$ are connected by a path, we conclude $w = \lambda v$.

Existence of $U$ is immediate, given that the kernel of $X_m$ has a basis of the form $\{e_1,\ldots,e_{d_m}\}$, $\{0\}$, or $\{v\}$.

□

### 2.3.4 Factor-width-k matrices

A generalization of $\mathcal{SDD}^d$ (and diagonal matrices $\mathcal{D}^d$) arises from notion of *factor-width* [19]. The factor-width of a matrix $X$ is the smallest integer $k$ for which $X$ can be written as the sum of psd matrices that are nonzero only on a single $k\times k$ principal sub-matrix.

Letting $\mathcal{FW}^d_k$ denote the set of $d\times d$ matrices of factor-width no greater than $k$, we have that $\mathcal{SDD}^d = \mathcal{FW}^d_2$ and $\mathcal{D}^d = \mathcal{FW}^d_1$. To represent $\mathcal{FW}^d_k$ as a cone of the form $\mathcal{C}(\mathbb{W})^*$, we set $\mathbb{W}$ to be the set of $d\times k$ matrices $W_j$ for which $W_j^T X W_j$ returns a $k\times k$ principal sub-matrix of $X$. Note that there are $\binom{d}{k}$ such matrices, so a complete parametrization of $\mathcal{FW}^d_k$ is not always practical using this representation. Also note $\mathcal{FW}^d_k$ equals $\mathbb{S}^d_+$ when $k = d$.

## 2.4 Faces of minimal dimension via rank maximization

Each iteration of the partial facial reduction algorithm (Algorithm 2.1) computes a face $\mathcal{F}\cap S^\perp$, where $\mathcal{F}$ is the current face at the start of the iteration and $S$ is the solution of a conic feasibility problem. If $\mathcal{F} = U\mathbb{S}^d_+U^T$, the dimension of $\mathcal{F}\cap S^\perp$ is determined by the co-rank of $U^TSU$. Precisely, if $U^TSU$ has co-rank $r$, then $\mathcal{F}\cap S^\perp$ equals $V\mathbb{S}^{d-r}V^T$ for an appropriate $V$. Hence, one can minimize $d-r$ by replacing the feasibility problem with

$$\begin{array}{ll} \text{maximize} & \operatorname{rank} U^TSU \\ \text{subject to} & U^TSU \in (\hat{\mathbb{S}}^d_+)^* \\ & S^\perp \text{ contains } \mathcal{A} \end{array} \tag{2.6}$$

where $(\hat{\mathbb{S}}^d)^*$ is the chosen approximation of $\mathbb{S}^d_+$. In this section, we show (2.6) can be solved by a single convex problem when $\hat{\mathbb{S}}^d$ is an approximation of the form $\mathcal{C}(\mathbb{W})$ studied in Section 2.3. Precisely, we show:

**Theorem 2.4.1.** *Suppose $\hat{\mathbb{S}}^d_+ = \mathcal{C}(\mathbb{W})$. A solution to the rank maximization problem*

(2.6) *is given by any* $S$ *that for some* $T_k, \bar{S}_k$ *solves*

$$\begin{array}{lll} \text{maximize} & \sum_{k=1}^{|\mathbb{W}_i|} \operatorname{Tr} T_k & \\ \text{subject to} & U^T S U = \sum_{k=1}^{|\mathbb{W}_i|} W_k \bar{S}_k W_k^T, & \textit{i.e.,}\ U^T S U \in \mathcal{C}(\mathbb{W}_i)^* \\ & \bar{S}_k \succeq T_k & \forall k \in \{1, \ldots, |\mathbb{W}_i|\} \\ & I \succeq T_k \succeq 0 & \forall k \in \{1, \ldots, |\mathbb{W}_i|\} \\ & S^\perp \supseteq \mathcal{A} & \end{array}$$

which will be an immediate corollary of Lemmas 2.4.1-2.4.2, to be stated and proven next.

To ease notation, we note (2.6) falls into a simple problem class: it is a rank maximization problem over the intersection of two convex cones—the dual cone of the approximation $\hat{\mathbb{S}}^d_+$ and the linear subspace $\{U^T S U \in \mathbb{S}^d : S^\perp \supseteq \mathcal{A}\}$. This observation motivates the following two lemmas, where we take $\hat{\mathbb{S}}^d_+$ to be of form $\mathcal{C}(\mathbb{W})$ and replace the subspace with an arbitrary convex cone $\mathcal{M}$.

**Lemma 2.4.1.** *Let* $\mathcal{M} \subseteq \mathbb{S}^d$ *be a convex cone. If* $X^\star := \sum_{i=1}^{|\mathbb{W}|} W_i X_i^\star W_i^T$ *maximizes* $\sum_{i=1}^{|\mathbb{W}|} \operatorname{rank} X_i$ *over* $\mathcal{M} \cap \mathcal{C}(\mathbb{W})^*$*, then* $X^\star$ *maximizes* $\operatorname{rank} X$ *over* $\mathcal{M} \cap \mathcal{C}(\mathbb{W})^*$*.*

*Proof.* We will argue the kernel of $X^\star$ is contained in the kernel of any $X \in \mathcal{M} \cap \mathcal{C}(\mathbb{W})^*$, which immediately implies $\operatorname{rank} X^\star \geq \operatorname{rank} X$.

To begin, we first argue for any $X = \sum_{i=1}^{|\mathbb{W}|} W_i X_i W_i^T \in \mathcal{M} \cap \mathcal{C}(\mathbb{W})^*$ that $\operatorname{null} X_i^\star \subseteq \operatorname{null} X_i$ for all $i \in \{1, \ldots, |\mathbb{W}|\}$. To see this, first note that for any $X \in \mathcal{M} \cap \mathcal{C}(\mathbb{W})^*$ the matrix

$$X^\star + X = \sum_{i=1}^{|\mathbb{W}|} W_i (X_i^\star + X_i) W_i^T$$

is also in $\mathcal{M} \cap \mathcal{C}(\mathbb{W})^*$ and satisfies $\operatorname{rank}(X_i^\star + X_i) \geq \operatorname{rank} X_i^\star$. Now suppose for some $d \in \{1, \ldots, |\mathbb{W}|\}$ that $\operatorname{null} X_d^\star \not\subseteq \operatorname{null} X_d$. This implies that $\operatorname{null}(X_d^\star + X_d) = \operatorname{null} X_d^\star \cap \operatorname{null} X_d \subset \operatorname{null} X_d^\star$ which in turn implies $\operatorname{rank}(X_d^\star + X_d) > \operatorname{rank} X_d^\star$. But this contradicts our assumption that $X^\star$ maximizes $\sum_i \operatorname{rank} X_i$. Hence, $\operatorname{null} X_i^\star \subseteq \operatorname{null} X_i$ for all $i \in \{1, \ldots, |\mathbb{W}|\}$.

Now suppose an $X \in \mathcal{M} \cap \mathcal{C}(\mathbb{W})^*$ exists for which $X^\star w = 0$ but $Xw \neq 0$ for some $w$. Since $Xw = 0$ if and only if $X_i W_i^T w = 0$ for all $i$, we must have for some $d$ that $W_d^T w$ is in the kernel of $X_d^\star$ but not in the kernel of $X_d$. But we have already established that $\operatorname{null} X_d^\star \subseteq \operatorname{null} X_d$. Hence, $w$ cannot exist. We therefore have that $\operatorname{null} X^\star \subseteq \operatorname{null} X$ for any $X \in \mathcal{M} \cap \mathcal{C}(\mathbb{W})^*$, which completes the proof. □

We can use this condition to formulate an SDP whose optimal solutions yield maximum rank matrices of $\mathcal{M} \cap \mathcal{C}(\mathbb{W})^*$. To maximize $\sum_{i=1}^{|\mathbb{W}|} \operatorname{rank} X_i$, we introduce matrices

$T_i$ constrained such that their traces $\operatorname{Tr} T_i$ lower bound $\operatorname{rank} X_i$. We then optimize the sum of their traces.

**Lemma 2.4.2.** *Let $\mathcal{M} \subseteq \mathbb{S}^d$ be a convex cone. A matrix $X$ maximizing $\sum_{i=1}^{|\mathbb{W}|} \operatorname{rank} X_i$ over $\mathcal{M} \cap \mathcal{C}(\mathbb{W})^*$ is given by any optimal solution $(X, X_i, T_i)$ to the following SDP:*

$$\begin{array}{lll} \text{maximize} & \sum_{i=1}^{|\mathbb{W}|} \operatorname{Tr} T_i & \\ \text{subject to} & X \in \mathcal{M}, & \\ & X = \sum_{i=1}^{|\mathbb{W}|} W_i X_i W_i^T & \text{i.e. } X \in \mathcal{C}(\mathbb{W})^* \\ & X_i \succeq T_i & \forall i \in \{1, \ldots, |\mathbb{W}|\} \\ & I \succeq T_i \succeq 0 & \forall i \in \{1, \ldots, |\mathbb{W}|\}. \end{array} \tag{2.7}$$

*Proof.* Let $r_{\max}$ equal the maximum of $\sum_{i=1}^{|\mathbb{W}|} \operatorname{rank} X_i$ over the set of feasible $X_i$. We will show at optimality $\sum_{i=1}^{|\mathbb{W}|} \operatorname{rank} X_i = r_{\max}$.

To begin, the constraint $I \succeq T_i \succeq 0$ implies the eigenvalues of $T_i$ are less than one. Hence, $\operatorname{rank} T_i \geq \operatorname{Tr} T_i$. Since $X_i \succeq T_i$ , we also have $\operatorname{rank} X_i \geq \operatorname{rank} T_i$. Thus, any feasible $(X_i, T_i)$ pair satisfies

$$r_{\max} \geq \sum_{i=1}^{|\mathbb{W}|} \operatorname{rank} X_i \geq \sum_{i=1}^{|\mathbb{W}|} \operatorname{rank} T_i \geq \sum_{i=1}^{|\mathbb{W}|} \operatorname{Tr} T_i. \tag{2.8}$$

Now note for any feasible $(X, X_i)$ we can pick $\alpha > 0$ and construct a feasible point $(\alpha X, \alpha X_i, \hat{T}_i)$ that satisfies $\sum_{i=1}^{|\mathbb{W}|} \operatorname{Tr} \hat{T}_i = \sum_{i=1}^{|\mathbb{W}|} \operatorname{rank} X_i$; if $X_i$ has eigen-decomposition $\sum_j \lambda_j u_j u_j^T$ for $\lambda_j > 0$, simply take $\hat{T}_i = \sum_j u_j u_j^T$ and $\alpha$ equal to

$$\max \bigcup_i \left\{ \frac{1}{\lambda} : \lambda \text{ is a positive eigenvalue of } X_i \right\}.$$

Hence, some feasible point $(\hat{X}, \hat{X}_i, \hat{T}_i)$ satisfies $\sum_{i=1}^{|\mathbb{W}|} \operatorname{Tr} \hat{T}_i = r_{\max}$. Therefore, the optimal $(X, X_i, T_i)$ satisfies

$$\sum_{i=1}^{|\mathbb{W}|} \operatorname{Tr} T_i \geq r_{\max}.$$

Combining this inequality with (2.8) yields that at optimality

$$\sum_{i=1}^{|\mathbb{W}|} \operatorname{Tr} T_i = \sum_{i=1}^{|\mathbb{W}|} \operatorname{rank} X_i = r_{\max},$$

which completes the proof. □

Combining the previous two lemmas shows solving (2.7) also maximizes rank over $\mathcal{M} \cap \mathcal{C}(\mathbb{W})^*$; hence, Theorem 2.4.1 is proven by taking $\mathcal{M}$ equal to the linear subspace

$\{U^T SU \in \mathbb{S}^d : S^\perp \supseteq \mathcal{A}\}$.

## ■ 2.5 Sparse reformulations

In this section, we show that sparse reformulations can be obtained by using scaled-diagonally-dominant approximations. To fix ideas, suppose our original SDP is of the form

$$\begin{aligned} \text{minimize} \quad & \operatorname{Tr} CX \\ \text{subject to} \quad & \operatorname{Tr} A_i X = b_i \qquad i \in \{1, \ldots, m\}, \\ & X \in \mathbb{S}^n_+, \end{aligned}$$

and consider the face $\mathcal{F} = \left\{ U\hat{X}U^T : \hat{X} \in \mathbb{S}^d_+ \right\}$ defined by a fixed $U \in \mathbb{R}^{n \times d}$ with $d < n$. We can reformulate this SDP over $\mathcal{F}$ by replacing $X$ with $U\hat{X}U^T$. Applying the cyclic property of trace to the equations and objective function

$$\operatorname{Tr} A_i U\hat{X}U^T = \operatorname{Tr} U^T A_i U\hat{X}, \qquad \operatorname{Tr} CU\hat{X}U^T = \operatorname{Tr} U^T CU\hat{X},$$

then yields a projected reformulation (Chapter 1.2.5) over $\mathbb{S}^d_+$:

$$\begin{aligned} \text{minimize} \quad & \operatorname{Tr}(U^T CU)\hat{X} \\ \text{subject to} \quad & \operatorname{Tr}(U^T A_i U)\hat{X} = b_i \;\; i \in \{1, \ldots, m\}, \\ & \hat{X} \in \mathbb{S}^d_+. \end{aligned}$$

Depending on $U$, the matrices $U^T CU$ and $U^T A_i U$, though of order $d < n$, may be dense even if $C$ and $A_i$ are sparse. In this section, we show how to *guarantee* these matrices are sparse when $A_i$ and $C$ are sparse. We do this by guaranteeing the matrix $U$ is structured—specifically, that its columns have disjoint support:

**Lemma 2.5.1.** *Suppose the columns of $U$ have disjoint support, i.e., $U_{ij} \neq 0$ and $U_{ik} \neq 0$ implies $j = k$. Then,*

$$\operatorname{nnz}(U^T XU) \leq \operatorname{nnz}(X) \qquad \forall X \in \mathbb{S}^n,$$

*where* $\operatorname{nnz}(X)$ *denotes the number of nonzero entries of $X$.*

*Proof.* Under the disjoint support assumption, each column of $XU$ is a combination of *distinct* columns of $X$; hence, $\operatorname{nnz}(XU) \leq \operatorname{nnz}(X)$. Further, each row of $U^T XU$ is a combination of distinct rows of $XU$. Hence, $\operatorname{nnz}(U^T XU) \leq \operatorname{nnz}(XU)$. □

We now show it is possible to always obtain a face $\mathcal{F} = U\mathbb{S}^d_+ U^T$ for which the columns of $U$ have disjoint support using Algorithm 2.1. The key is to use scaled diagonally-dominant approximations ($\mathcal{SDD}^d$) at each iteration. Formally:

**Theorem 2.5.1.** *Let $\mathcal{F} = U\mathbb{S}^d_+U^T$ for $U$ in $\mathbb{R}^{n\times d}$, where the columns of $U$ have disjoint support. Suppose $S \in \{X \in \mathbb{S}^n : U^TXU \in \mathcal{SDD}^d\} \subseteq \mathcal{F}^*$. Then, there exists $r \leq d$ and $V \in \mathbb{R}^{n\times r}$ for which*

$$\mathcal{F} \cap S^{\perp} = V\mathbb{S}^r_+V^T$$

*where the columns of $V$ have disjoint support.*

*Proof.* We can take $V$ equal to $V = UB$ where $B$ is any matrix whose columns $B$ form a basis for the kernel of $U^TSU$. By Proposition 2.3.1, we can find a $B$ whose columns have disjoint support. Since the columns of $UB$ have disjoint support when the columns of $U$ and $B$ do, the claim follows. □

Note this statement also applies to diagonal ($\mathcal{D}^d$) and diagonally-dominant ($\mathcal{DD}^d$) approximations, since both $\mathcal{D}^d$ and $\mathcal{DD}^d$ are subsets of $\mathcal{SDD}^d$. Indeed, if we replace $\mathcal{SDD}^d$ with $\mathcal{D}^d$, we can even pick the columns of $U$ to be standard basis vectors, implying $U^TXU$ is a principal submatrix of $X$ for all $X$—that is, we construct $U^TXU$ simply by deleting rows and columns from $X$.

## ■ 2.6 Illustrative Examples

We now consider simple examples that illustrate when approximations are effective. We also illustrate how they lead to sparse reformulations (Theorem 2.5.1).

### ■ 2.6.1 Example with diagonal approximations ($\mathcal{D}^d$)

$$\begin{array}{ll} \text{Find} & y \in \mathbb{R}^4 \\ \text{subject to} & \end{array}$$

$$A(y) = \begin{pmatrix} y_1 & 0 & 0 & 0 & 0 \\ 0 & -y_1 & y_2 & 0 & 0 \\ 0 & y_2 & y_2 - y_3 & 0 & 0 \\ 0 & 0 & 0 & y_3 & 0 \\ 0 & 0 & 0 & 0 & y_4 \end{pmatrix} \in \mathbb{S}^5_+.$$

where $A : \mathbb{R}^4 \to \mathbb{S}^n$ is a linear map and $\mathcal{A} = \{A(y) : y \in \mathbb{R}^4\}$. Taking $U_0$ equal to the identity matrix and the initial face equal to $\mathcal{F}_0 = U_0\mathbb{S}^5_+U_0$, we seek a matrix $S_0$ orthogonal to $\mathcal{A}$ for which $U_0^TS_0U_0$ is nonnegative and diagonal. An $S_0$ satisfying this constraint and a basis $B$ for null $U_0^T S_0 U_0$ is given by:

$$S_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Taking $U_1 = U_0 B = B$, yields the face $\mathcal{F}_1 = U_1 \mathbb{S}_+^3 U_1^T$, i.e., the set of psd matrices in $\mathbb{S}_+^5$ with vanishing first and second rows/cols.

Continuing to the next iteration, we seek a matrix $S_1$ orthogonal to $\mathcal{A}$ for which $U_1^T S_1 U_1$ is nonnegative and diagonal. An $S_1$ satisfying this constraint and a basis $B$ for null $U_1^T S_1 U_1$ is given by:

$$S_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Setting $U_2 = U_1 B$ gives the face $\mathcal{F}_2 = U_2 \mathbb{S}_+^1 U_2^T$, where $U_2 = (0, 0, 0, 0, 1)^T$.

Terminating the algorithm, we now formulate a reduced SDP over $\mathcal{F}_2$. Letting $V$ denote a basis for null $U_2^T$ yields:

$$\begin{array}{ll} \text{Find} & y \in \mathbb{R}^4 \\ \text{subject to} & U_2^T A(y) U_2 \in \mathbb{S}_+^1 \\ & U_2^T A(y) V = 0 \\ & V^T A(y) V = 0, \end{array}$$

which simplifies to

$$\begin{array}{ll} \text{Find} & y \in \mathbb{R}^4 \\ \text{subject to} & y_4 \geq 0 \\ & y_1 = y_2 = y_3 = 0. \end{array}$$

**Geometric interpretation** Corollary 1.3.2 states that existence of $S_i \in \mathcal{F}_{i,outer}^* \setminus \mathcal{F}_{i,outer}^\perp$ implies $\mathcal{A} \cap \operatorname{relint} \mathcal{F}_{i,outer}$ is empty. We now verify this fact. Clearly, $\mathcal{A}$ is contained in $\operatorname{relint} \mathcal{F}_{0,outer}$ only if the inequalities

$$y_1 \geq 0 \qquad -y_1 \geq 0$$

are strictly satisfied, which cannot hold. Similarly, $\mathcal{A}$ is contained in relint $\mathcal{F}_{1,outer}$ only if $y_1 = y_2 = 0$ and the inequalities

$$y_3 \geq 0 \qquad y_2 - y_3 \geq 0$$

are strictly satisfied, which again cannot hold.

## 2.6.2 Example with diagonally-dominant approximations ($\mathcal{DD}^d$)

In this next example, we equip Algorithm 2.1 with diagonally-dominant approximations; i.e. at iteration $i$, the face $\mathcal{F}_i := U_i \mathbb{S}_+^{d_i} U_i^T$ is approximated by the set $\mathcal{F}_{i,outer} = U_i \mathcal{C}(\mathbb{W}_i) U_i^T$, where $\mathcal{C}(\mathbb{W}_i)^*$ equals $\mathcal{DD}^{d_i}$, the set of $d_i \times d_i$ matrices that are diagonally-dominant. An exposing vector $S_i$ is found in $\mathcal{F}_{i,outer}^*$, the set of matrices $X$ for which $U_i^T X U_i$ is in $\mathcal{DD}^{d_i}$. We apply the algorithm to the SDP

$$\begin{array}{ll} \text{Find} & y \in \mathbb{R}^3 \\ \text{subject to} & \\ & A(y) = \begin{pmatrix} 1 & -y_1 & 0 & -y_3 \\ -y_1 & 2y_2 - 1 & y_3 & 0 \\ 0 & y_3 & 2y_1 - 1 & -y_2 \\ -y_3 & 0 & -y_2 & 1 \end{pmatrix} \in \mathbb{S}_+^4, \end{array}$$

where $A : \mathbb{R}^3 \to \mathbb{S}^n$ is an affine map and $\mathcal{A} = \{A(y) : y \in \mathbb{R}^3\}$. Taking $U_0$ equal to the identity, a matrix $S_0$ orthogonal to $\mathcal{A}$ for which $U_0^T S_0 U_0$ is diagonally-dominant and a basis $B$ for null $U_0^T S_0 U_0$ is given by

$$S_0 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \quad B = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \end{pmatrix}.$$

Taking $U_1 = U_0 B = B$ yields the face $\mathcal{F}_1 = U_1 \mathbb{S}_+^2 U_1^T$. (Note the columns of $U$ have disjoint support, a reflection of Theorem 2.5.1.)

Terminating the algorithm, we can describe $\mathcal{A} \cap \mathcal{F}_1$ using $S_0$. Specifically, $\mathcal{A} \cap \mathcal{F}_1$ equals the psd matrices of the form $A(y)$ that satisfy $A(y)S_0 = 0$. These latter constraints hold if and only if $y_1 = y_2 = 1$ and $y_3 = 0$, showing $\mathcal{A} \cap \mathcal{F}_1$ consists of a single point $A\left((1,1,0)^T\right)$.

**Geometric interpretation** As was the case in the previous example, we can interpret this example geometrically. That is, we can verify emptiness of $\mathcal{A} \cap$ relint $\mathcal{F}_{i,outer}$. At the first (and only) iteration, if $A(y) \in \mathcal{A}$ is also in $\mathcal{F}_{0,outer}$, then $w_k^T A(y) w_k \geq 0$, where $w_k w_k^T$ is any extreme ray of $\mathcal{DD}^4$. Taking $w_1 = (1,1,0,0)^T$ and $w_2 = (0,0,1,1)^T$, we

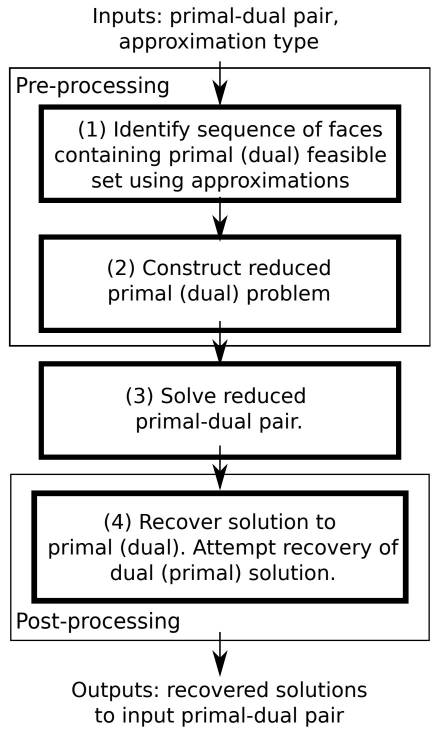



**Figure 2.2:** Flow of MATLAB implementation

have that $\mathcal{A}$ is contained in relint $\mathcal{F}_{0,outer}$ only if the inequalities

$$w_1^T A(y) w_1 = 2y_2 - 2y_1 \geq 0$$
$$w_2^T A(y) w_2 = 2y_1 - 2y_2 \geq 0$$

are strictly satisfied for some $y$, which cannot hold.

## ■ 2.7 Implementation

The discussed techniques have been implemented as a suite of MATLAB scripts we dub `frlib`, available at at `www.mit.edu/~fperment`. The basic work flow is depicted in Figure 2.2. The implemented code takes as input a primal-dual SDP *pair* and applies Algorithm 2.1 to either the primal problem or the dual problem. This is an important feature since either the primal or the dual may model the problem of interest.

### ■ 2.7.1 Input formats

The implementation takes in SeDuMi-formatted inputs `A,b,c,K`, where `A,b,c`, define the subspace constraint and objective function and `K` specifies the sizes of the semidefinite constraints [129]. Conventionally, the primal problem described by `A,b,c,K` refers to an SDP defined by equations $A_i \cdot X = b_i$. Similarly, the dual problem described by A,b,c,K refers to an SDP defined by generators $C - \sum_i y_i A_i$.

## 2.7.2 Reduction of the primal problem

Given A,b,c,K; the following syntax is used to reduce the primal problem, solve the reduced primal-dual pair, and recover solutions to the original primal-dual pair via our implementation:

```
prg = frlibPrg(A,b,c,K);
prgR = prg.ReducePrimal('d');
[x_reduced,y_reduced] = sedumi(prgR.A, prgR.b, prgR.c, prgR.K);
[x,y,dual_recov_success] = prgR.Recover(x_reduced,y_reduced);
```

The call to `prg.ReducePrimal` reduces the primal problem using diagonal ( `'d'` ) approximations by executing a variant of Algorithm 1.1. To execute Algorithm 2.1, it solves a series of LPs (defined by the diagonal approximation) that can be solved using a handful of supported solvers. The returned object `prgR` has member variables

`prgR.A, prgR.b, prgR.c, prgR.K,`

which describe the reduced primal-dual pair. For a single semidefinite constraint, this reduced primal-dual pair is given by:

$$\begin{array}{lll} \text{min.} & C \cdot U\hat{X}U^T \\ \text{subj. to} & A_i \cdot U\hat{X}U^T = b_i \ \ \forall i \in \{1,\ldots,m\} \\ & \hat{X} \in \mathbb{S}^d_+ \end{array} \qquad \begin{array}{ll} \text{max.} & b^T y \\ \text{subj. to} & U^T(C - \sum_{i=1}^m y_i A_i)U \in \mathbb{S}^d_+, \\ & \end{array}$$

where $U\mathbb{S}^d_+U^T$ is a face identified by `prg.ReducePrimal`. The reduced primal and its dual are solved by calling SeDuMi.

The primal solution `x_reduced` returned by SeDuMi represents an optimal $\hat{X}$. The function `prgR.Recover` computes from $\hat{X}$ a solution $U\hat{X}U^T$ to the original primal problem. It then attempts to find a solution to the original dual using a variant of the recovery procedure described in Chapter 3 (Algorithm 3.1). The flag `dual_recov_success` indicates success of this recovery procedure.

## 2.7.3 Reduction of the dual problem

The above syntax can be modified to reduce the dual problem described by A,b,c,K. This is done replacing the relevant line above with:

```
prgR = prg.ReduceDual('d');
```

As above, the object `prgR` contains a description of the primal-dual pair.

With `prgR` created in this manner, a call to `prgR.Recover` (though syntactically identical) now returns a solution to the original *dual* and attempts to recover a solution to the original *primal* using techniques discussed in Chapter 3 (which don't always succeed). In other words, a call of the form

```
[x,y,prim_recov_success] = prgR.Recover(x_reduced,y_reduced);
```

returns a solution `y` to the original dual problem and attempts to recover a solution `x` to the original primal problem. The flag `prim_recov_success` indicates successful recovery of `x`.

### ■ 2.7.4 Solution recovery

As suggested by the flags `prim_recov_success` and `dual_recov_success` in the preceding examples, solution recovery is only guaranteed for the problem that is reduced, i.e. if the primal (resp. dual) is reduced, recovery of the original dual (resp. primal). Thus, it is important to reduce the primal only if it is the problem of interest, and similarly for the dual.

## ■ 2.8 Examples

This section gives larger examples that illustrate effectiveness of our method. For each example, the same type of approximation (e.g. diagonal or diagonally-dominant) is used at each facial reduction iteration. Many examples are also over products of cones, e.g. $\mathcal{K} = \mathbb{S}^{n_1} \times \mathbb{S}^{n_2} \times \cdots \times \mathbb{S}^{n_k}$. In these cases, we use the same type of approximation for each cone $\mathbb{S}^{n_i}$. For each example, we report one or more of the following items $(1-4)$:

**1) Complexity parameters and sparsity** For each example, we report a list of numbers describing the size and sparsity of the problem, denoted

$$n; r; \text{nnz}\,.$$

Here, $n$ gives the size(s) of the psd cone(s) and $r$ the dimension of the affine subspace that together define the feasible set. The number nnz is the total number of nonzero entries of the matrix `A` and cost vector `c` used to describe the problem in SeDuMi format. These results show problem size is often significantly reduced and sparsity enhanced by our method.

**2) DIMACS errors and distance to face** We report a tuple $(e_1, \ldots, e_6)$ of DIMACS errors [93] for the original problem and reduced problem. We also report the distance $d_{face}$ (in norm induced by the trace inner product) of the solution to the subspace spanned by the identified face. Specifically, if $X \in \mathbb{S}^n$ is an optimal solution to the original

SDP (obtained by direct solution or recovered from a solution to the reduced SDP), we report

$$d_{face} = \|\Phi(X) - X\|_F,$$

where $\Phi : \mathbb{S}^n \to \mathbb{S}^n$ is the orthogonal projection map onto the mentioned subspace and $\|\cdot\|_F$ denotes the Frobenius norm. Note if the face equals $U\mathbb{S}^d_+U^T$ for $U$ with orthonormal columns, then

$$\Phi(X) = UU^T XUU^T.$$

Note $d_{face}$ should always be zero if $X$ is an exact solution and the face has been identified in exact arithmetic.

The reported errors show the reduced SDP can be solved just as accurately as the original in terms of DIMACS error. They also show that by the measure $d_{face}$, solutions recovered from the reduced SDP are significantly more accurate. That $d_{face}$ is larger for the original SDP reflects the fact DIMACS error (a measure of backwards-error) can be a poor measure of forwards-error when strict feasibility fails. (A phenomena observed in [130].)

**3) Reduction error** When one iteration of facial reduction is performed, we report the minimum eigenvalue of the exposing vector $S$ and a measure of the containment $S^\perp \supseteq \mathcal{A}$, where $\mathcal{A}$ is the affine set of the SDP.

**4) Solve times** For larger instances, we give solve times before and after reductions and report the total time $t_{LPs}$ spent solving LPs. These solve times are reported for an Intel(R) Core(TM) i7-2600K CPU @ 3.40GHz machine with 16 gigabytes of RAM using the LP solver of MOSEK and the SDP solver SeDuMi called from MATLAB 2014a running Ubuntu. For these instances, solve time is significantly reduced and the cost of solving LPs is negligible.

## 2.8.1 Lower bounds for optimal multi-period investment

Our first example arises from SDP-based lower bounds of optimal multi-period investment strategies. The strategies and specific SDP formulations are given in [24]. For each strategy, an SDP produces a quadratic lower bound on the value function arising in the dynamic programming solution to the underlying optimization problem. These bounds are produced using the $S$-procedure, an SDP-based method for showing emptiness of sets defined by quadratic polynomials (see, e.g., [22]). We report reductions using diagonal ($\mathcal{D}^d$) approximations, DIMACs error, reduction error, and solve time in Tables 2.2-2.5. Scripts that generate the SDPs are found here (and require the package CVX [64]):

www.stanford.edu/~boyd/papers/matlab/port_opt_bound/port_opt_code.tgz

## 2.8.2 Copositivity of quadratic forms

Our next example pertains to SDPs that demonstrate *copositivity* of certain quadratic forms. A quadratic form $x^T Jx$ is copositive if and only if $x^T Jx \geq 0$ for all $x$ in the nonnegative orthant. Deciding copositivity is NP-hard, but a sufficient condition can be checked using sum-of-squares techniques and semidefinite programming, as we now illustrate.

**The Horn form** An example of a copositive polynomial is the Horn form $f(x) := x^T Jx$, where

$$J = \begin{pmatrix} 1 & -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 & 1 \end{pmatrix}, \qquad x = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \end{pmatrix}^T.$$

This polynomial, originally introduced by A. Horn, appeared previously in [44] [113]. To see how copositivity can be demonstrated using SDP, first note copositivity of $f(x)$ is equivalent to global nonnegativity of $f(z_1^2, z_2^2, z_3^2, z_4^2, z_5^2)$, where we have substituted each variable $x_i$ with the square of a new indeterminate $z_i^2$. Next, note global nonnegativity of the latter polynomial can be demonstrated by showing

$$g(z) = \left(\sum_{i=1}^{5} z_i^2\right) f(z_1^2, z_2^2, z_3^2, z_4^2, z_5^2) \tag{2.9}$$

is a sum-of-squares, which is equivalent to feasibility of a particular SDP over $\mathbb{S}_+^n$ where $n = \binom{5+2}{3}$, the number of degree-three monomials in 5 variables (see Chapter 3 of [15] for details on constructing this SDP).

**Generalized Horn forms** The Horn form $f(x)$ generalizes to a family of copositive forms in $n = 3m + 2$ variables ($m \geq 1$):

$$B(x; m) = \left(\sum_{i=1}^{3m+2} x_i\right)^2 - 2 \sum_{i=1}^{3m+2} x_i \sum_{j=0}^{m} x_{i+3j+1},$$

where we let the subscript for the indeterminate $x$ wrap cyclically, i.e. $x_{r+n} = x_r$. This family was studied in [11], and the Horn form corresponds to the case $m = 1$. As with the Horn form, we can show copositivity of $B(x; m)$ by showing a polynomial analogous to (2.9) is a sum-of-squares. We formulate SDPs that demonstrate copositivity of

| Example | $n$ | $r$ | nnz |
|---|---|---|---|
| `long_only` | $(91 \times 100, 30 \times 100)$ | 59095 | 853011 |
| `unconstrained` | $(121 \times 100, 30 \times 100)$ | 62095 | 874011 |
| `sector_neutral` | $(121 \times 100, 30 \times 100)$ | 62392 | 1373000 |
| `leverage_limit` | $(151 \times 100, 30 \times 100)$ | 68195 | 915993 |

**(a)** Original

| Example | $n$ | $r$ | nnz |
|---|---|---|---|
| `long_only` | $(61 \times 100, 30 \times 100)$ | 56095 | 832011 |
| `unconstrained` | $(61 \times 100, 30 \times 100)$ | 56095 | 840891 |
| `sector_neutral` | $(61 \times 100, 30 \times 100)$ | 56392 | 1342880 |
| `leverage_limit` | $(61 \times 100, 30 \times 100)$ | 59195 | 873873 |

**(b)** Reduced

**Table 2.2:** Dimension $r$ of subspace and order $(n_1, \ldots, n_{200})$ of cone $\mathbb{S}^{n_1}_+ \times \cdots \times \mathbb{S}^{n_{200}}_+$ describing feasible set. The column 'nnz' shows number of nonzero entries of SDP data matrices.

| Example | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_5$ | $e_6$ | $d_{face}$ |
|---|---|---|---|---|---|---|---|
| `long_only` | 4.3e-08 | 0 | 0 | 4.4e-11 | -4.0e-05 | -3.9e-05 | 2.3e-06 |
| `unconstrained` | 7.1e-08 | 0 | 0 | 1.1e-11 | -2.5e-06 | -1.6e-06 | 2.1e-05 |
| `sector_neutral` | 4.2e-07 | 0 | 0 | 1.1e-10 | -1.4e-08 | 3.1e-05 | 1.6e-04 |
| `leverage_limit` | 7.3e-08 | 0 | 0 | 1.0e-11 | -1.6e-06 | -6.4e-07 | 1.2e-05 |

**(a)** Original

| Example | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_5$ | $e_6$ | $d_{face}$ |
|---|---|---|---|---|---|---|---|
| `long_only` | 3.7e-08 | 0 | 0 | 1.5e-11 | -6.8e-06 | -5.8e-06 | 2.9e-17 |
| `unconstrained` | 4.9e-08 | 0 | 0 | 1.2e-11 | -4.9e-07 | 3.9e-07 | 3.1e-17 |
| `sector_neutral` | 3.5e-07 | 0 | 0 | 1.0e-10 | -3.5e-08 | 3.0e-05 | 3.5e-17 |
| `leverage_limit` | 4.8e-08 | 0 | 0 | 9.7e-12 | -1.1e-06 | -2.9e-07 | 4.3e-14 |

**(b)** Reduced

**Table 2.3:** DIMACS errors $e_i$ and distance $d_{face}$ to linear span of identified face.

| Example | $\lvert C \cdot S \rvert$ | $\max_i \lvert A_i \cdot S \rvert$ | $\lambda_{\min}(S)$ |
|---|---|---|---|
| `long_only` | 0 | 0 | 0 |
| `unconstrained` | 0 | 0 | 0 |
| `sector_neutral` | 0 | 0 | 0 |
| `leverage_limit` | 0 | 0 | 0 |

**Table 2.4:** Reduction error. The first two columns measure containment of the SDP's affine subspace in the hyperplane $S^\perp$. The last denotes the minimum eigenvalue of the exposing vector $S$.

| Example | Original | Reduced | $t_{LPs}$ |
|---|---|---|---|
| `long_only` | 651 | 613 | 0.33 |
| `unconstrained` | 800 | 574 | 0.71 |
| `sector_neutral` | 760 | 496 | 0.70 |
| `leverage_limit` | 976 | 617 | 1.2 |

**Table 2.5:** Solve times (sec) for original and reduced SDPs. The reduced SDP was formulated by solving LPs over diagonal approximations, i.e., by taking $\mathcal{C}(\mathbb{W}) = \mathcal{D}^d$. These LPs took $t_{LPs}$ seconds to solve.

$B(x;m)$ in this way for each $m \in \{1,\ldots,5\}$. We report reductions using diagonally-dominant ($\mathcal{DD}^d$) approximations, DIMACs error, reduction error, and solve time in Tables 2.6-2.9. (Errors and solve time are omitted for $m > 3$ since the SDPs are too large to solve.)

## ■ 2.8.3 Lower bounds on completely positive rank

A matrix $A \in \mathbb{S}^n$ is *completely positive* (CP) if there exist $r$ nonnegative vectors $v_i \in \mathbb{R}^n$ for which

$$A = \sum_{i=1}^{r} v_i v_i^T. \tag{2.10}$$

The completely positive rank of $A$, denoted $\operatorname{rank}_{\text{cp}} A$, is the smallest $r$ for which $A$ admits the decomposition (2.10). It follows trivially that

$$\operatorname{rank} A \leq \operatorname{rank}_{\text{cp}} A.$$

In [53], Fawzi and the second author give an SDP formulation that improves this lower bound for a fixed matrix $A$. This bound, denoted $\tau_{cp}^{sos}(A)$ in [53], equals the optimal value of the following semidefinite program:

$$\begin{array}{ll} \text{minimize } t & \\ \text{subject to} & \\ & \begin{pmatrix} t & \operatorname{vect} A^T \\ \operatorname{vect} A & X \end{pmatrix} \succeq 0 \\ & X_{ij,ij} \leq A_{ij}^2 \quad \forall i,j \in \{1,\ldots,n\} \\ & X \preceq A \otimes A \\ & X_{ij,kl} = X_{il,jk} \quad \forall (1,1) \leq (i,j) < (k,l) \leq (n,n), \end{array}$$

where $A \otimes A$ denotes the *Kronecker product* and $\operatorname{vect} A$ denotes the $n^2 \times 1$ vector obtained by stacking the columns of $A$. Here, the double subscript $ij$ indexes the $n^2$ rows (or columns) of $X$ and the inequalities on $(i,j)$ hold iff they hold element-wise (see [53] for further clarification on this notation).

In this example, we formulate SDPs as above for computing $\tau_{cp}^{sos}(Z)$, $\tau_{cp}^{sos}(Z \otimes Z)$, and $\tau_{cp}^{sos}(Z \otimes Z \otimes Z)$, where $Z$ is the completely positive matrix:

$$Z = \begin{pmatrix} 4 & 0 & 1 \\ 0 & 4 & 1 \\ 1 & 1 & 3 \end{pmatrix}.$$

Notice that since $Z$ is CP, the Kronecker products $Z \otimes Z$ and $Z \otimes Z \otimes Z$ are CP (using the fact that $A \otimes B$ is $CP$ when $A$ and $B$ are CP [12]). Also notice that since $Z$ contains

| Example | $n$ | $r$ | nnz |
|---|---|---|---|
| $m = 1$ | 35 | 420 | 1225 |
| $m = 2$ | 120 | 5544 | 14400 |
| $m = 3$ | 286 | 33033 | 81796 |
| $m = 4$ | 560 | 129948 | 313600 |
| $m = 5$ | 969 | 395352 | 938961 |

**(a)** Original

| Example | $n$ | $r$ | nnz |
|---|---|---|---|
| $m = 1$ | 25 | 165 | 1200 |
| $m = 2$ | 96 | 3132 | 14312 |
| $m = 3$ | 242 | 21879 | 81554 |
| $m = 4$ | 490 | 494143 | 313040 |
| $m = 5$ | 867 | 303399 | 937822 |

**(b)** Reduced

**Table 2.6:** Dimension $r$ of subspace and order $n$ of cone $\mathbb{S}^n$ describing feasible set. The column 'nnz' shows number of nonzero entries of SDP data matrices.

| Example | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_5$ | $e_6$ | $d_{face}$ |
|---|---|---|---|---|---|---|---|
| $m = 1$ | 8.63e-10 | 0 | 0 | 9.99e-11 | 1.99e-10 | 3.14e-08 | 8.47e-06 |
| $m = 2$ | 9.34e-09 | 0 | 0 | 3.68e-10 | 8.12e-10 | 6.05e-07 | 3.96e-05 |
| $m = 3$ | 1.87e-09 | 0 | 0 | 1.01e-10 | 1.97e-10 | 4.16e-07 | 3.83e-05 |

**(a)** Original

| Example | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_5$ | $e_6$ | $d_{face}$ |
|---|---|---|---|---|---|---|---|
| $m = 1$ | 7.82e-10 | 0 | 0 | 2.42e-11 | 7.64e-11 | 2.04e-08 | 9.27e-16 |
| $m = 2$ | 1.23e-09 | 0 | 0 | 1.59e-10 | 3.82e-10 | 5.84e-08 | 3.26e-16 |
| $m = 3$ | 4.00e-10 | 0 | 0 | 7.08e-11 | 2.25e-10 | 7.93e-08 | 7.48e-16 |

**(b)** Reduced

**Table 2.7:** DIMACs errors $e_i$ and distance $d_{face}$ to linear span of identified face.

| Example | $\lvert b^T y \rvert$ | $\lVert S - \sum_i y_i A_i \rVert_F$ | $\lambda_{\min}(S)$ |
|---|---|---|---|
| $m = 1$ | 0 | 3.33e-16 | 0 |
| $m = 2$ | 0 | 1.67e-16 | 0 |
| $m = 3$ | 0 | -1.28e-15 | 0 |

**Table 2.8:** Reduction error. The first two columns measure containment of the SDP's affine subspace in the hyperplane $S^{\perp}$. The last denotes the minimum eigenvalue of exposing vector $S$.

| Example | Original | Reduced | $t_{LPs}$ |
|---|---|---|---|
| $m = 1$ | .81 | .23 | .047 |
| $m = 2$ | 11 | 9.2 | .58 |
| $m = 3$ | 3900 | 3200 | 4.3 |

**Table 2.9:** Solve times (sec) for original and reduced SDPs. The reduced SDP was formulated by solving LPs over diagonal approximations, i.e., by taking $\mathcal{C}(\mathbb{W}) = \mathcal{DD}^d$. These LPs took $t_{LPs}$ seconds to solve.

zeros, the constraint $X_{ij,ij} \leq Z_{ij}^2$ implies that $X$ has rows and columns identically zero; in other words, because $Z$ has elements equal to zero, the SDP for computing $\tau_{cp}^{sos}(Z)$ cannot have a strictly feasible solution.

To reduce the formulated SDPs, we first observe that each is actually a cone program over $\mathbb{R}_+^{n_1} \times \mathbb{S}_+^{n_2} \times \mathbb{S}_+^{n_3}$, i.e., each SDP has a mix of linear inequalities and semidefinite constraints. To find reductions, we first treat the linear equalities as a semidefinite constraint on a diagonal matrix. We report reductions using diagonal ($\mathcal{D}^d$) approximations, DIMACs error, reduction error, and solve time in Tables 2.10-2.13. (Solve times and errors are omitted for $Z \otimes Z \otimes Z$ since the SDP is too large to solve.)

## 2.8.4 Lyapunov Analysis of a Hybrid Dynamical System

The next example arises from SDP-based stability analysis of a *rimless wheel*, a hybrid dynamical system and simple model for walking robots studied in [110] by Posa, Tobenkin, and Tedrake. The SDP includes several coupled semidefinite constraints that impose Lyapunov-like stability conditions accounting for Coulomb friction and the contact dynamics of the rimless wheel. We report reductions using diagonally-dominant ($\mathcal{DD}^d$) and diagonal ($\mathcal{D}^d$) approximations, DIMACs error, and solve time in Tables 2.14-2.16. (Reduction error is omitted since multiple facial reduction iterations were performed.)

## 2.8.5 Multi-affine polynomials, matroids, and the half-plane property

A multivariate polynomial $f(z) : \mathbb{C}^n \rightarrow \mathbb{C}$ has the *half-plane* property if it is nonzero when each variable $z_i$ has positive real part. A polynomial is *multi-affine* if each indeterminate is raised to at most the first power. As proven in [32], if a multi-affine, homogeneous polynomial with unit coefficients has the half-plane property, it is the *basis generating polynomial* of a *matroid*. In this section, we reduce SDPs that arise in the study of the converse question: given a matroid, does its basis generating polynomial have the half-plane property? Or more precisely, given a rank-$r$ matroid $M$ (over the ground-set $\{1, \ldots, n\}$) with set of bases $B(M)$, does the multi-affine, degree-$r$ polynomial

$$f_M(z_1, \ldots, z_n) := \sum_{\substack{\{i_1, i_2, \ldots, i_r\} \\ \in B(M)}} z_{i_1} z_{i_2} \cdots z_{i_r} \tag{2.11}$$

have the half-plane property?

**The role of polynomial nonnegativity** This converse question is related to global nonnegativity of so-called *Rayleigh differences* of $f_M(z)$, which are polynomials over $\mathbb{R}^n$

| Example | $n$ | $r$ | nnz |
|---|---|---|---|
| $Z$ | $(9, 10, 9)$ | 37 | 260 |
| $Z \otimes Z$ | $(81, 82, 81)$ | 2026 | 18344 |
| $Z \otimes Z \otimes Z$ | $(729, 730, 729)$ | 142885 | 1428692 |

**(a)** Complexity parameters - original

| Example | $n$ | $r$ | nnz |
|---|---|---|---|
| $Z$ | $(7, 8, 9)$ | 20 | 187 |
| $Z \otimes Z$ | $(49, 50, 81)$ | 464 | 8336 |
| $Z \otimes Z \otimes Z$ | $(343, 344, 729)$ | 13262 | 408403 |

**(b)** Complexity parameters - reduced

**Table 2.10:** Dimension $r$ of subspace and order $n$ of cone $\mathbb{R}^{n_1}_+ \times \mathbb{S}^{n_1}_+ \times \mathbb{S}^{n_2}_+$ describing feasible set. The column 'nnz' shows number of nonzero entries of SDP data matrices.

| Example | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_5$ | $e_6$ | $d_{face}$ |
|---|---|---|---|---|---|---|---|
| $Z$ | 2.08e-11 | 0 | 0 | 1.09e-10 | -1.36e-09 | -1.44e-09 | 1.02e-05 |
| $Z \otimes Z$ | 6.58e-09 | 0 | 0 | 1.72e-10 | 2.82e-06 | 2.68e-06 | 2.42e-03 |

**(a)** Original

| Example | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_5$ | $e_6$ | $d_{face}$ |
|---|---|---|---|---|---|---|---|
| $Z$ | 1.27e-11 | 0 | 0 | 7.87e-11 | -1.54e-09 | -1.59e-09 | 0 |
| $Z \otimes Z$ | 3.91e-08 | 0 | 0 | 0 | 8.50e-06 | 7.56e-06 | 0 |

**(b)** Reduced

**Table 2.11:** DIMACs errors $e_i$ and distance $d_{face}$ to linear span of identified face.

| Example | $\lvert C \cdot S\rvert$ | $\max_i \lvert A_i \cdot S\rvert$ | $\lambda_{\min}(S)$ |
|---|---|---|---|
| $Z$ | 0 | 0 | 0 |
| $Z \otimes Z$ | 0 | 0 | 0 |

**Table 2.12:** Reduction error. The first two columns measure containment of the SDP's affine subspace in the hyperplane $S^{\perp}$. The last denotes the minimum eigenvalue of the exposing vector $S$.

| Example | Original | Reduced | $t_{LPs}$ |
|---|---|---|---|
| $Z$ | .4 | .7 | .0084 |
| $Z \otimes Z$ | 131 | 10.5 | .016 |

**Table 2.13:** Solve times (sec) for original and reduced SDPs. The reduced SDP was formulated by solving LPs over diagonal approximations, i.e., by taking $\mathcal{C}(\mathbb{W}) = \mathcal{D}^d$. These LPs took $t_{LPs}$ seconds to solve.

| Problem | $n$ | $r$ | nnz |
|---|---|---|---|
| Original | $(6, 108, 11, 11, 11, 11, 11, 11, 11, 11, 11, 11)$ | 4334 | 16864 |
| Reduced, $\mathcal{C}(\mathbb{W}) = \mathcal{D}^d$ | $(6, 56, 11, 1, 1, 0, 11, 1, 1, 0, 11, 11)$ | 1138 | 6661 |
| Reduced, $\mathcal{C}(\mathbb{W}) = \mathcal{DD}^d$ | $(6, 34, 8, 1, 1, 0, 8, 1, 1, 0, 9, 7)$ | 452 | 4007 |

**Table 2.14:** The feasible set is an $r$-dimensional subspace intersected with the cone $\mathbb{S}_+^{n_1} \times \mathbb{S}_+^{n_2} \cdots \times \mathbb{S}_+^{n_{12}}$.

| Problem | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_5$ | $e_6$ | $d_{face}$ |
|---|---|---|---|---|---|---|---|
| Original | 2.56e-07 | 0 | 0 | 2.48e-10 | 9.78e-08 | 1.70e-05 | 8.99e-02 |
| Red., $\mathcal{C}(\mathbb{W}) = \mathcal{D}^d$ | 2.77e-08 | 0 | 0 | 0 | 1.76e-08 | 8.29e-06 | 0 |
| Red., $\mathcal{C}(\mathbb{W}) = \mathcal{DD}^d$ | 6.65e-08 | 0 | 0 | 0 | 4.29e-08 | 1.20e-05 | 3.82e-15 |

**Table 2.15:** DIMACS error bounds $e_i$ and distance $d_{face}$ to the linear span of identified face.

| Problem | Solve time | $t_{LPs}$ |
|---|---|---|
| Original | 111 | – |
| Reduced, $\mathcal{C}(\mathbb{W}) = \mathcal{D}^d$ | 5 | .05 |
| Reduced, $\mathcal{C}(\mathbb{W}) = \mathcal{DD}^d$ | 1.8 | 0.82 |

**Table 2.16:** Solve times (sec) for original and reduced SDPs. The reduced SDP was formulated by solving LPs over the indicated approximation ($\mathcal{C}(\mathbb{W}) = \mathcal{D}^d$ or $\mathcal{C}(\mathbb{W}) = \mathcal{DD}^d$) which took $t_{LPs}$ seconds to solve.

defined for each $\{i,j\} \subset \{1,\ldots,n\}$ as follows:

$$\Delta_{ij} f_M(x) := \frac{\partial f_M}{\partial z_i}(x)\frac{\partial f_M}{\partial z_j}(x) - \frac{\partial^2 f_M}{\partial z_i \partial z_j}(x) \cdot f_M(x).$$

A theorem of Brändén [25] states $f_M(z)$ has the half-plane property if and only if all of $\binom{n}{2}$ Rayleigh differences are globally nonnegative, i.e., $\Delta_{ij} f_M(x) \geq 0$ for all $x \in \mathbb{R}^n$. An equivalent criterion, stated in terms of global nonnegativity of a single Rayleigh difference (and so-called *contractions* and *deletions* of $M$), appears in [136].

**The role of semidefinite programming** Since semidefinite programming can demonstrate a given polynomial is a sum-of-squares, it is a natural tool for proving a given Rayleigh difference $\Delta_{ij} f_M(x)$ is globally nonnegative. In this section, we formulate and then apply our reduction technique to SDPs that test the sum-of-squares condition for various $\Delta_{ij} f_M(x)$ and various matroids $M$. As is standard, the SDPs are formulated using the set of monomial exponents in $\frac{1}{2}\mathcal{N}(\Delta_{ij} f_M) \cap \mathbb{N}^n$, where $\mathcal{N}(\Delta_{ij} f_M)$ denotes the *Newton polytope* of $\Delta_{ij} f_M$ (see Chapter 3 of [15] for details on this formulation).

We report reductions using diagonally-dominant $\mathcal{DD}^d$ approximations, DIMACs error, and reduction error in Tables 2.17-2.19. (Solve time is omitted since the original SDPs are small.) We now elaborate on each matroid in these tables.

### Various matroids with the half-plane property

The first set of matroids were studied by Wagner and Wei [136]. Specifically, Wagner and Wei [136] demonstrate that $\Delta_{ij} f_M$ (for specific $\{i,j\}$) is a sum-of-squares for matroids $M$ they denote $\mathcal{F}_7^{-4}$, $\mathcal{W}^{3+}$, $\mathcal{W}^3{+}e$, $\mathcal{P}_7'$, $n\mathcal{P}\backslash 1$, $n\mathcal{P}\backslash 9$, and $\mathcal{V}_8$. (We refer the reader to [136] for definitions of these matroids and the explicit polynomials $\Delta_{ij} f_M$.) Note Wagner and Wei demonstrate each sum-of-squares condition via ad-hoc construction, instead of by solving an SDP.

Notice from Table 2.17 that for matroids $\mathcal{W}^{3+}$, $\mathcal{W}^3 + e$, $\mathcal{P}_7'$, $n\mathcal{P} \setminus 1$ and $n\mathcal{P} \setminus 9$, the reduced SDP is described by a zero-dimensional affine subspace. In other words, the SDP demonstrating the sum-of-squares condition has a feasible set containing a single point.

### Extended Vmos matroid

The other matroid considered was studied by Burton, Vinzant, and Youm in [28]. There, the authors use semidefinite programming to show $\Delta_{ij} f_{\mathcal{V}_{10}}$ is a sum-of-squares for a specific $\{i,j\}$, where $\mathcal{V}_{10}$ denotes the *extended Vámos matroid* defined over the ground set $\{1,\ldots,10\}$. The bases of $\mathcal{V}_{10}$ are all cardinality-four subsets of $\{1,\ldots,10\}$

| Matroid | $\{i,j\}$ | $n$ | $r$ | nnz |
|---|---|---|---|---|
| $\mathcal{F}_7^{-4}$ | $\{1,2\}$ | 8 | 5 | 64 |
| $\mathcal{W}^{3+}$ | $\{1,2\}$ | 8 | 5 | 64 |
| $\mathcal{W}^3+e$ | $\{1,2\}$ | 9 | 7 | 81 |
| $\mathcal{P}_7'$ | $\{1,2\}$ | 8 | 4 | 64 |
| $n\mathcal{P}\setminus 1$ | $\{2,4\}$ | 12 | 14 | 144 |
| $n\mathcal{P}\setminus 9$ | $\{1,2\}$ | 12 | 14 | 144 |
| $\mathcal{V}_8$ | $\{1,2\}$ | 16 | 33 | 256 |
| $\mathcal{V}_{10}$ | $\{3,4\}$ | 52 | 657 | 2704 |

**(a)** Original

| Matroid | $\{i,j\}$ | $n$ | $r$ | nnz |
|---|---|---|---|---|
| $\mathcal{F}_7^{-4}$ | $\{1,2\}$ | 5 | 1 | 25 |
| $\mathcal{W}^{3+}$ | $\{1,2\}$ | 3 | 0 | 9 |
| $\mathcal{W}^3+e$ | $\{1,2\}$ | 5 | 0 | 27 |
| $\mathcal{P}_7'$ | $\{1,2\}$ | 4 | 0 | 16 |
| $n\mathcal{P}\setminus 1$ | $\{2,4\}$ | 6 | 0 | 40 |
| $n\mathcal{P}\setminus 9$ | $\{1,2\}$ | 5 | 0 | 27 |
| $\mathcal{V}_8$ | $\{1,2\}$ | 13 | 17 | 185 |
| $\mathcal{V}_{10}$ | $\{3,4\}$ | 41 | 327 | 2087 |

**(b)** Reduced

**Table 2.17:** Dimension $r$ of subspace and order $n$ of cone $\mathbb{S}^n_+$ describing feasible set. The column 'nnz' shows number of nonzero entries of SDP data matrices.

| Matroid | $\{i,j\}$ | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_5$ | $e_6$ | $d_{face}$ |
|---|---|---|---|---|---|---|---|---|
| $\mathcal{F}_7^{-4}$ | $\{1,2\}$ | 2.10e-12 | 0 | 0 | 9.32e-13 | 7.50e-12 | 2.06e-11 | 3.25e-12 |
| $\mathcal{W}^{3+}$ | $\{1,2\}$ | 7.57e-11 | 0 | 0 | 3.52e-11 | 6.47e-11 | 7.29e-10 | 4.41e-06 |
| $\mathcal{W}^3+e$ | $\{1,2\}$ | 8.14e-09 | 0 | 0 | 1.05e-11 | 2.21e-11 | 4.04e-10 | 1.74e-06 |
| $\mathcal{P}_7'$ | $\{1,2\}$ | 2.31e-10 | 0 | 0 | 8.55e-10 | 1.39e-09 | 2.19e-09 | 6.54e-06 |
| $n\mathcal{P}\setminus 1$ | $\{2,4\}$ | 9.04e-10 | 0 | 0 | 1.29e-10 | 2.40e-10 | 9.31e-09 | 4.51e-06 |
| $n\mathcal{P}\setminus 9$ | $\{1,2\}$ | 2.45e-09 | 0 | 0 | 3.54e-10 | 7.07e-10 | 1.87e-08 | 4.15e-08 |
| $\mathcal{V}_8$ | $\{1,2\}$ | 5.29e-11 | 0 | 0 | 8.32e-11 | 1.78e-10 | 6.49e-10 | 4.19e-06 |
| $\mathcal{V}_{10}$ | $\{3,4\}$ | 3.64e-11 | 0 | 0 | 1.40e-09 | 2.73e-09 | 9.78e-09 | 5.37e-06 |

**(a)** Original

| Matroid | $\{i,j\}$ | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_5$ | $e_6$ | $d_{face}$ |
|---|---|---|---|---|---|---|---|---|
| $\mathcal{F}_7^{-4}$ | $\{1,2\}$ | 4.65e-10 | 0 | 0 | 2.16e-08 | 6.61e-08 | 6.86e-08 | 0 |
| $\mathcal{W}^{3+}$ | $\{1,2\}$ | 1.35e-15 | 0 | 0 | 2.91e-12 | 5.55e-12 | 5.55e-12 | 4.57e-16 |
| $\mathcal{W}^3+e$ | $\{1,2\}$ | 8.86e-15 | 0 | 0 | 1.27e-11 | 2.03e-11 | 2.03e-11 | 4.62e-16 |
| $\mathcal{P}_7'$ | $\{1,2\}$ | 6.91e-16 | 0 | 0 | 1.13e-11 | 1.62e-11 | 1.62e-11 | 4.73e-16 |
| $n\mathcal{P}\setminus 1$ | $\{2,4\}$ | 5.27e-11 | 0 | 0 | 1.89e-08 | 3.67e-08 | 3.75e-08 | 3.46e-16 |
| $n\mathcal{P}\setminus 9$ | $\{1,2\}$ | 1.18e-10 | 0 | 0 | 3.18e-08 | 4.29e-08 | 4.35e-08 | 3.44e-16 |
| $\mathcal{V}_8$ | $\{1,2\}$ | 1.43e-11 | 0 | 0 | 4.35e-11 | 7.77e-11 | 1.97e-10 | 0 |
| $\mathcal{V}_{10}$ | $\{3,4\}$ | 2.90e-11 | 0 | 0 | 1.11e-09 | 2.09e-09 | 3.15e-09 | 5.15e-17 |

**(b)** Reduced

**Table 2.18:** DIMACs errors $e_i$ and distance $d_{face}$ to linear span of identified face.

| Matroid | $\{i,j\}$ | $\lvert b^Ty\rvert$ | $\lVert S-\sum_i y_iA_i\rVert_F$ | $\lambda_{\min}(S)$ |
|---|---|---|---|---|
| $\mathcal{F}_7^{-4}$ | $\{1,2\}$ | 0 | 0 | 0 |
| $\mathcal{W}^{3+}$ | $\{1,2\}$ | 2.22e-16 | 0 | 0 |
| $\mathcal{W}^3+e$ | $\{1,2\}$ | 0 | 0 | 0 |
| $\mathcal{P}_7'$ | $\{1,2\}$ | 0 | 0 | 0 |
| $n\mathcal{P}\setminus 1$ | $\{2,4\}$ | 6.66e-16 | 0 | 0 |
| $n\mathcal{P}\setminus 9$ | $\{1,2\}$ | 0 | 0 | 0 |
| $\mathcal{V}_8$ | $\{1,2\}$ | 0 | 0 | 0 |
| $\mathcal{V}_{10}$ | $\{3,4\}$ | 1.78e-15 | 0 | 0 |

**Table 2.19:** Reduction error. The first two columns measure containment of the SDP's affine subspace in the hyperplane $S^\perp$. The last denotes the minimum eigenvalue of the exposing vector $S$.

| Example | Original Primal $n; r$ | Reduced Primal $n; r$ | Original Dual $n; r$ | Reduced Dual $n; r$ |
|---|---|---|---|---|
| `Example1` | $3; 4$ | $2; 2$ | $3; 2$ | $1; 1$ |
| `Example2` | $3; 4$ | $2; 2$ | $3; 2$ | $2; 1$ |
| `Example3` | $3; 2$ | $2; 2$ | $3; 4$ | $2; 2$ |
| `Example4` | $3; 3$ | $1; 0$ | $3; 3$ | $1; 1$ |
| `Example5` | $10; 50$ | $10; 50$ | $10; 5$ | $10; 5$ |
| `Example6` | $8; 28$ | $5; 11$ | $8; 8$ | $4; 4$ |
| `Example7` | $5; 12$ | $4; 8$ | $5; 3$ | $1; 1$ |
| `Example9a` | $100; 4950$ | $1; 0$ | $100; 100$ | $1; 1$ |
| `Example9b` | $20; 190$ | $1; 0$ | $20; 20$ | $1; 1$ |

**Table 2.20:** Complexity parameters for the primal-dual SDP pairs given in [31]. The feasible set of each SDP is an $r$-dimensional subspace intersected with the cone $\mathbb{S}^n_+$. To formulate each reduced SDP, a face was identified by solving LPs over diagonally-dominant approximations ($\mathcal{DD}^d$). These LPs took (in total) $t_{LPs}$ seconds to solve.

excluding

$$\{1,2,6,7\}, \{1,3,6,8\}, \{1,4,6,9\}, \{1,5,6,10\}, \{2,3,7,8\}, \{3,4,8,9\}, \text{ and } \{4,5,9,10\}.$$

From these bases, we construct $f_{\mathcal{V}_{10}}$ via (2.11) and formulate an SDP demonstrating $\Delta_{34} f_{\mathcal{V}_{10}}$ is a sum-of-squares (as was done in [28]).

## ■ 2.8.6 Facial Reduction Benchmark Problems

In [31], Cheung, Schurr, and Wolkowicz developed a facial reduction procedure for identifying faces in a numerically stable manner. They also created a set of benchmark problems for testing their method. These problem instances are available at the URL below:

`http://www.math.uwaterloo.ca/~hwolkowi/henry/reports/SDPinstances.tar`.

Each problem is a primal-dual pair hand-crafted so that both the primal and dual have no strictly feasible solution. We apply our technique to each primal problem and each dual problem individually, using diagonally-dominant ($\mathcal{DD}^d$) approximations. Results are shown in Table 2.20. Since some of the examples have duality gaps, we do not show DIMACs errors nor we do show solve time given the small sizes. We also omit reduction error since multiple facial reduction iterations were performed.

| Example | $n;r$ Original | $n;r$ Reduced |
|---|---|---|
| `CompactDim2R1` | 3;4 | 1;1 |
| `CompactDim2R2` | (6,3,3,3); 25 | (1,0,1,1); 1 |
| `CompactDim2R3` | (10,6,6,6); 91 | (1,0,1,1); 1 |
| `CompactDim2R4` | (15,10,10,10); 241 | (1,0,1,1); 1 |
| `CompactDim2R5` | (21,15,15,15); 526 | (1,0,1,1); 1 |
| `CompactDim2R6` | (28,21,21,21); 1009 | (1,0,1,1); 1 |
| `CompactDim2R7` | (36,28,28,28); 1765 | (1,0,1,1); 1 |
| `CompactDim2R8` | (45,36,36,36); 2881 | (1,0,1,1); 1 |
| `CompactDim2R9` | (55,45,45,45); 4456 | (1,0,1,1); 1 |
| `CompactDim2R10` | (66,55,55,55); 6601 | (1,0,1,1); 1 |

**Table 2.21:** Complexity parameters for weakly-infeasible SDPs studied in [137]. The feasible set of each SDP is an $r$-dimensional subspace intersected with the cone $\mathbb{S}^n_+$. To formulate each reduced SDP, a face was identified by solving LPs defined by diagonal approximations ($\mathcal{D}^d$). These LPs took (in total) $t_{LPs}$ seconds to solve.

## ■ 2.8.7 Difficult SDPs arising in polynomial nonnegativity

In [137] and [140], Waki et al. study two sets of SDPs that are difficult to solve. For one set of SDPs, SeDuMi fails to find certificates of infeasibility [137]. For the other set, SeDuMi reports an incorrect optimal value [140]. The sets of SDPs are available at:

`https://sites.google.com/site/hayatowaki/Home/difficult-sdp-problems`.

It turns out for each primal-dual pair in these sets, the problem defined by equations $A_i \cdot X = b_i$ is not strictly feasible. We apply our technique to both sets of SDPs using diagonal approximations $\mathcal{D}^d$ and arrive at SDPs that are more easily solved. In particular, certificates of infeasibility are found for the SDPs in [137] and correct optimal values are found for the SDPs in [140] by solving the reduced SDPs with SeDuMi. Problem size reductions are shown in Table 2.21 and Table 2.22. We omit solve time comparisons and DIMACs errors since the reduced problem is a trivial SDP in each case. We omit reduction error since multiple facial reduction iterations were performed.

## ■ 2.8.8 DIMACS Controller Design Problems

Our final examples are the controller design problems `hinf12` and `hinf13` of the DIMACS library [104]—which evidently are SDPs in the library with no strictly feasible solution. Results are shown in Tables 2.23-2.25, where we apply facial reduction to

| Example | $n;r$ Original | $n;r$ Reduced | Optimal Value Reduced |
|---|---|---|---|
| unboundDim1R2 | (3,2,2); 8 | (1,1,0); 1 | 1.080478e-13 |
| unboundDim1R3 | (4,3,3); 16 | (1,1,0); 1 | 1.080478e-13 |
| unboundDim1R4 | (5,4,4); 27 | (1,1,0); 1 | 1.080478e-13 |
| unboundDim1R5 | (6,5,5); 41 | (1,1,0); 1 | 1.080478e-13 |
| unboundDim1R6 | (7,6,6); 58 | (1,1,0); 1 | 1.080478e-13 |
| unboundDim1R7 | (8,7,7); 78 | (1,1,0); 1 | 1.080478e-13 |
| unboundDim1R8 | (9,8,8); 101 | (1,1,0); 1 | 1.080478e-13 |
| unboundDim1R9 | (10,9,9); 127 | (1,1,0); 1 | 1.080478e-13 |
| unboundDim1R10 | (11,11,10); 156 | (1,1,0); 1 | 1.080478e-13 |

**Table 2.22:** Complexity parameters for the SDPs in [140]. The feasible set of each SDP is an $r$-dimensional subspace intersected with the cone $\mathbb{S}^n_+$. To formulate each reduced SDP, a face was identified by solving LPs defined by diagonal approximations ($\mathcal{D}^d$). These LPs took (in total) $t_{LPs}$ seconds to solve. For these examples, SeDuMi incorrectly returns an optimal value of one for the original problem. The optimal value returned for the reduced problem is very near the correct optimal value of zero.

the primal problem of both SDPs (using $\mathcal{DD}^d$ for `hinf12` and $\mathcal{SDD}^d$ for `hinf13`). As observed in [93], these problem instances are extremely difficult for SDP solvers. For purposes of comparison, we therefore report DIMACS errors for both SeDuMi and SDPT3 [131]. Solution times are omitted given the small sizes of these SDPs.

## ■ 2.9 Conclusion

We presented a general technique for facial reduction that utilizes approximations of the positive semidefinite cone. The technique is effective on examples arising in practice and for simple approximation is a practical pre-processing routine for SDP solvers. An implementation has been made available. Through the chosen approximation, one controls the cost of facial reduction and the sparsity of the obtained reformulations.

| Matroid | $n$ | $r$ | nnz |
|---|---|---|---|
| `hinf12` | (6, 6, 12) | 77 | 990 |
| `hinf13` | (7, 9, 14) | 121 | 2559 |

**(a)** Original

| Problem | $n$ | $r$ | nnz |
|---|---|---|---|
| `hinf12` | (6, 2, 6) | 23 | 583 |
| `hinf13` | (1, 9, 7) | 45 | 1465 |

**(b)** Reduced

**Table 2.23:** Dimension $r$ of subspace and order $n$ of cone $\mathbb{S}^n_+$ describing feasible set. The column 'nnz' shows number of nonzero entries of SDP data matrices. For `hinf12`, we used $\mathcal{DD}^d$. For `hinf13`, we used $\mathcal{SDD}^d$.

| Problem | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_5$ | $e_6$ | $d_{face}$ |
|---|---|---|---|---|---|---|---|
| `hinf12/sedumi` | 5.04e-09 | 0 | 0 | 0 | -1.55e-02 | 2.23e-01 | 1.17e-08 |
| `hinf13/sedumi` | 6.21e-05 | 0 | 0 | 2.63e-06 | -3.68e-03 | 2.30e-02 | 1.00e+00$^\dagger$ |
| `hinf12/sdpt3` | 1.67e-11 | 0 | 1.72e-05 | 0 | -1.72e-06 | 2.36e-05 | 3.81e-12 |
| `hinf13/sdpt3` | 9.97e-06 | 0 | 5.73e-07 | 0 | -2.35e-04 | 1.94e-04 | 1.43e-02 |

**(a)** Original

| Problem | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_5$ | $e_6$ | $d_{face}$ |
|---|---|---|---|---|---|---|---|
| `hinf12/sedumi` | 4.99e-09 | 0 | 0 | 0 | -5.62e-02 | 2.82e-01 | 0 |
| `hinf13/sedumi` | 6.39e-05 | 0 | 0 | 1.51e-06 | -2.76e-04 | 1.93e-03 | 0 |
| `hinf12/sdpt3` | 1.58e-11 | 0 | 3.18e-06 | 0 | -2.06e-06 | 3.33e-05 | 0 |
| `hinf13/sdpt3` | 3.84e-05 | 0 | 7.09e-08 | 0 | -6.61e-04 | 1.07e-05 | 0 |

**(b)** Reduced

**Table 2.24:** DIMACs errors $e_i$ and distance $d_{face}$ to linear span of identified face. Normalized by solution norm, the outlier, marked $^\dagger$, equals $2.53e-04$.

| Example | $\lvert b^T y\rvert$ | $\lVert S - \sum_i y_i A_i \rVert_F$ | $\lambda_{\min}(S)$ |
|---|---|---|---|
| `hinf12` | 0 | 0 | 0 |
| `hinf13` | 0 | 8.31e-10 | 0 |

**Table 2.25:** Reduction error. The first two columns measure containment of the SDP's affine subspace in the hyperplane $S^\perp$. The last denotes the minimum eigenvalue of the exposing vector $S$.

Chapter 3

# Dual solution recovery

In this chapter we give and study a simple post-processing algorithm for dual solution recovery. Dual solution recovery is critical for primal-dual solvers since they are often agnostic to which problem—primal or dual—is of actual interest. It is necessary because facial reduction *relaxes* the dual problem (Section 1.3.6). Note that simple recovery is not always possible. Indeed, dual solutions may not even exist for the original problem. Hence, we give conditions characterizing success of our procedure.

## ■ 3.1 Approach

To begin, consider the following primal-dual pair

$$\mathbf{P}(\mathcal{K}): \quad \begin{array}{ll} \text{maximize} & \langle b, y\rangle \\ \text{subject to} & s = c - Ay, \\ & (s,y) \in \mathcal{K} \times \mathbb{R}^m, \end{array} \qquad \mathbf{D}(\mathcal{K}): \quad \begin{array}{ll} \text{minimize} & \langle c, x\rangle \\ \text{subject to} & A^*x = b, \\ & x \in \mathcal{K}^*, \end{array} \tag{3.1}$$

where $\mathcal{V}$ is an inner product space, $\mathcal{K} \subseteq \mathcal{V}$ is a closed, convex cone, $\mathcal{K}^*$ is the dual cone, $A : \mathcal{V} \to \mathbb{R}^m$ is a linear map and $c \in \mathcal{V}$. Here, we use a convention opposite to that of Chapter 4 by designating the primal problem as the maximization problem. We use this convention to enable easier comparison with results in the literature.

Applying the facial reduction algorithm to $\mathbf{P}(\mathcal{K})$ yields a face $\mathcal{F}$ of $\mathcal{K}$ that contains the image of $\mathbb{R}^m$ under the affine map $\mathcal{A}(y) := c - Ay$. This in turns yields a reformulation over $\mathcal{F}$ and its dual cone $\mathcal{F}^*$ :

$$\mathbf{P}(\mathcal{F}): \quad \begin{array}{ll} \text{maximize} & \langle b, y\rangle \\ \text{subject to} & s = c - Ay, \\ & (s,y) \in \mathcal{F} \times \mathbb{R}^m, \end{array} \qquad \mathbf{D}(\mathcal{F}): \quad \begin{array}{ll} \text{minimize} & \langle c, x\rangle \\ \text{subject to} & A^*x = b, \\ & x \in \mathcal{F}^*. \end{array}$$

Since the feasible sets of $\mathbf{P}(\mathcal{K})$ and $\mathbf{P}(\mathcal{F})$ are equal, any solution of $\mathbf{P}(\mathcal{F})$ solves $\mathbf{P}(\mathcal{K})$. On the other hand, a solution to $\mathbf{D}(\mathcal{F})$ is not necessarily even a feasible point of $\mathbf{D}(\mathcal{K})$ since $\mathcal{K}^* \subseteq \mathcal{F}^*$. While recovering a solution to $\mathbf{D}(\mathcal{K})$ from a solution to $\mathbf{D}(\mathcal{F})$ may seem hopeless, the facial reduction algorithm (Algorithm 1.1) finds faces

$$\mathcal{K} = \mathcal{F}_0 \supset \mathcal{F}_1 \supset \cdots \supset \mathcal{F}_N = \mathcal{F}$$

and $s_i \in \mathcal{F}_i^* \setminus (\operatorname{span} \mathcal{F})_i^\perp$ orthogonal to $\mathcal{A}(y)$ for all $y$, where $\mathcal{F}_{i+1} = \mathcal{F}_i \cap s_i^\perp$, that may be useful. This leads to the following problem statement:

**Problem 3.1.1** (Recovery of dual solutions)**.** *Given a solution $x$ to $\mathbf{D}(\mathcal{F})$ and $s_i$ satisfying*

$$\begin{array}{rcl} \langle c, s_i \rangle &=& 0 \\ A^* s_i &=& 0 \\ s_i &\in& \mathcal{F}_i^* \setminus (\operatorname{span} \mathcal{F})_i^\perp \\ \mathcal{F}_{i+1} &:=& \mathcal{F}_i \cap s_i^\perp \qquad\qquad (\textit{which implies} \quad \mathcal{F}_{i+1}^* = \overline{\mathcal{F}_i^* + \operatorname{span} s_i}) \\ \mathcal{F}_0 &:=& \mathcal{K}, \ \ \mathcal{F} := \mathcal{F}_N, \end{array}$$

*find a solution to $\mathbf{D}(\mathcal{K})$.*

To solve this problem, we generalize a recovery procedure described in [103] for *well-behaved* SDPs. (See the discussion following [103, Theorem 5].) First, we observe that each $s_i$ is a feasible direction for $\mathbf{D}(\mathcal{F})$ that does not increase the dual objective $\langle c, x \rangle$. We also observe that $\mathcal{F}_{i+1}^* = \overline{\mathcal{F}_i^* + \operatorname{span} s_i}$ (since $\mathcal{K}$, and hence $\mathcal{F}_i$, is closed). This implies if $\mathcal{F}_i^* + \operatorname{span} s_i$ is closed, then one can, for any $x \in \mathcal{F}_{i+1}^*$, find an $\alpha$ such that $x + \alpha s_i \in \mathcal{F}_i^*$. We conclude if $\mathcal{F}_i^* + \operatorname{span} s_i$ is closed for each $i$, then a sequence of line searches, given explicitly by Algorithm 3.1, constructs a solution to $\mathbf{D}(\mathcal{K})$.

**Algorithm 3.1:** Recovery of dual solutions

Input: A solution $x \in \mathcal{F}^*$ to $\mathbf{D}(\mathcal{F})$ and $s_0, \ldots, s_{N-1}$
Output: A solution $x$ to $\mathbf{D}(\mathcal{K})$ or flag indicating failure.
**for** $i \leftarrow N-1$ down to 0 **do**

1. Using a line search, find $\alpha$ s.t. $x + \alpha s_i \in \mathcal{F}_i^*$.
2. If no $\alpha$ exists, return FAIL. Else, set $x \leftarrow x + \alpha s_i$.

**end**

The following properties of this algorithm are immediate:

**Lemma 3.1.1.** *Algorithm 3.1 has the following properties:*

1. Sufficient condition for recovery. *Algorithm 3.1 succeeds if $\mathcal{F}_i^* + \operatorname{span} s_i$ is closed for all $i$.*

*2.* Necessary condition for recovery. *Suppose* $\mathbf{P}(\mathcal{F})$ *and* $\mathbf{D}(\mathcal{F})$ *have equal optimal values. Then, Algorithm 3.1 succeeds only if* $\mathbf{P}(\mathcal{K})$ *and* $\mathbf{D}(\mathcal{K})$ *have equal optimal values.*

When $\mathcal{K}$ is polyhedral, the above sufficient condition *always* holds. On the other hand, it rarely holds for SDP: for $S \in \mathbb{S}^n_+$, the set $\mathbb{S}^n_+ + \operatorname{span} S$ is closed only if $S$ is zero or positive definite. (To show this, one can use essentially the same argument that proves Lemma 2.2 of [115]. This lemma shows that for a face $\mathcal{F}$ of $\mathbb{S}^n_+$, the set $\mathbb{S}^n_+ + \operatorname{lin} \mathcal{F}$ is closed only if $\mathcal{F} = \{0\}$ or $\mathcal{F} = \mathbb{S}^n_+$.) Hence, better sufficient conditions for SDP are needed.

In the next section, we give a necessary and sufficient condition (Condition 3.2.1) for SDP assuming $N = 1$, i.e., assuming one iteration of facial reduction was performed. This condition is on the specific solution of $\mathbf{D}(\mathcal{F})$ used to initialize Algorithm 3.1. We also give a sufficient condition valid for arbitrary initialization points (Condition 3.4.1).

**Remark 3.1.1.** *Closedness of* $\mathcal{K}^* + \operatorname{span} s$ *for* $s \in \mathcal{K}^*$ *appears in other contexts. Borwein and Wolkowicz use this condition to simplify their generalized optimality conditions for convex programs (see Remark 6.2 of [20]). Tunçel and Wolkowicz use failure of a related condition, namely closedness of* $\mathcal{K}^* + \operatorname{lin} \mathcal{F}$ *for a face* $\mathcal{F}$ *of* $\mathcal{K}^*$*, to construct primal-dual pairs with infinite duality gaps.*

## ■ 3.2 A necessary and sufficient condition for dual solution recovery

In this section, we give a necessary and sufficient condition (Condition 3.2.1) for solution recovery assuming $\mathcal{K} = \mathbb{S}^n_+$ and one iteration of facial reduction. In this case, the primal-dual pair is

$$\mathbf{P}_{SDP}(\mathbb{S}^n_+): \quad \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & C - \sum_{i=1}^m y_i A_i \in \mathbb{S}^n_+, \end{array}$$

$$\mathbf{D}_{SDP}(\mathbb{S}^n_+): \quad \begin{array}{lll} \text{minimize} & \operatorname{Tr} CX & \\ \text{subject to} & \operatorname{Tr} A_i X = b_i & \forall i \in \{1, \ldots, m\}, \\ & X \in \mathbb{S}^n_+, & \end{array}$$

and its reformulation over $\mathcal{F} := \mathbb{S}^n_+ \cap S^\perp$ and $\mathcal{F}^* = \overline{\mathbb{S}^n_+ + \operatorname{span} S}$ is

$$\mathbf{P}_{SDP}(\mathcal{F}): \quad \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & \mathcal{A}(y) = C - \sum_{i=1}^m y_i A_i, \\ & U^T \mathcal{A}(y) U \in \mathbb{S}^d_+, \\ & U^T \mathcal{A}(y) V = 0, \\ & V^T \mathcal{A}(y) V = 0, \end{array}$$

$$\mathbf{D}_{SDP}(\mathcal{F}): \quad \begin{array}{ll} \text{minimize} & \operatorname{Tr} CX \\ \text{subject to} & \operatorname{Tr} A_i X = b_i \qquad \forall i \in \{1, \dots, m\}, \\ & X = (U, V) \begin{pmatrix} W & Z \\ Z^T & R \end{pmatrix} (U, V)^T, \\ & W \in \mathbb{S}^d_+, R \in \mathbb{S}^{n-d}, Z \in \mathbb{R}^{d \times (n-d)}, \end{array}$$

where $(U, V)$ is an invertible matrix satisfying $S = VV^T$ and $\operatorname{range} U = \operatorname{null} S$. Note to reformulate the primal-dual pair, we have used the facts

$$\mathbb{S}^n_+ \cap S^\perp = \{X \in \mathbb{S}^n_+ : XS = 0\} = \{X \in \mathbb{S}^n_+ : XV = 0\} = \left\{X \in \mathbb{S}^n_+ : (U, V)^T XV = 0\right\}.$$

Algorithm 3.1 constructs a solution to $\mathbf{D}_{SDP}(\mathbb{S}^n_+)$ from a solution $X$ to $\mathbf{D}_{SDP}(\mathcal{F})$ if and only if $X$ is in $\mathbb{S}^n_+ + \operatorname{span} S$. The following shows this is equivalent to the condition that $\operatorname{null} W \subseteq \operatorname{null} Z^T$. We give a direct proof of this fact, but note it also follows (essentially) by combining [103, Lemma 3] with [101, Lemma 3.2.1].

**Lemma 3.2.1.** *Given a face $\mathcal{F} = \mathbb{S}^n_+ \cap S^\perp$, let $(U, V) \in \mathbb{R}^{n \times n}$ be an invertible matrix satisfying* $\operatorname{range} V = \operatorname{range} S$*. A matrix $X$ in $\mathcal{F}^* = \overline{\mathbb{S}^n_+ + \operatorname{span} S}$, i.e., a matrix $X$ of the form*

$$X = (U, V) \begin{pmatrix} W & Z \\ Z^T & R \end{pmatrix} (U, V)^T \qquad \text{for some } W \in \mathbb{S}^d_+, R \in \mathbb{S}^{n-d}, Z \in \mathbb{R}^{d \times n-d}, \tag{3.2}$$

*is in $\mathbb{S}^n_+ + \operatorname{span} S$ if and only if* $\operatorname{null} W \subseteq \operatorname{null} Z^T$.

*Proof.* For the "only if" direction, suppose $X$ is in $\mathbb{S}^n_+ + \operatorname{span} S$, i.e. for an $\alpha \in \mathbb{R}$ suppose

$$X + \alpha VV^T = (U, V) \begin{pmatrix} W & Z \\ Z^T & R + \alpha I \end{pmatrix} (U, V)^T \in \mathbb{S}^n_+.$$

Here, membership in $\mathbb{S}^n_+$ holds only if $Z^T(I - WW^\dagger) = 0$, where $(I - WW^\dagger)$ is the orthogonal projector onto $\operatorname{null} W$ (see, e.g. A.5 of [21]). But this implies that $\operatorname{null} W \subseteq \operatorname{null} Z^T$, as desired.

To see the converse direction, suppose $X$ is such that $Z$ and $W$ satisfy $\operatorname{null} W \subseteq$

$\operatorname{null} Z^T$. The result follows by finding $\alpha$ for which $X + \alpha S \succeq 0$. To do this, we show existence of $\alpha_1$ and $\alpha_2$ for which

$$X - VRV^T + \alpha_1 S \in \mathbb{S}^n_+ \qquad \text{and} \qquad VRV^T + \alpha_2 S \in \mathbb{S}^n_+.$$

Adding these two matrices then shows that $X + (\alpha_1 + \alpha_2)S \in \mathbb{S}^n_+$.

Existence of $\alpha_2$ is obvious given that

$$VRV^T + \alpha_2 S = V(R + \alpha_2 I)V^T.$$

To show existence of $\alpha_1$, we note that

$$X - VRV^T + \alpha_1 S = (U, V) \begin{pmatrix} W & Z \\ Z^T & \alpha_1 I \end{pmatrix} (U, V)^T.$$

By a Schur complement, the above is psd if and only if

$$W - \frac{1}{\alpha_1} ZZ^T \in \mathbb{S}^d_+.$$

But since $\operatorname{null} W \subseteq \operatorname{null} Z^T$, the matrix $ZZ^T$ is contained in the face

$$\mathcal{G} = \left\{ T \in \mathbb{S}^d_+ : \operatorname{range} T \subseteq \operatorname{range} W \right\}.$$

Since $W$ is in the relative interior of $\mathcal{G}$, there exists $\alpha_1 > 0$ for which $W - \frac{1}{\alpha_1} ZZ^T \in \mathcal{G} \subseteq \mathbb{S}^d_+$ (Proposition 1.3.1).

□

The above characterization of $\mathbb{S}^n_+ + \operatorname{span} S$ yields a necessary and sufficient condition for success of Algorithm 3.1 assuming one iteration of facial reduction ($N = 1$):

**Condition 3.2.1.** *The solution $X$ to $\mathbf{D}_{SDP}(\mathcal{F})$ satisfies* $\operatorname{null} W \subseteq \operatorname{null} Z^T$.

The following example illustrates success and failure of Condition 3.2.1.

**Example 3.2.1.** *Consider the following primal-dual pair:*

$$\begin{array}{lll} \text{max.} & y_3 + 2y_2 & \\ \text{subj. to} & & \\ & \mathcal{A}(y) = \begin{pmatrix} y_1 & y_2 & 0 \\ y_2 & -y_3 & y_2 \\ 0 & y_2 & y_3 \end{pmatrix} \in \mathbb{S}^3_+ & \end{array} \qquad \begin{array}{ll} \text{min.} & 0 \\ \text{subj. to} & x_{33} - x_{22} = -1, \\ & x_{12} + x_{21} + x_{23} + x_{32} = -2, \\ & x_{11} = 0, \\ & X \in \mathbb{S}^3_+ \end{array}$$

*and let $S = VV^T$, with $V = (e_2, e_3)$. Clearly $S$ is orthogonal to $\mathcal{A}(y)$ for all $y$ and exposes a face $\mathcal{F} := \mathbb{S}^n_+ \cap S^\perp$ equal to $U\mathbb{S}^1_+ U^T$ for $U = e_1 = (1, 0, 0)^T$. Rewriting the* *primal-dual pair over $\mathcal{F}$ and $\mathcal{F}^*$ gives:*

$$\begin{array}{llcll}
\text{max.} & y_3 + 2y_2 & & \text{min.} & 0 \\
\text{subj. to} & & & \text{subj. to} & x_{33} - x_{22} = -1, \\
& V^T \mathcal{A}(y) V = 0, & & & x_{12} + x_{21} + x_{23} + x_{32} = -2, \\
& U^T \mathcal{A}(y) V = 0, & & & x_{11} = 0, \\
& U^T \mathcal{A}(y) U \geq 0, & & & X \in \mathbb{S}^3, U^T X U = x_{11} \geq 0. \\
& \mathcal{A}(y) = \begin{pmatrix} y_1 & y_2 & 0 \\ y_2 & -y_3 & y_2 \\ 0 & y_2 & y_3 \end{pmatrix} & & &
\end{array}$$

*A solution of the dual problem satisfying Condition 3.2.1 is*

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

*To see that the condition holds, note $Z = (x_{12}, x_{13}) = (0,0)$ and $W = x_{11} = 0$. Hence,* $\operatorname{null} Z^T$ *contains (indeed, equals)* $\operatorname{null} W$*. We therefore see that solution recovery succeeds, i.e. for (say) $\alpha = 2$:*

$$X + \alpha S = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \alpha & -1 \\ 0 & -1 & \alpha - 1 \end{pmatrix} \in \mathbb{S}^3_+.$$

*On the other hand, the following solution fails Condition 3.2.1:*

$$X = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{3.3}$$

*Here, $Z = (-1, 0)$ and $W = 0$. Hence,* $\operatorname{null} Z^T = \{0\}$ *does not contain* $\operatorname{null} W = \mathbb{R}$ *and recovery must fail. In other words, there is no $\alpha$ for which*

$$X + \alpha S = \begin{pmatrix} 0 & -1 & 0 \\ -1 & \alpha & 0 \\ 0 & 0 & \alpha - 1 \end{pmatrix} \in \mathbb{S}^3_+,$$

*which is easily seen.*

## ■ 3.3 Strong duality is not sufficient for dual recovery

Recall that zero duality gap between the original primal-dual pair $\mathbf{P}(\mathcal{K})$ and $\mathbf{D}(\mathcal{K})$ is a necessary condition for successful recovery when $\mathbf{P}(\mathcal{F})$ and $\mathbf{D}(\mathcal{F})$ have no gap

(Lemma 3.1.1). Example 3.2.1 shows this is *not* a sufficient condition: both primal-dual pairs have no gap.

**Corollary 3.3.1.** *The dual solution recovery procedure of Algorithm 3.1 can fail even if both the original primal-dual pair* $\mathbf{P}(\mathcal{K})$ *and* $\mathbf{D}(\mathcal{K})$ *and the primal-dual pair* $\mathbf{P}(\mathcal{F})$ *and* $\mathbf{D}(\mathcal{F})$ *have zero duality gap.*

## ■ 3.4 Ensuring successful dual recovery

Condition 3.2.1 lets one determine if recovery is possible by a simple null space computation. Unfortunately, this check must be done *after* solving the dual problem $\mathbf{D}_{SDP}(\mathcal{F})$ as it depends on the obtained solution. In this section, we give a simple sufficient condition that is checkable *prior* to solving $\mathbf{D}_{SDP}(\mathcal{F})$. If this condition holds, one can modify $\mathbf{P}_{SDP}(\mathcal{F})$ and $\mathbf{D}_{SDP}(\mathcal{F})$ to guarantee successful recovery, independent of the solution found for $\mathbf{D}_{SDP}(\mathcal{F})$.

The idea is simple: when can one assume $Z = 0$ and hence ensure Condition 3.2.1 holds without changing the dual optimal value? By [103, Theorem 5], this assumption can be made if $\mathbf{P}_{SDP}(\mathcal{F})$ satisfies Slater's condition and $\mathbf{P}_{SDP}(\mathbb{S}^n_+)$ is *well-behaved*—where $\mathbf{P}_{SDP}(\mathbb{S}^n_+)$ is well-behaved if, for *all* cost vectors $b$, the optimal values of $\mathbf{P}_{SDP}(\mathbb{S}^n_+)$ and $\mathbf{D}_{SDP}(\mathbb{S}^n_+)$ are equal and $\mathbf{D}_{SDP}(\mathbb{S}^n_+)$ attains it optimal value when it is finite. It turns out we can assume $Z = 0$ under a related but purely linear-algebraic condition inspired by a characterization of well-behaved SDPs [103, Theorem 3]. The condition and statement follow.

**Condition 3.4.1.** *The equations of* $\mathbf{P}_{SDP}(\mathcal{F})$ *have the following property:*

$$\left\{y \in \mathbb{R}^m : V^T\mathcal{A}(y)V = 0\right\} = \left\{y \in \mathbb{R}^m : V^T\mathcal{A}(y)V = 0, V^T\mathcal{A}(y)U = 0\right\},$$

*that is,* $V^T\mathcal{A}(y)V = 0$ *implies* $V^T\mathcal{A}(y)U = 0$.

**Proposition 3.4.1.** *Suppose Condition 3.4.1 holds. If* $\mathbf{D}_{SDP}(\mathcal{F})$ *has an optimal solution, then it has an optimal solution with* $Z = 0$.

*Proof.* Let $X$ be an optimal solution to $\mathbf{D}_{SDP}(\mathcal{F})$, which, for some $W \in \mathbb{S}^d_+$, $Z \in \mathbb{R}^{d\times(n-d)}$, and $R \in \mathbb{S}^{n-d}$ satisfies

$$X = (U, V)\begin{pmatrix} W & Z \\ Z^T & R \end{pmatrix}(U, V)^T. \tag{3.4}$$

We will construct a new solution $\hat{X}$ by setting $Z$ to zero and replacing $R$ with $R + \hat{R}$ for a particular $\hat{R}$.

Towards this, we first show existence of $X$ implies the set $\left\{y \in \mathbb{R}^m : V^T\mathcal{A}(y)V = 0\right\}$ is non-empty. If it were empty, then, by Farkas lemma, there exists $\tilde{R}$ satisfying

$\operatorname{Tr} \tilde{R}(V^T A_i V) = 0$ and $\operatorname{Tr} \tilde{R}(V^T C V) < 0$, which implies

$$\tilde{X} = X + (U, V) \begin{pmatrix} 0 & 0 \\ 0 & \tilde{R} \end{pmatrix} (U, V)^T$$

is a feasible point of $\mathbf{D}_{SDP}(\mathcal{F})$ with strictly better cost, contradicting optimality of $X$. Hence, there exists $y_0 \in \Big\{y \in \mathbb{R}^m : V^T \mathcal{A}(y) V = 0\Big\}$.

Now, consider the linear maps $L_1 : \mathbb{R}^m \to \mathbb{S}^{n-d}$ and $L_2 : \mathbb{R}^m \to \mathbb{R}^{d\times(n-d)}$

$$L_1(y) = \sum_{i=1}^m y_i(V^T A_i V), \quad L_2(y) = \sum_{i=1}^m y_i(U^T A_i V),$$

where $\mathbb{R}^{d\times(n-d)}$ and $\mathbb{S}^{n-d}$ are equipped with trace inner-product. With these definitions, $L_1(y) = V^T CV$ if and only if $V^T \mathcal{A}(y) V = 0$ and $L_2(y) = U^T CV$ if and only if $V^T \mathcal{A}(y) U = 0$. Further, $X$ of form (3.4) satisfies the equations $\operatorname{Tr} A_i X = b_i$ for $\mathbf{D}_{SDP}(\mathcal{F})$ if and only if

$$L_1^*(R) + 2L_2^*(Z) = b - \Big(\operatorname{Tr}(U^T A_1 U)W, \ldots, \operatorname{Tr}(U^T A_m U)W\Big)^T.$$

Now suppose Condition 3.4.1 holds. Given existence of $y_0$, it follows that $\operatorname{null} L_1 \subseteq \operatorname{null} L_2$—otherwise, we could construct solutions to $L_1(y) = V^T CV$ that do not solve $L_2(y) = U^T CV$, a contradiction of Condition 3.4.1. But $\operatorname{null} L_1 \subseteq \operatorname{null} L_2$ holds if and only if $\operatorname{range} L_1^* \supseteq \operatorname{range} L_2^*$. Hence, we can find a $\hat{R}$ satisfying

$$sL_1^*(\hat{R}) = 2L_2^*(Z), \tag{3.5}$$

which implies the matrix

$$\hat{X} = (U, V) \begin{pmatrix} W & 0 \\ 0 & R + \hat{R} \end{pmatrix} (U, V)^T$$

satisfies $\operatorname{Tr} A_i \hat{X} = b_i$. Since $W \in \mathbb{S}^d_+$, it follows $\hat{X}$ is feasible for $\mathbf{D}_{SDP}(\mathcal{F})$.

We now show $\operatorname{Tr} CX = \operatorname{Tr} C\hat{X}$, proving $\hat{X}$ is also optimal. For this, it suffices to show $\operatorname{Tr} CV\hat{R}V^T = 2 \operatorname{Tr} CUZV^T$. Since $L_1(y_0) = V^T CV$ and, by Condition 3.4.1, $L_2(y_0) = U^T CV$, we conclude

$$\langle L_1^*(\hat{R}), y_0\rangle = \langle \hat{R}, L_1(y_0)\rangle = \langle \hat{R}, V^T CV\rangle = \operatorname{Tr} V^T CV\hat{R},$$
$$\langle L_2^*(Z), y_0\rangle = \langle Z, L_2(y_0)\rangle = \langle Z, U^T CV\rangle = \operatorname{Tr} V^T CUZ,$$

which, by (3.5), shows $\operatorname{Tr} CV\hat{R}V^T = 2 \operatorname{Tr} CUZV^T$, as desired. □

We conclude one can fix $Z$ to zero in $\mathbf{D}_{SDP}(\mathcal{F})$ and omit the equations $V^T \mathcal{A}(y) U = 0$

from $\mathbf{P}_{SDP}(\mathcal{F})$ under Condition 3.4.1. This leads to a modified primal-dual pair:

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & \mathcal{A}(y) = C - \sum_{i=1}^m y_i A_i, \\ & U^T \mathcal{A}(y) U \in \mathbb{S}^d_+, \\ & V^T \mathcal{A}(y) V = 0, \end{array} \qquad \begin{array}{ll} \text{minimize} & \operatorname{Tr} CX \\ \text{subject to} & \operatorname{Tr} A_i X = b_i \qquad \forall i \in \{1, \ldots, m\}, \\ & X = (U, V) \begin{pmatrix} W & 0 \\ 0 & R \end{pmatrix} (U, V)^T, \\ & W \in \mathbb{S}^d_+, R \in \mathbb{S}^{n-d}, \end{array}$$

where any solution of the modified primal solves $\mathbf{P}_{SDP}(\mathbb{S}^n_+)$ and any solution of the modified dual satisfies Condition 3.2.1 (by construction).

**Comparison with well-behavedness** We now illustrate differences between Condition 3.4.1 and well-behavedness of $\mathbf{P}_{SDP}(\mathbb{S}^n_+)$. Suppose one constructs $\mathbf{P}_{SDP}(\mathcal{F})$ after a single iteration of facial reduction. From [103, Theorem 3], it follows Condition 3.4.1 and well-behavedness of $\mathbf{P}_{SDP}(\mathbb{S}^n_+)$ are equivalent if $\mathbf{P}_{SDP}(\mathcal{F})$ satisfies Slater's condition. The following examples show this equivalence can fail if Slater's condition does not hold.

**Example 3.4.1** (A well-behaved SDP and failure of Condition 3.4.1). *Consider the following SDP:*

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & \\ & \mathcal{A}(y) = \begin{pmatrix} y_1 & 0 & y_2 & 0 \\ 0 & -y_1 & 0 & 0 \\ y_2 & 0 & y_2 & 0 \\ 0 & 0 & 0 & -y_2 \end{pmatrix} \in \mathbb{S}^4_+. \end{array}$$

*For any $b$, the optimal value is zero. Further, for any $b$, a nonnegative diagonal matrix $X$ satisfying $b_1 = x_{22} - x_{11}$ and $b_2 = x_{44} - x_{33}$ is dual optimal. Hence, the SDP is well-behaved.*

*The matrix $S = e_1 e_1 + e_2 e_2^T$ is orthogonal to $\mathcal{A}(y)$ for all $y$. Hence, we can take $\mathcal{F} = \mathbb{S}^4_+ \cap S^\perp$ and formulate $\mathbf{P}_{SDP}(\mathcal{F})$ as follows*

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & \\ & U^T \mathcal{A}(y) U = \begin{pmatrix} y_2 & 0 \\ 0 & -y_2 \end{pmatrix} \in \mathbb{S}^2_+, \qquad V^T \mathcal{A}(y) U = \begin{pmatrix} y_2 & 0 \\ 0 & 0 \end{pmatrix} = 0_{2\times 2}, \\ & V^T \mathcal{A}(y) V = \begin{pmatrix} y_1 & 0 \\ 0 & -y_1 \end{pmatrix} = 0_{2\times 2}, \end{array}$$

*where $V = (e_1, e_2)$, $U = (e_3, e_4)$. Clearly $U^T \mathcal{A}(y) U$ cannot be positive definite, hence*

$\mathbf{P}_{SDP}(\mathcal{F})$ *fails Slater's condition. In addition, Condition 3.4.1 fails, i.e.,*

$$\left\{y \in \mathbb{R}^m : V^T \mathcal{A}(y) V = 0\right\} \neq \left\{y \in \mathbb{R}^m : V^T \mathcal{A}(y) V = 0, V^T \mathcal{A}(y) U = 0\right\}.$$

*Note one avoids this failure by taking $S = I$, which is also orthogonal to $\mathcal{A}(y)$ for all $y$.*

**Example 3.4.2** (An SDP not well-behaved and success of Condition 3.4.1)**.** *The following SDP, based on [103, Example 1], has an unattained dual optimal value when $b = (1, 0, 0)^T$ and is hence not well-behaved:*

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & \\ & \mathcal{A}(y) = \begin{pmatrix} 1 & -y_1 & 0 & 0 \\ -y_1 & y_3 & 0 & 0 \\ 0 & 0 & y_2 & y_3 \\ 0 & 0 & y_3 & -y_2 \end{pmatrix} \in \mathbb{S}^4_+. \end{array}$$

*For the rank two matrix $S = e_3 e_3 + e_4 e_4^T$, we have that $\mathbf{P}_{SDP}(\mathcal{F})$ takes the form:*

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & \\ & U^T \mathcal{A}(y) U = \begin{pmatrix} 1 & -y_1 \\ -y_1 & y_3 \end{pmatrix} \in \mathbb{S}^2_+, \quad V^T \mathcal{A}(y) U = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0_{2\times 2}, \\ & V^T \mathcal{A}(y) V = \begin{pmatrix} y_2 & y_3 \\ y_3 & -y_2 \end{pmatrix} = 0_{2\times 2}, \end{array}$$

*where $V = (e_3, e_4)$ and $U = (e_1, e_2)$. Since $y_3 = 0$ if $y$ is feasible, Slater's condition fails. Condition 3.4.1, on the other hand, holds given that $V^T \mathcal{A}(y) U = 0$ imposes no constraints on $y$. This is despite the fact the SDP is not well-behaved.*

## 3.5 Recovering solutions to an extended dual

We close by discussing recovery for an alternative dual program intimately related to facial reduction—an *extended dual* [102, 114, 115]. Solutions to the extended dual simultaneously identify a chain of faces $\mathcal{F}_1, \ldots, \mathcal{F}_N$ (where $N = n - 1$) and a solution $X \in \mathcal{F}_N^*$ to $\mathbf{D}_{SDP}(\mathcal{F}_N)$. The description of faces relies on the following.

**Lemma 3.5.1.** *The following statements are true:*

1. *For any face $\mathcal{F}$ of $\mathbb{S}^n_+$, $\mathcal{F}^* = \mathbb{S}^n_+ + (\operatorname{span} \mathcal{F})^\perp$.*
2. *If $\mathcal{F} = \mathbb{S}^n_+ \cap S^\perp$ for $S \in \mathbb{S}^n_+$, then*

$$(\operatorname{span} \mathcal{F})^\perp = \left\{ W + W^T : \begin{pmatrix} S & W \\ W^T & \alpha I \end{pmatrix} \in \mathbb{S}^{2n}_+ \text{ for some } \alpha \in \mathbb{R} \right\}.$$

3. *Let* $\mathcal{F}_0 := \mathbb{S}^n_+$ *and consider the chain of faces defined by matrices* $S_i$

$$\mathcal{F}_{i+1} := \mathcal{F}_i \cap S_i^\perp,$$

*where* $S_i$ *is in* $\mathcal{F}_i^*$, *i.e.* $S_i = \bar{S}_i + V_i$ *for* $\bar{S}_i \in \mathbb{S}^n_+$ *and* $V_i \in (\operatorname{span} \mathcal{F})_i^\perp$. *The following relationship holds:*

$$\mathcal{F}_{i+1} = \mathbb{S}^n_+ \cap (\sum_{j=0}^{i} \bar{S}_j)^\perp.$$

*Proof.* The first statement holds because $\mathbb{S}^n_+$ is a *nice* cone [102]. The other statements are shown by Proposition 1 and Theorem 3 of [102]. □

The variables and constraints of the extended dual arise directly from this lemma. It is given below as an optimization problem over $X, \bar{X}, W_N, S_i, \bar{S}_i, W_i, \alpha_i$:

$$\begin{array}{lrl}
\text{minimize} & \operatorname{Tr} CX & \\
\text{subject to} & \operatorname{Tr} A_j X = b_j & \\
 & \operatorname{Tr} CS_i = 0,\ \operatorname{Tr} A_j S_i = 0 & (S_i^\perp \text{ contains } \mathcal{A}(y) \text{ for all } y) \\
 & X = \bar{X} + W_N + W_N^T & (X \in \mathcal{F}_N^*) \\
 & S_i = \bar{S}_i + W_i + W_i^T & (S_i \in \mathcal{F}_i^*) \\
 & \begin{pmatrix} \sum_{j=0}^i \bar{S}_j & W_{i+1} \\ W_{i+1}^T & \alpha_i I \end{pmatrix} \in \mathbb{S}^{2n}_+ & \\
 & \bar{S}_i \in \mathbb{S}^n_+, \bar{X} \in \mathbb{S}^n_+, W_0 = 0, &
\end{array}$$

where $i$ ranges from 0 to $N-1$ and $j$ ranges from 1 to $m$ (indexing $m$ linear equations $\operatorname{Tr} A_j X = b_j$).

**Recovering a solution** Suppose $\mathcal{F}_i = U_i \mathbb{S}^{d_i}_+ U_i^T$ for $i = 0, \ldots, N$ is a sequence of faces identified by facial reduction suitably padded so that the length of the sequence is $N$, i.e., $\mathcal{F}_0, \ldots, \mathcal{F}_M = \mathbb{S}^n_+$ for some $M < N$. Let $S_i \in \mathcal{F}_i^*$ be the exposing vector of $\mathcal{F}_{i+1}$ and let $X$ be a solution to $\mathbf{D}_{SDP}(\mathcal{F})$. One can construct a feasible point of the extended dual by decomposing $S_i$ (and similarly $X$) into the form $S_i = \bar{S}_i + W_i + W_i^T$, for $\bar{S}_i \in \mathbb{S}^n_+$ and $W_i + W_i^T \in (\operatorname{span} \mathcal{F}_i)^\perp$. Supposing $U_i$ has orthonormal columns, this is done by taking

$$\begin{array}{lll}
\bar{S}_i = U_i U_i^T S_i U_i U_i^T, & W_i = \frac{1}{2}(S - \bar{S}_i) & \forall i \in \{0, \ldots, N-1\}, \\
\bar{X} = U_N U_N^T X U_N U_N^T, & W_N = \frac{1}{2}(X - \bar{X}). &
\end{array}$$

One can then pick each $\alpha_i$ (individually) to satisfy the corresponding semidefinite constraint.

## ■ 3.6 Conclusion

We gave a post-processing procedure for dual solution recovery that applies generally to cone programs preprocessed using facial reduction. This recovery procedure always succeeds when the cone is polyhedral, but may fail otherwise, illustrating an interesting difference between linear programming and optimization over nonpolyhedral cones. We gave sufficient conditions for successful recovery and explored the connection with well-behaved SDPs—a subset of SDPs for which recovery is always possible.

Chapter 4

# Self-dual embeddings and facial reduction

Recall that a cone program is *pathological* if none of the following three objects exist: a solution to the optimality conditions, a feasible point paired with an improving ray, or a strictly separating hyperplane, i.e., a dual improving ray. By applying facial reduction to a cone program and/or its dual, we can guarantee the resulting cone program is not pathological. There is no algorithm, however, that does facial reduction *only if* a given program is pathological. Such an algorithm would be useful if one only wanted to pay the costs of facial reduction for pathological instances which, without facial reduction, are perhaps unsolvable. This chapter provides such an algorithm. Indeed, this algorithm actually solves arbitrary instances of cone programs, performing facial reduction if and only if an instance is pathological before returning the optimal value and a point attaining it if one exists.

Note that the facial reduction algorithms [20, 102, 138] are based on failure of Slater's condition. Failure of Slater's condition is necessary for pathologies but not sufficient. Hence, our algorithm is based on a different object: the self-dual embedding, an auxiliary cone program constructed from a given instance. When the instance isn't pathological, solutions of the embedding provide solutions of the optimality conditions or improving rays. When the instance is pathological, we will show solutions to the embedding provide the needed hyperplanes for facial reduction. Note that the self-dual embedding is the basis of widely used solvers [128, 94]. Indeed, implementing our algorithm involves only minimal code changes to these solvers that only modify execution for pathological instances. Nevertheless, numerical experiments will show such implementations still face the significant numerical challenges encountered by other facial reduction algorithms. Hence, the contributions of this chapter are, at this point, mostly theoretical. We now overview the self-dual embedding and explain contributions of this chapter in more detail. First we fix the form of the cone program of interest.

**Problem of interest** In this chapter, the cone program of interest has decision variable $x \in \mathcal{V}$ and its dual has decision variables $s \in \mathcal{V}$ and $y \in \mathbb{R}^m$:

$$\begin{array}{llcll} \text{minimize} & \langle c, x\rangle & \qquad & \text{maximize} & \langle b, y\rangle \\ \text{subject to} & Ax = b, & & \text{subject to} & c - A^* y = s, \\ & x \in \mathcal{K}, & & & s \in \mathcal{K}^*,\ y \in \mathbb{R}^m. \end{array} \tag{4.1}$$

As in previous chapters, $\mathcal{V}$ is a finite-dimensional inner product space identified with its dual space $\mathcal{V}^*$, the map $A : \mathcal{V} \to \mathbb{R}^m$ is linear with adjoint $A^* : \mathbb{R}^m \to \mathcal{V}$, the points $b \in \mathbb{R}^m$ and $c \in \mathcal{V}$ are fixed, and the cone $\mathcal{K} \subseteq \mathcal{V}$ is closed and convex with dual cone $\mathcal{K}^* := \{s \in \mathcal{V} : \langle s, x\rangle \geq 0 \ \ \forall x \in \mathcal{K}\}$.

Recall from Chapter 1.2.1 that the optimality conditions of (4.1) are

$$Ax = b,\ x \in \mathcal{K}, \qquad s = c - A^T y,\ s \in \mathcal{K}^*, \qquad \langle c, x\rangle = b^T y,$$

where we've used the fact that $\langle c, x\rangle = b^T y$ is, for feasible points, equivalent to the complementary slackness condition $\langle s, x\rangle = 0$. Given this equivalence, we will call $(x, s, y) \in \mathcal{V} \times \mathcal{V} \times \mathbb{R}^m$ a *complementary solution* if it solves the optimality conditions. Similarly, a *primal improving ray* $x \in \mathcal{V}$ and *dual improving ray* $(s, y) \in \mathcal{V} \times \mathbb{R}^m$ are points satisfying

$$Ax = 0,\ x \in \mathcal{K},\ \langle c, x\rangle < 0 \qquad s = -A^* y,\ s \in \mathcal{K}^*,\ b^T y > 0.$$

Also recall from Chapter 1.2.1 that the hyperplane $\{x \in \mathcal{V} : \langle s, x\rangle = 0\}$ strictly separates $\mathcal{K}$ from the solutions to $Ax = b$ if $(s, y)$ is a dual improving ray. Hence, dual improving rays certify primal infeasibility. Similar remarks apply to primal improving rays and dual infeasibility.

**Self-dual embeddings** The *self-dual embedding*, also called the *self-dual homogeneous model*, is a cone program whose constraints simultaneously describe improving rays and complementary solutions [63, 87, 40, 97, 111]. It includes two extra variables $\tau \in \mathbb{R}_+$ and $\kappa \in \mathbb{R}_+$ and takes the following form

$$\begin{array}{rcl} Ax - b\tau & = & 0, \\ -A^* y - s + c\tau & = & 0, \\ \langle b, y\rangle - \langle c, x\rangle - \kappa & = & 0, \\ (x, s, y, \tau, \kappa) \in \mathcal{K} \times \mathcal{K}^* \times \mathbb{R}^m \times \mathbb{R}_+ \times \mathbb{R}_+. & & \end{array} \tag{4.2}$$

Any solution $(x, s, y, \tau, \kappa)$ of the embedding satisfies the complementarity condition $\tau\kappa = 0$; hence, any solution falls into one of three categories. If $\tau > 0$, then $\frac{1}{\tau}(x, s, y)$ solves the optimality conditions of (4.1), i.e., $\frac{1}{\tau}(x, s, y)$ is a complementary solution. If $\kappa > 0$, then $x$ and/or $(s, y)$ are improving rays for the primal and/or dual. If, on the other hand, $\tau = \kappa = 0$, then the solution $(x, s, y, \tau, \kappa)$ reveals nothing about the primal or dual optimal value—either could be finite or unbounded from above or below.[1] Further, $\tau = \kappa = 0$ holds for *all* solutions if and only if complementary solutions and improving rays do not exist, i.e., if and only if the cone program of interest is pathological.

**Theoretical contributions** In this chapter, we reexamine the solution set of the embedding in the pathological case. To summarize our results, we first note that the solution set of the embedding (4.2) is a convex cone. For pathological instances, we show points in the relative interior of this cone identify faces of $\mathcal{K}$ and/or $\mathcal{K}^*$ containing the primal and/or dual feasible set, which is precisely the information needed for facial reduction. This allows one to reformulate the embedding (4.2) over a face and resolve, repeating until a complementary solution or an improving ray is obtained. This ultimately leads to an algorithm that finds the optimal value and a point attaining it (if one exists) for any cone programming instance.

This algorithm, of course, needs to find relative interior solutions of the embedding (4.2). As we show, such solutions are obtained from relative interior solutions of the *extended-embedding* [145], a strictly feasible cone program with a strictly feasible dual. Note that if $\mathcal{K}$ is a symmetric cone, the central path of the extended embedding converges to the relative interior of its solution set [66, 116]. Further, reformulating (4.2) over a face of $\mathcal{K}$ or $\mathcal{K}^*$ requires only basic linear algebra (see, e.g., Chapter 1.3.1). Hence, implementations of our algorithms are conceptually simple for symmetric cones, involving only linear algebra and repeated calls to an interior point method.

**Numerical experiments** Though the aforementioned implementations are *conceptually* simple, they nevertheless face significant *practical* barriers. Numerical experiments will be given to illustrate the following three issues. First, interior-point methods only approximate the limit point of the central path; hence, in practice, only approximate solutions of the embedding are obtained. Second, facial reduction with these approximate solutions can change the optimal value by an arbitrary amount, e.g., by changing a feasible problem to an infeasible problem. Note that this second issue is rooted in the fact at least one problem—primal or dual—is *ill-posed* in the sense of Renegar [118] when $\tau = \kappa = 0$ for all solutions;[2] that is, at least one problem has optimal value infinitely sensitive to perturbations of $A$, $b$ and $c$. Third, strict complementarity can fail badly, which can cause poor convergence of interior-point algorithms. Indeed, the extended-embedding never has a strictly complementary solution when $\tau = \kappa = 0$

[1] Though the solution reveals nothing, the asymptotic behavior of central-path-following techniques gives information in certain cases. See Luo et al. [87] and de Klerk et al. [41].

[2] In this case, no complementary solution or improving ray exists. At least one problem—primal or dual—is thus feasible and fails Slater's condition or infeasible and fails to have a dual improving ray. In other words, at least one problem is *weakly feasible* or *weakly infeasible* in the sense of [88] and hence ill-posed in the sense of Renegar [118].

| Case | Interpretation |
|---|---|
| $\tau > 0, \kappa = 0$ | Complementary solution |
| $\tau = 0, \kappa > 0$ | Improving ray(s) |
| $\tau = 0, \kappa = 0$ | Hyperplane(s) for facial reduction |

**Table 4.1:** Interpretation of a relative interior solution to the self-dual embedding (4.2) in terms of the conic optimization problem (4.1). A main observation of this paper, given by Corollary 4.2.1, is summarized by the last row.

holds for all solutions, since, by definition, strict complementarity requires $\tau\kappa = 0$ and $\tau + \kappa > 0$. We emphasize that these issues currently preclude a practical implementation of the mentioned algorithms. Nevertheless, experiments indicate "good" solutions to the embedding can be obtained for many pathological instances, which may have practical value; indeed, in Section 4.4.4, we use approximate solutions to explain the poor performance [93] of primal-dual solvers on a DIMACS library instance [104] by offering numerical evidence the dual optimal value is unattained.

**Outline** This chapter is organized as follows. Section 4.1 establishes additional notation and definitions. Section 4.2 studies the solution set of the embedding and the extended-embedding. Section 4.3 gives a theoretical algorithm for solving arbitrary instances of a given cone program. Section 4.4 gives numerical experiments, highlighting barriers to practical implementation.

## 4.1 Notation and definitions

To enable succinct and precise statements, we need additional definitions. First, we let $\mathbf{P}(\mathcal{C})$ and $\mathbf{D}(\mathcal{C})$ denote the following primal-dual pair parametrized by a closed, convex cone $\mathcal{C} \subseteq \mathcal{V}$:

$$\mathbf{P}(\mathcal{C}) : \begin{array}{ll} \text{minimize} & \langle c, x \rangle \\ \text{subject to} & Ax = b, \\ & x \in \mathcal{C}, \end{array} \qquad \mathbf{D}(\mathcal{C}) : \begin{array}{ll} \text{maximize} & \langle b, y \rangle \\ \text{subject to} & s = c - A^* y, \\ & (s, y) \in \mathcal{C}^* \times \mathbb{R}^m. \end{array} \tag{4.3}$$

Here, $(A, b, c)$ is the problem data of (4.1); indeed, with this notation, $\mathbf{P}(\mathcal{K})$ and $\mathbf{D}(\mathcal{K})$ denote (4.1). We also parametrized the solution set of the self-dual embedding by $\mathcal{C}$.

**Definition 4.1.1.** *For a nonempty, closed, convex cone $\mathcal{C} \subseteq \mathcal{V}$, and the problem data $A : \mathcal{V} \to \mathbb{R}^m$, $b \in \mathbb{R}^m$ and $c \in \mathcal{V}$ of* (4.1), *define $\mathbf{H}(\mathcal{C})$ as the convex cone of solutions*

$(x, s, y, \tau, \kappa)$ *to the system*

$$\begin{aligned} Ax - b\tau &= 0, \\ -A^*y - s + c\tau &= 0, \\ \langle b, y\rangle - \langle c, x\rangle - \kappa &= 0, \\ (x, s, y, \tau, \kappa) \in \mathcal{C} \times \mathcal{C}^* &\times \mathbb{R}^m \times \mathbb{R}_+ \times \mathbb{R}_+. \end{aligned} \tag{4.4}$$

Note that if $\mathcal{C} = \mathcal{K}$, then $\mathbf{H}(\mathcal{C})$ equals the solution set of the self-dual embedding (4.2).

Finally, we name the hyperplanes of facial reduction *facial reduction certificates.*

**Definition 4.1.2.** *For a nonempty, closed, convex cone $\mathcal{C} \subseteq \mathcal{V}$, and the problem data $A : \mathcal{V} \to \mathbb{R}^m$, $b \in \mathbb{R}^m$ and $c \in \mathcal{V}$ of* (4.1)*, define facial reduction certificates as follows:*

- *Call $s \in \mathcal{C}^*$ a* facial reduction certificate for $\mathbf{P}(\mathcal{C})$ *if the hyperplane $s^\perp$ contains the affine set $\{x \in \mathcal{V} : Ax = b\}$ and $\mathcal{C} \cap s^\perp \subseteq \mathcal{C}$ holds strictly.*
- *Call $x \in \mathcal{C}$ a* facial reduction certificate for $\mathbf{D}(\mathcal{C})$ *if the hyperplane $x^\perp$ contains the affine set $\{c - A^*y : y \in \mathbb{R}^m\}$ and $\mathcal{C}^* \cap x^\perp \subseteq \mathcal{C}^*$ holds strictly.*

We also define a notion of optimality for certificates.

**Definition 4.1.3.** *Let $\mathcal{C} \subseteq \mathcal{V}$ be a nonempty, closed, convex cone. Let $Z_p \subseteq \mathcal{V}$ denote the set of facial reduction certificates for $\mathbf{P}(\mathcal{C})$ and $Z_d \subseteq \mathcal{V}$ the set of facial reduction certificates for $\mathbf{D}(\mathcal{C})$.*

- *$s \in Z_p$ is an* optimal facial reduction certificate *for $\mathbf{P}(\mathcal{C})$ if $\mathcal{C} \cap s^\perp$ satisfies*
$$\mathcal{C} \cap s^\perp \subseteq \mathcal{C} \cap \hat{s}^\perp \quad \textit{for all } \hat{s} \in Z_p.$$
- *$x \in Z_d$ is an* optimal facial reduction certificate *for $\mathbf{D}(\mathcal{C})$ if $\mathcal{C}^* \cap x^\perp$ satisfies*
$$\mathcal{C}^* \cap x^\perp \subseteq \mathcal{C}^* \cap \hat{x}^\perp \quad \textit{for all } \hat{x} \in Z_d.$$

Note that the sum of two facial reduction certificates is a certificate. Further, the sum of a maximal set of linearly independent certificates is optimal given the identities

$$\mathcal{C} \cap (s_1 + s_2)^\perp = \mathcal{C} \cap s_1^\perp \cap s_2^\perp, \qquad \mathcal{C}^* \cap (x_1 + x_2)^\perp = \mathcal{C}^* \cap x_1^\perp \cap x_2^\perp,$$

which hold for any $s_1, s_2 \in \mathcal{C}^*$ and $x_1, x_2 \in \mathcal{C}$.

## ■ 4.2 Solutions to self-dual embeddings

In this section, we study the solution set of the self-dual embedding $\mathbf{H}(\mathcal{C})$. Not every point in $\mathbf{H}(\mathcal{C})$ provides information about the primal-dual pair $\mathbf{P}(\mathcal{C})$ and $\mathbf{D}(\mathcal{C})$; indeed,

$\mathbf{H}(\mathcal{C})$ contains the zero vector. We therefore focus on two subsets of $\mathbf{H}(\mathcal{C})$ whose points all provide information. We classify the points in these subsets based on $\tau$ and $\kappa$, leveraging the fact that $\tau\kappa = 0$; that is, for $(x, s, y, \tau, \kappa) \in \mathbf{H}(\mathcal{C})$,

$$0 \le \langle x, s\rangle = \langle x, c\tau - A^*y\rangle = \tau(\langle c, x\rangle - \langle b, y\rangle) = -\tau\kappa \le 0. \tag{4.5}$$

In other words, for each subset, we consider the three cases $\tau > 0$, $\kappa > 0$ and $\tau = \kappa = 0$, providing novel insight into this latter case.

The first subset of interest is the relative interior of $\mathbf{H}(\mathcal{C})$. Our main result (Theorem 4.2.1) implies the results of Table 4.1; specifically, it shows that points in the relative interior yield complementary solutions when $\tau > 0$, improving rays when $\kappa > 0$, and optimal facial reduction certificates when $\tau = \kappa = 0$. Specialized to SDP, this expands the analysis of [41, 40].

We next consider the subset $M \cap \mathbf{H}(\mathcal{C})$ where $M$ is a distinguished type of hyperplane introduced in Ye et al. [145] and studied in Freund [56]. As with the relative interior, points in this subset yield complementary solutions when $\tau > 0$ and improving rays when $\kappa > 0$. When $\tau = \kappa = 0$, we obtain either an improving ray or a facial reduction certificate; in other words, we obtain a certificate that Slater's condition has failed.

Finally, we show how to find points in both subsets by solving the extended-embedding of Ye et al. [145]—a strictly feasible conic problem with strictly feasible dual.

## 4.2.1 The relative interior

The following theorem classifies $(x, s, y, \tau, \kappa) \in \operatorname{relint} \mathbf{H}(\mathcal{C})$ by the values of $\tau$ and $\kappa$. A corollary follows restating key statements in terms of complementary solutions, improving rays, and facial reduction certificates for the primal-dual pair given by $\mathbf{P}(\mathcal{C})$ and $\mathbf{D}(\mathcal{C})$. The key observation for the case of $\tau = \kappa = 0$ is the following: if $(x, s, y, 0, 0) \in \mathbf{H}(\mathcal{C})$, then, by inspection, so is at least one of the points $(0, s, y, 0, \langle b, y\rangle)$ or $(x, 0, 0, 0, -\langle c, x\rangle)$. This will imply $\langle c, x\rangle = \langle b, y\rangle = 0$ when $\tau = \kappa = 0$ holds for points in the relative interior.

**Theorem 4.2.1.** *Let $\mathcal{C} \subseteq \mathcal{V}$ be a nonempty, closed, convex cone with $(x, s, y, \tau, \kappa) \in \operatorname{relint} \mathbf{H}(\mathcal{C})$. Then, the complementarity condition $\tau\kappa = 0$ holds. The following statements also hold.*

1. *If $\tau > 0$, then $\frac{1}{\tau}(Ax) = b$, $\frac{1}{\tau}(A^*y + s) = c$, and $\langle b, y\rangle = \langle c, x\rangle$.*
2. *If $\kappa > 0$, then $Ax = 0$, $A^*y + s = 0$, and $\langle b, y\rangle > \langle c, x\rangle$.*
3. *If $\tau = \kappa = 0$, then $\hat{\tau} = \hat{\kappa} = 0$ and $\langle c, \hat{x}\rangle = \langle b, \hat{y}\rangle = 0$ for all $(\hat{x}, \hat{s}, \hat{y}, \hat{\tau}, \hat{\kappa}) \in \mathbf{H}(\mathcal{C})$. Further, letting $\mathcal{F}_p := \mathcal{C} \cap s^{\perp}$, $\mathcal{F}_d := \mathcal{C}^* \cap x^{\perp}$, $\mathcal{A}_p := \{x \in \mathcal{V} : Ax = b\}$ and $\mathcal{A}_d := \{c - A^*y : y \in \mathbb{R}^m\}$,*

(a) *the hyperplane* $s^{\perp}$ *contains* $\mathcal{A}_p$*;*

(b) *the hyperplane* $x^{\perp}$ *contains* $\mathcal{A}_d$*;*

(c) *the face* $\mathcal{F}_p$ *is proper if and only if* $\mathcal{A}_p \cap \operatorname{relint} \mathcal{C}$ *is empty;*

(d) *the face* $\mathcal{F}_d$ *is proper if and only if* $\mathcal{A}_d \cap \operatorname{relint} \mathcal{C}^*$ *is empty;*

(e) *at least one of the faces* $\mathcal{F}_p$ *or* $\mathcal{F}_d$ *is proper;*

(f) *the inclusion* $\mathcal{F}_p \subseteq \mathcal{C} \cap \hat{s}^{\perp}$ *holds for all* $\hat{s} \in \mathcal{C}^*$ *satisfying* $\mathcal{A}_p \subseteq \hat{s}^{\perp}$*;*

(g) *the inclusion* $\mathcal{F}_d \subseteq \mathcal{C}^* \cap \hat{x}^{\perp}$ *holds for all* $\hat{x} \in \mathcal{C}$ *satisfying* $\mathcal{A}_d \subseteq \hat{x}^{\perp}$*.*

*Proof.* We only prove the third statement, noting that the first two are well known and follow easily from the complementarity condition $\tau\kappa = 0$ implied by (4.5). To begin, let $w := (x, s, y, \tau, \kappa) \in \operatorname{relint} \mathbf{H}(\mathcal{C})$ and assume $\tau = \kappa = 0$. Now pick arbitrary $\hat{w} := (\hat{x}, \hat{s}, \hat{y}, \hat{\tau}, \hat{\kappa}) \in \mathbf{H}(\mathcal{C})$. First note that $\hat{\tau} = \hat{\kappa} = 0$; otherwise, we'd have $w - \alpha\hat{w} \notin \mathbf{H}(\mathcal{C})$ for every $\alpha > 0$, contradicting the fact $w \in \operatorname{relint} \mathbf{H}(\mathcal{C})$. Since $\hat{\kappa} = 0$, it follows that $\langle c, \hat{x} \rangle - \langle b, \hat{y} \rangle = 0$, i.e., that $\langle c, \hat{x} \rangle = \langle b, \hat{y} \rangle$. We will now tighten this to $\langle c, \hat{x} \rangle = \langle b, \hat{y} \rangle = 0$. To begin, let $\hat{\theta} := \langle c, \hat{x} \rangle = \langle b, \hat{y} \rangle$. By inspection, at least one point—$(0, \hat{s}, \hat{y}, 0, \hat{\theta})$ or $(\hat{x}, 0, 0, 0, -\hat{\theta})$—is in $\mathbf{H}(\mathcal{C})$. But as just argued, the $\kappa$-coordinate of any point in $\mathbf{H}(\mathcal{C})$ must be zero. We conclude $\hat{\theta} = 0$; hence $\langle c, \hat{x} \rangle = \langle b, \hat{y} \rangle = 0$, as desired. We now use this to show statements (3a)-(3b).

To see (3a)-(3b), first note that $\langle \hat{x}, s \rangle = -\langle \hat{x}, A^* y \rangle = -\langle b, y \rangle = 0$ for all solutions $\hat{x}$ of $Ax = b$. Hence, the solution set $\{x \in \mathcal{V} : Ax = b\}$ is contained in the hyperplane $s^{\perp}$. Likewise, $\langle x, \hat{s} \rangle = \langle c, x \rangle = 0$ for all $\hat{s} \in \{c - A^* y : y \in \mathbb{R}^m\}$; hence, $x^{\perp}$ contains $\{c - A^* y : y \in \mathbb{R}^m\}$.

We now show (3d). One direction is trivial; if $\mathcal{A}_d$ is contained in a proper face of $\mathcal{C}^*$, then $\mathcal{A}_d \cap \operatorname{relint} \mathcal{C}^*$ must be empty. For the converse direction, suppose that $\mathcal{A}_d \cap \operatorname{relint} \mathcal{C}^*$ is empty. The separating hyperplane theorem [121, Theorem 11.3] states that a hyperplane exists *properly* separating these sets. Using [121, Theorem 11.7], we can additionally assume this hyperplane passes through the origin since $\mathcal{C}^*$ is a cone. In other words, there exists $\hat{x} \in \mathcal{C}^{**} = \mathcal{C}$ satisfying

$$\langle \hat{x}, c - A^* y \rangle \leq 0 \text{ for all } y \in \mathbb{R}^m,$$
$$\langle \hat{x}, z \rangle \neq 0 \text{ for some } z \in \{c - A^* y : y \in \mathbb{R}^m\} \cup \mathcal{C}^*.$$

It follows that $\langle c, \hat{x} \rangle \leq \langle A^* y, \hat{x} \rangle$ for arbitrary $y \in \mathbb{R}^m$, which implies $A\hat{x} = 0$ and $\langle c, \hat{x} \rangle \leq 0$. But $\langle c, \hat{x} \rangle = 0$, otherwise $\hat{x}$ is an improving ray for $\mathbf{P}(\mathcal{C})$, and $(\hat{x}, 0, 0, 0, -\langle c, \hat{x} \rangle) \in \mathbf{H}(\mathcal{C})$ with $\kappa > 0$. Hence, the hyperplane $\hat{x}^{\perp}$ contains $\{c - A^* y : y \in \mathbb{R}^m\}$ implying $\langle \hat{x}, z \rangle \neq 0$ for some $z \in \mathcal{C}^*$ given proper separation of the sets. That is, $\hat{x}$ exposes a proper face of $\mathcal{C}^*$. We now use $\hat{x}$ to show that $x$ exposes a proper face of $\mathcal{C}^*$ as claimed. Clearly, $\hat{w} := (\hat{x}, 0, 0, 0, 0) \in \mathbf{H}(\mathcal{C})$. Since $w$ is in the relative interior of $\mathbf{H}(\mathcal{C})$, it holds that $w \pm \alpha\hat{w} \in \mathbf{H}(\mathcal{C})$ and $x \pm \alpha\hat{x} \in \mathcal{C}$ for some $\alpha > 0$. Hence, for any $u \in \mathcal{C}^*$, the inequality $\langle u, x \pm \alpha\hat{x}\rangle \geq 0$ holds, which in turn implies $\langle u, \hat{x}\rangle = 0$ when $\langle u, x\rangle = 0$. In other words, $\mathcal{C}^* \cap x^\perp$ is contained in a proper face, i.e.,

$$\mathcal{C}^* \cap x^\perp \subseteq \mathcal{C}^* \cap \hat{x}^\perp, \tag{4.6}$$

and is hence proper.

Applying the argument of the previous paragraph to the set $\mathcal{A}_p \cap \operatorname{relint} \mathcal{C}$ shows (3c).

Statement (3e) follows if at least one of the sets—$\{x \in \mathcal{V} : Ax = b\} \cap \operatorname{relint} \mathcal{C}$ or $\{c - A^*y : y \in \mathbb{R}^m\} \cap \operatorname{relint} \mathcal{C}^*$—is empty. Suppose this is not the case. Then, Slater's condition is satisfied for both $\mathbf{P}(\mathcal{C})$ and $\mathbf{D}(\mathcal{C})$ showing the existence of an optimal primal-dual solution with zero duality gap [13, Section 7.2.2]. Hence, there exists a point in $\mathbf{H}(\mathcal{C})$ with $\tau > 0$, contradicting the assumption that $(x, s, y, \tau, \kappa)$ is in the relative interior of $\mathbf{H}(\mathcal{C})$.

The same argument that shows the containment (4.6) shows (3f) and (3g). □

The following corollary arises simply from definitions. Variants of the first two statements (and their converses) are well known for semidefinite optimization [41, Theorem 5.3.2]. Note that taking $\mathcal{C} = \mathcal{K}$ yields Table 4.1.

**Corollary 4.2.1.** *Let $\mathcal{C} \subseteq \mathcal{V}$ be a nonempty, closed, convex cone with $(x, s, y, \tau, \kappa) \in \operatorname{relint} \mathbf{H}(\mathcal{C})$. The following statements hold for the primal-dual pair $\mathbf{P}(\mathcal{C})$ and $\mathbf{D}(\mathcal{C})$.*

1. *If $\tau > 0$, then $\frac{1}{\tau}(x, s, y)$ is a complementary solution for $\mathbf{P}(\mathcal{C})$ and $\mathbf{D}(\mathcal{C})$.*
2. *If $\kappa > 0$, then $x$ is an improving ray for $\mathbf{P}(\mathcal{C})$ and/or $(s, y)$ is an improving ray for $\mathbf{D}(\mathcal{C})$.*
3. *If $\tau = \kappa = 0$, then $x$ and/or $s$ are facial reduction certificates.*

*Moreover, converses of the first two statements hold: if $\mathbf{P}(\mathcal{C})$ and $\mathbf{D}(\mathcal{C})$ have a complementary solution, then $\tau > 0$; if $\mathbf{P}(\mathcal{C})$ and/or $\mathbf{D}(\mathcal{C})$ have an improving ray, then $\kappa > 0$.*

We now strengthen the third statement of Corollary 4.2.1, showing that facial reduction certificates are obtained for both problems if they both fail Slater's condition. We then illustrate the second statement does not have an analogous strengthening; that is, improving rays are not necessarily obtained for both problems even if they exist.

**Facial reduction certificates for the primal and dual.** The third statement of Corollary 4.2.1 does not guarantee that $s$ and $x$ are *both* facial reduction certificates when such certificates exist for *both* primal and dual problems. Statements (3c)-(3d) of Theorem 4.2.1 provide this guarantee. Moreover, statements (3f)-(3g) imply that these certificates are optimal. Formally:

**Corollary 4.2.2.** *Let $\mathcal{C} \subseteq \mathcal{V}$ be a nonempty, closed, convex cone with $(x, s, y, \tau, \kappa) \in \operatorname{relint} \mathbf{H}(\mathcal{C})$. If $\tau = \kappa = 0$, the following statements hold.*

- *The set $\{x \in \mathcal{V} : Ax = b\} \cap \operatorname{relint} \mathcal{C}$ is empty if and only if $s$ is a facial reduction certificate for $\mathbf{P}(\mathcal{C})$.*
- *The set $\{c - A^*y : y \in \mathbb{R}^m\} \cap \operatorname{relint} \mathcal{C}^*$ is empty if and only if $x$ is a facial reduction certificate for $\mathbf{D}(\mathcal{C})$.*

*Moreover, if $x$ (resp., $s$) is a facial reduction certificate, then $x$ (resp. $s$) is optimal in the sense of Definition 4.1.3.*

This can be compared to Corollary 1.3.2 (the basis of the facial reduction algorithm [102]), which asserted that, for feasible problems, existence of facial reduction certificates is equivalent to failure of Slater's condition. Also note that the optimality of certificates is ensured by the restriction to $\operatorname{relint} \mathbf{H}(\mathcal{C})$; note that [130, Procedure 1] and [138, Algorithm 4.1] find optimal certificates using similar restrictions.

The next example illustrates the first two statements of Corollary 4.2.2. Here, the primal-dual pair has finite but non-zero duality gap. This implies $\tau = \kappa = 0$ for all points in $\operatorname{relint} \mathbf{H}(\mathcal{C})$ and that Slater's condition fails for both the primal and dual.

**Example 4.2.1** (Example with positive duality gap [4].)**.** *Let $\mathcal{Q}^n := \left\{ x_1 \geq \sqrt{\sum_{i=2}^n x_i^2} \right\}$ denote the Lorentz cone. The following primal-dual pair has a duality gap of one:*

$$\begin{array}{ll} \text{minimize} & x_3 \\ \text{subject to} & x_1 + x_2 + x_4 + x_5 = 0 \\ & -x_3 + x_4 = 1 \\ & x \in \mathcal{Q}^3 \times \mathcal{Q}^2 \end{array} \qquad \begin{array}{ll} \text{maximize} & y_2 \\ \text{subject to} & \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} y_1 \\ y_1 \\ -y_2 \\ y_1 + y_2 \\ y_1 \end{pmatrix} = s \\ & s \in \mathcal{Q}^3 \times \mathcal{Q}^2. \end{array}$$

*Indeed, if $x \in \mathcal{Q}^3 \times \mathcal{Q}^2$, then $x_1 + x_2 \geq 0$ and $x_4 + x_5 \geq 0$; if, in addition, $x_1 + x_2 + x_4 + x_5 = 0$, then $x_1 + x_2 = 0$, implying $x_3 = 0$ if $(x_1, x_2, x_3) \in \mathcal{Q}^3$. On the other hand, dual feasible points satisfy $s_1 = s_2$, which in turn implies $s_3 = 0$, i.e., $y_2 = -1$. Since both problems are feasible, the duality gap is $0 - (-1) = 1$.*

*Since there is a duality gap of one, any point in $\operatorname{relint} \mathbf{H}(\mathcal{Q}^3 \times \mathcal{Q}^2)$ satisfies $\tau = \kappa = 0$, e.g.,*

$$\hat{x} = (1, -1, 0, 0, 0)^T, \quad \hat{s} = (1, 1, 0, 1, 1)^T, \quad \hat{y} = (-1, 0)^T, \quad \hat{\tau} = \hat{\kappa} = 0.$$

*We see that $\hat{s}$ and $\hat{x}$ are facial reduction certificates for the primal and dual, as predicted by Corollary 4.2.2.*

**Improving rays for the primal and dual.** Statement 2 of Corollary 4.2.1 does not guarantee $s$ and $x$ are *both* improving rays when such rays exists for *both* primal and dual problems. Unfortunately, this statement cannot be strengthened—in the $\kappa > 0$ case, there are instances for which relative interior points do not yield improving rays for both problems, even if these rays exist. This is a known shortcoming of the self-dual embedding that occurs even in the linear programming case (see, e.g., [145]). The following example illustrates this shortcoming.

**Example 4.2.2.** *Consider the following primal-dual pair of linear programs*

$$\begin{array}{ll} \text{minimize} & -x_1 \\ \text{subject to} & x_1 - x_2 = 0, \\ & -(x_1 - x_2) = 1, \\ & x \in \mathbb{R}^2_+, \end{array} \qquad \begin{array}{ll} \text{maximize} & y_2 \\ \text{subject to} & \binom{-1}{0} - \begin{pmatrix} y_1 - y_2 \\ -(y_1 - y_2) \end{pmatrix} = s, \\ & s \in \mathbb{R}^2_+, \end{array}$$

*where both the primal and dual problem are infeasible. Indeed, the point* $(\hat{x}, \hat{s}, \hat{y}, \hat{\tau}, \hat{\kappa}) \in \operatorname{relint} \mathbf{H}(\mathbb{R}^2_+)$ *yields an improving ray* $\hat{x}$ *for the primal and an improving ray* $\hat{y}$ *for the dual, where*

$$\hat{x} = (1,1)^T, \quad \hat{s} = (0,0)^T, \quad \hat{y} = (1,1)^T, \quad \hat{\tau} = 0, \quad \hat{\kappa} = 2.$$

*Nevertheless, the entire family of points*

$$\tilde{x} = (r,r)^T, \quad \tilde{s} = (0,0)^T, \quad \tilde{y} = (t,t)^T, \quad \tilde{\tau} = 0, \quad \tilde{\kappa} = r + t, \quad \textit{for } r > -t \geq 0$$

*are also in the relative interior of solutions to the self-dual embedding, and only give improving rays for the primal problem.*

Note that this example illustrates one cannot in general decide infeasibility of the primal and dual from a single point in $\operatorname{relint} \mathbf{H}(\mathcal{C})$ with $\kappa > 0$. We will revisit this issue in Section 4.3.3, as it complicates a presented solution algorithm for finding optimal values.

## 4.2.2 The intersection with distinguished hyperplanes

The next subset of $\mathbf{H}(\mathcal{C})$ of interest is its intersection with a distinguished type of hyperplane. Such a hyperplane $M$ is defined by $\mu > 0$ and a fixed point $(\hat{x}, \hat{s}, \hat{\tau}, \hat{\kappa}) \in \operatorname{relint}(\mathcal{C} \times \mathcal{C}^* \times \mathbb{R}_+ \times \mathbb{R}_+)$ via

$$M := \{(x, s, y, \tau, \kappa) : \langle \hat{s}, x \rangle + \langle \hat{x}, s \rangle + \hat{\tau}\kappa + \hat{\kappa}\tau = \mu\}.$$

For a particular $\mu$, membership in $M$ is an implicit constraint of the extended-embedding of [145]. Also, as shown in [56], it can be interpreted as a norm constraint on $(x, s, \tau, \kappa) \in \mathcal{C} \times \mathcal{C}^* \times \mathbb{R}_+ \times \mathbb{R}_+$ when $\mathcal{C}$ is a *proper cone*, i.e., a cone that is full-dimensional, pointed, convex, and closed.

As we now show, from $(x, s, y, \tau, \kappa) \in \mathbf{H}(\mathcal{C}) \cap M$, one always obtains one of the following objects: a complementary solution or a certificate that the primal or dual has failed Slater's condition. As with relint $\mathbf{H}(\mathcal{C})$, complementary solutions are obtained when $\tau > 0$ and improving rays are obtained when $\kappa > 0$. The case of $\tau = \kappa = 0$, however, is now more delicate: we either obtain a facial reduction certificate *or* an improving ray. Facial reduction certificates are also not necessarily optimal nor are they necessarily obtained for *both* primal and dual if *both* fail Slater's condition.

**Theorem 4.2.2.** *Let $\mathcal{C} \subseteq \mathcal{V}$ be a nonempty closed, convex cone. For a scalar $\mu > 0$ and point $(\hat{x}, \hat{s}, \hat{\tau}, \hat{\kappa})$ in the relative interior of $\mathcal{C} \times \mathcal{C}^* \times \mathbb{R}_+ \times \mathbb{R}_+$ consider the hyperplane*

$$M := \{(x, s, y, \tau, \kappa) : \langle \hat{s}, x\rangle + \langle \hat{x}, s\rangle + \hat{\tau}\kappa + \hat{\kappa}\tau = \mu\}.$$

*For $(x, s, y, \tau, \kappa) \in M \cap \mathbf{H}(\mathcal{C})$, the complementarity condition $\tau\kappa = 0$ holds. The following statements also hold.*

1. *If $\tau > 0$, then $\frac{1}{\tau}(x, s, y)$ is a complementary solution of $\mathbf{P}(\mathcal{C})$ and $\mathbf{D}(\mathcal{C})$.*
2. *If $\kappa > 0$, then $x$ is an improving ray for $\mathbf{P}(\mathcal{C})$ and/or $(s, y)$ is an improving ray for $\mathbf{D}(\mathcal{C})$.*
3. *If $\tau = \kappa = 0$, at least one of the following statements is true:*
   - *(a) $x$ is an improving ray for $\mathbf{P}(\mathcal{C})$;*
   - *(b) $(s, y)$ is an improving ray for $\mathbf{D}(\mathcal{C})$;*
   - *(c) $x$ is a facial reduction certificate for $\mathbf{D}(\mathcal{C})$, i.e., $\mathcal{C}^* \cap x^\perp \subseteq \mathcal{C}^*$ holds strictly and the hyperplane $x^\perp$ contains $\mathcal{A}_d := \{c - A^*y : y \in \mathbb{R}^m\}$;*
   - *(d) $s$ is a facial reduction certificate for $\mathbf{P}(\mathcal{C})$, i.e., $\mathcal{C} \cap s^\perp \subseteq \mathcal{C}$ holds strictly and the hyperplane $s^\perp$ contains $\mathcal{A}_p := \{x \in \mathcal{V} : Ax = b\}$.*

*Proof.* The first two statements are trivial from the complementarity condition $\tau\kappa = 0$. For the third statement, we note from $\kappa = 0$ that $\langle b, y\rangle = \langle c, x\rangle$ and consider the three cases $\langle b, y\rangle = \langle c, x\rangle > 0$, $\langle b, y\rangle = \langle c, x\rangle < 0$ and $\langle c, x\rangle = \langle b, y\rangle = 0$. In the first case, it is trivial to check that $(s, y)$ is an improving ray and in the second that $x$ is an improving ray. In the third case, where $\langle c, x\rangle = \langle b, y\rangle = 0$, it follows $x^\perp$ and $s^\perp$ contain $\mathcal{A}_d$ and $\mathcal{A}_p$, respectively; see the proof for statements (3a)-(3b) of Theorem 4.2.1. In addition, at least one exposes a proper face since $\tau = \kappa = 0$ implies $\langle \hat{s}, x\rangle + \langle \hat{x}, s\rangle = \mu > 0$; that is, the face exposed by $x^\perp$ doesn't contain $\hat{s}$ if $\langle \hat{s}, x\rangle > 0$ and the face exposed by $s^\perp$ does not contain $\hat{x}$ if $\langle \hat{x}, s\rangle > 0$. □

We conclude with descriptions of relint$(M \cap \mathbf{H}(\mathcal{C}))$. We will use this description in the next section to find points in relint $\mathbf{H}(\mathcal{C})$.

**Theorem 4.2.3.** *Let $\mathcal{C} \subseteq \mathcal{V}$ be a nonempty closed, convex cone. For a scalar $\mu > 0$ and point $(\hat{x}, \hat{s}, \hat{\tau}, \hat{\kappa})$ in the relative interior of $\mathcal{C} \times \mathcal{C}^* \times \mathbb{R}_+ \times \mathbb{R}_+$ consider the hyperplane*

$$M := \{(x, s, y, \tau, \kappa) : \langle \hat{s}, x\rangle + \langle \hat{x}, s\rangle + \hat{\tau}\kappa + \hat{\kappa}\tau = \mu\}.$$

*The following statements hold.*

1. $\operatorname{relint}\left(M \cap \mathbf{H}(\mathcal{C})\right) = M \cap \operatorname{relint} \mathbf{H}(\mathcal{C})$
2. *The conic hull of* $(M \cap \operatorname{relint} \mathbf{H}(\mathcal{C}))$ *equals* $\operatorname{relint} \mathbf{H}(\mathcal{C})$.

*Proof.* We first show the second statement. The inclusion $\subseteq$ is trivial: since $M \cap \operatorname{relint}\left(\mathbf{H}(\mathcal{C})\right)$ is a subset of $\operatorname{relint} \mathbf{H}(\mathcal{C})$, its conic hull is contained in $\operatorname{relint} \mathbf{H}(\mathcal{C})$ since $\operatorname{relint} \mathbf{H}(\mathcal{C})$ is a cone. For the reverse, let $(x, s, y, \tau, \kappa) \in \operatorname{relint} \mathbf{H}(\mathcal{C})$. By Corollary 4.2.1, $\tau > 0$, $\kappa > 0$, $s \notin \mathcal{C}^\perp$, or $x \notin (\mathcal{C}^*)^\perp$. Hence, the sum $\langle \hat{x}, s\rangle + \langle x, \hat{s}\rangle + \hat{\tau}\kappa + \tau\hat{\kappa}$ equals some positive number, say, $\alpha$. It follows that $\lambda(x, s, y, \tau, \kappa) \in M \cap \operatorname{relint} \mathbf{H}(\mathcal{C})$ for the strictly positive number $\lambda = \frac{\mu}{\alpha}$ by the fact $\operatorname{relint} \mathbf{H}(\mathcal{C})$ is a cone.

The second statement implies that $M \cap \operatorname{relint} \mathbf{H}(\mathcal{C})$ is nonempty; hence, the first statement follows from Corollary 6.5.1 of [121].

□

**Remark 4.2.1.** *Note the special structure of $M$ is crucial to statements* (3c) *and* (3d) *of Theorem 4.2.2. If, for instance, we replaced $M$ with an arbitrary hyperplane, the inclusions $\mathcal{C} \cap s^\perp \subseteq \mathcal{C}$ and $\mathcal{C}^* \cap x^\perp \subseteq \mathcal{C}^*$ wouldn't necessarily hold strictly. It is also crucial to Theorem 4.2.3: an arbitrary hyperplane doesn't necessarily intersect the relative interior of* $\mathbf{H}(\mathcal{C})$.

### ■ 4.2.3 Finding solutions via extended-embeddings

We now discuss how to find points in the two analyzed subsets, $\operatorname{relint} \mathbf{H}(\mathcal{C})$ and $M \cap \mathbf{H}(\mathcal{C})$, considered in Section 4.2.1 and 4.2.2. Recall that $M$ is a hyperplane of the distinguished type defined by a fixed point $(\hat{x}, \hat{s}, \hat{\tau}, \hat{\kappa}) \in \operatorname{relint}(\mathcal{C} \times \mathcal{C}^* \times \mathbb{R}_+ \times \mathbb{R}_+)$ and fixed $\mu > 0$. Our analysis is based on the following hyperplane $M'$ for which $\mu = \langle \hat{x}, \hat{s}\rangle + \hat{\tau}\hat{\kappa}$:

$$M' := \{(x, s, y, \tau, \kappa) : \langle \hat{s}, x\rangle + \langle \hat{x}, s\rangle + \hat{\tau}\kappa + \hat{\kappa}\tau = \langle \hat{x}, \hat{s}\rangle + \hat{\tau}\hat{\kappa}\}. \tag{4.7}$$

As shown in Ye et al. [145], points in $M' \cap \mathbf{H}(\mathcal{C})$ are obtained from solutions to an *extended-embedding*. Hence, points in $M \cap \mathbf{H}(\mathcal{C})$ for arbitrary $\mu > 0$ are obtained by rescaling these solutions. We will review this result and then build on it, showing points in $\operatorname{relint} \mathbf{H}(\mathcal{C})$ are obtained from relative interior solutions to the extended-embedding (which is not obvious).

Picking $\hat{y} \in \mathbb{R}^m$ and letting $\alpha = \langle \hat{s}, \hat{x} \rangle + \hat{\tau}\hat{\kappa}$, the extended-embedding takes the following form

$$\begin{array}{ll} \text{minimize} & \alpha\theta \\ \text{subject to} & Ax - b\tau = r_p\theta, \\ & -A^*y - s + c\tau = r_d\theta, \\ & \langle b, y \rangle - \langle c, x \rangle - \kappa = r_g\theta, \\ & \langle r_p, y \rangle + \langle r_d, x \rangle + r_g\tau = -\alpha, \end{array} \tag{4.8}$$

where $(x, s, y, \tau, \kappa, \theta) \in \mathcal{C} \times \mathcal{C}^* \times \mathbb{R}^m \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}$ is the decision variable and $r_p$, $r_d$, and $r_g$ are additional parameters defined by $(\hat{x}, \hat{s}, \hat{y}, \hat{\tau}, \hat{\kappa})$ via

$$r_p = A\hat{x} - b\hat{\tau}, \qquad r_d = -A^*\hat{y} - \hat{s} + c\hat{\tau}, \qquad r_g = \langle b, \hat{y} \rangle - \langle c, \hat{x} \rangle - \hat{\kappa}.$$

As shown in [145], the point $(\hat{x}, \hat{s}, \hat{y}, \hat{\tau}, \hat{\kappa}, 1)$ is strictly feasible; further, the optimal value is zero and it is attained, implying $\theta = 0$ at optimality. Note this latter property is easy to see from duality. To be precise, the dual problem is to maximize $-\alpha\bar{\theta}$ over dual variables $(\bar{x}, \bar{s}, \bar{y}, \bar{\tau}, \bar{\kappa}, \bar{\theta})$ satisfying identical constraints. Hence, an optimal solution, which exists by strict feasibility of (4.8), can be viewed as a dual optimal solution. In conclusion, $\alpha\theta = -\alpha\theta$ for an optimal solution since there is no duality gap (again by strict feasibility). This shows that $\theta = 0$ at optimality since $\alpha > 0$.

Note that if $\theta = 0$, then the first three constraints of (4.8) reduce to the defining equations of $\mathbf{H}(\mathcal{C})$; i.e., a point $(x, s, y, \tau, \kappa, \theta)$ with $\theta = 0$ satisfies these constraints if and only if $(x, s, y, \tau, \kappa) \in \mathbf{H}(\mathcal{C})$. For points in $\mathbf{H}(\mathcal{C})$, the remaining constraint is equivalent to membership in $M'$:

$$\begin{aligned} -\alpha &= \langle r_p, y \rangle + \langle r_d, x \rangle + r_g\tau \\ &= \langle A\hat{x} - b\hat{\tau}, y \rangle + \langle -A^*\hat{y} - \hat{s} + c\hat{\tau}, x \rangle + (\langle b, \hat{y} \rangle - \langle c, \hat{x} \rangle - \hat{\kappa})\tau \\ &= \langle -\hat{x}, -A^*y + c\tau \rangle - \langle \hat{s}, x \rangle - \langle \hat{y}, Ax - b\tau \rangle - \hat{\tau}(\langle b, y \rangle - \langle c, x \rangle) - \hat{\kappa}\tau \\ &= -\langle \hat{x}, s \rangle - \langle \hat{s}, x \rangle - \hat{\tau}\kappa - \hat{\kappa}\tau. \end{aligned}$$

It follows a feasible point $(x, s, y, \tau, \kappa, \theta)$ is optimal if and only if $\theta = 0$ and $(x, s, y, \tau, \kappa) \in M' \cap \mathbf{H}(\mathcal{C})$—a result originally due to [145]. We restate this as the first statement of the following theorem. The second statement—which to our knowledge is new—describes the relative interior of the optimal solution set of (4.8) in terms of $\operatorname{relint} \mathbf{H}(\mathcal{C})$.

**Theorem 4.2.4.** *Let* $\Omega \times \{0\} \subseteq \mathcal{C} \times \mathcal{C}^* \times \mathbb{R}^m \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}$ *denote the set of optimal solutions of* (4.8) *and* $M'$ *the hyperplane* (4.7). *The following statements hold:*

1. $\Omega = M' \cap \mathbf{H}(\mathcal{C})$.

2. $\operatorname{relint} \Omega = M' \cap \operatorname{relint} \mathbf{H}(\mathcal{C})$; *further, the conic hull of* $\operatorname{relint} \Omega$ *equals* $\operatorname{relint} \mathbf{H}(\mathcal{C})$.

*Proof.* The first statement is from [145]. The second is from statement 1 and 2 of Theorem 4.2.3. □

In conclusion, we obtain points in $M' \cap \mathbf{H}(\mathcal{C})$ and $\operatorname{relint} \mathbf{H}(\mathcal{C})$ from points in $\Omega \times \{0\}$ and $\operatorname{relint} \Omega \times \{0\}$, respectively. We also note when $\mathcal{C}$ is a symmetric cone (and the range of $A$ equals $\mathbb{R}^m$), the central path of the extended-embedding exists and converges to a point in $\operatorname{relint} \Omega \times \{0\}$ by strict feasibility and results of [66, 116]. Hence, for the cases of semidefinite, linear, and second-order cone optimization, we can approximate points in $\operatorname{relint} \Omega \times \{0\}$—and therefore points in $\operatorname{relint} \mathbf{H}(\mathcal{C})$—using interior-point methods. We will investigate convergence behavior, focusing on cases where $\tau = \kappa = 0$ for all $(x, s, y, \tau, \kappa) \in \mathbf{H}(\mathcal{C})$ in Section 4.4. We next develop a solution algorithm assuming oracle access to $\mathbf{H}(\mathcal{C})$.

## 4.3 An algorithm based on self-dual embeddings and facial reduction

As we have shown, we always obtain complementary solutions, improving rays, or facial reduction certificates from two subsets of $\mathbf{H}(\mathcal{C})$: the relative interior of $\mathbf{H}(\mathcal{C})$ and, for a distinguished type of hyperplane $M$, the set $M \cap \mathbf{H}(\mathcal{C})$. We've also seen facial reduction certificates are obtained from $\operatorname{relint} \mathbf{H}(\mathcal{C})$ precisely when they are needed, i.e., when complementary solutions or improving rays do not exist (Corollary 4.2.1). These facts suggest an iterative procedure that finds a point in $\operatorname{relint} \mathbf{H}(\mathcal{C})$, *regularizes* the primal or dual if necessary (i.e., performs a facial reduction iteration), and repeats until an improving ray or complementary solution is obtained. In this section we develop such a procedure.

To design a useful procedure, one must make an upfront decision: should regularization leave the primal or the dual optimal value unchanged? (Recall from Section 1.3.6 we cannot guarantee that both are unchanged.) In this section, we choose the former, stating an algorithm that finds a cone $\mathcal{C}$ such that $\mathbf{P}(\mathcal{C})$ and $\mathbf{P}(\mathcal{K})$ have equal optimal values and a point $(x, s, y, \tau, \kappa) \in \operatorname{relint} \mathbf{H}(\mathcal{C})$ satisfying $\tau > 0$ or $\kappa > 0$. This procedure appears in Algorithm 4.1. We then interpret $(x, s, y, \tau, \kappa)$ in terms of $\mathbf{P}(\mathcal{K})$, showing that an optimal solution is obtained when $\tau > 0$ and that a certificate of infeasibility is obtained when $\kappa > 0$ and $\langle b, y \rangle > 0$. In the remaining case where, $\kappa > 0$ and $\langle b, y \rangle \leq 0$, the primal problem $\mathbf{P}(\mathcal{K})$ is either unbounded or infeasible. A trivial extension (given in Section 4.3.3) resolves this ambiguity.

### 4.3.1 Basic properties

As indicated, Algorithm 4.1 solves a sequence of self-dual embeddings, terminating when $\tau > 0$ or $\kappa > 0$. If $\tau = \kappa = 0$, primal regularization is performed if $s \notin \mathcal{C}^{\perp}$ and dual regularization is performed otherwise. Note that when $\tau = \kappa = 0$, the condition $s \notin \mathcal{C}^{\perp}$

**Algorithm 4.1:** Given $\mathcal{K}$, returns $(x, s, y, \tau, \kappa) \in \operatorname{relint} \mathbf{H}(\mathcal{C})$ satisfying $\tau > 0$ or $\kappa > 0$ where $\mathbf{P}(\mathcal{C})$ and $\mathbf{P}(\mathcal{K})$ have equal optimal values.

```
C ← K
repeat
    Find (x, s, y, τ, κ) in the relative interior of H(C)
    if τ > 0 or κ > 0 then
        return (x, s, y, τ, κ)
    else
        if s ∉ C^⊥ then
            Regularize primal: C ← C ∩ s^⊥
        else
            Regularize dual: C ← (C* ∩ x^⊥)*
        end
    end
until algorithm returns;
```

holds if and only if $\mathbf{P}(\mathcal{C})$ fails Slater's condition—a consequence of Theorem 4.2.1 (3c). Hence, if necessary, this algorithm performs primal regularization until $\mathbf{P}(\mathcal{C})$ satisfies Slater's condition, then switches to dual regularization, never switching back. Precise statements of this property and others follow.

**Theorem 4.3.1.** *Let $\mathcal{A}_p := \{x \in \mathcal{V} : Ax = b\}$ and let $\theta_p \in \mathbb{R} \cup \{\pm\infty\}$ denote the optimal value of $\mathbf{P}(\mathcal{K})$, i.e.,*

$$\theta_p := \inf \{\langle c, x\rangle : x \in \mathcal{A}_p \cap \mathcal{K}\}.$$

*Algorithm 4.1 has the following basic properties.*

1. *Algorithm 4.1 terminates in finitely-many iterations. Further, it terminates after one iteration, with $\mathcal{C} = \mathcal{K}$, if and only if a complementary solution, primal improving ray or dual improving exists for the primal-dual pair $\mathbf{P}(\mathcal{K})$ and $\mathbf{D}(\mathcal{K})$.*

2. *At each iteration, the optimal values of the primal problem $\mathbf{P}(\mathcal{K})$ and the regularized problem $\mathbf{P}(\mathcal{C})$ are equal, i.e.,*

$$\theta_p = \inf \{\langle c, x\rangle : x \in \mathcal{A}_p \cap \mathcal{C}\}.$$

3. *Suppose the dual regularization step $\mathcal{C} \leftarrow (\mathcal{C}^* \cap x^\perp)^*$ executes, and let $\mathcal{C}'$ and $\mathcal{C}''$ denote $\mathcal{C}$ just before and after execution. The following statements hold.*

   (a) *The primal problems $\mathbf{P}(\mathcal{C}')$ and $\mathbf{P}(\mathcal{C}'')$ satisfy Slater's condition, that is, $\mathcal{A}_p \cap \operatorname{relint} \mathcal{C}'$ and $\mathcal{A}_p \cap \operatorname{relint} \mathcal{C}''$ are both nonempty. Further, $\mathcal{A}_p \cap \operatorname{relint} \mathcal{C}' \subseteq$*

$\mathcal{A}_p \cap \operatorname{relint} \mathcal{C}''$.

(b) *The primal regularization step* $\mathcal{C} \leftarrow \mathcal{C} \cap s^\perp$ *is not executed at any following iteration.*

4. *The dual regularization step* $\mathcal{C} \leftarrow (\mathcal{C}^* \cap x^\perp)^*$ *executes if and only if one of the following statements hold.*

   (a) *The optimal value of* $\mathbf{P}(\mathcal{K})$ *is finite and unattained;*

   (b) *The optimal value of* $\mathbf{P}(\mathcal{K})$ *is unbounded below* ($\theta_p = -\infty$) *and the set of improving rays* $\{x \in \mathcal{K} : Ax = 0, \langle c, x\rangle < 0\}$ *is empty.*

5. *At termination,* $\mathcal{A}_p \cap \mathcal{K} \subseteq \mathcal{A}_p \cap \mathcal{C}$*, with strict inclusion only if* (4a) *or* (4b) *holds.*

*Proof.* In the arguments below, we call $\mathcal{C} \leftarrow \mathcal{C} \cap s^\perp$ the primal regularization step and $\mathcal{C} \leftarrow (\mathcal{C}^* \cap x^\perp)^*$ the dual regularization step. When one of these steps executes, we let $\mathcal{C}'$ and $\mathcal{C}''$ denote the cone $\mathcal{C}$ before and after execution, respectively.

*Statement 4.* We will show that $\mathbf{P}(\mathcal{C}')$ satisfies (4a) or (4b) when a dual regularization step executes. Using this, we then show that $\mathbf{P}(\mathcal{K})$ also satisfies (4a) or (4b). We use the following facts from Theorem 4.2.1: when the dual regularization step executes, all points in $\mathbf{H}(\mathcal{C}')$ satisfy $\tau = \kappa = 0$; and, when the dual regularization step executes, no facial reduction certificate exists for $\mathbf{P}(\mathcal{C}')$ (since $s \in \mathcal{C}'^\perp$).

To begin, we first show that $\mathbf{P}(\mathcal{C}')$ is feasible, and hence has finite optimal value or is unbounded below. Suppose that $\mathbf{P}(\mathcal{C}')$ is infeasible. If $\{x \in \mathcal{V} : Ax = b\}$ is empty, then there exists $\hat{y}$ for which $\langle b, \hat{y}\rangle = 1$ and $A^*\hat{y} = 0$. Hence $(0, 0, \hat{y}, 0, \langle b, \hat{y}\rangle)$ is a point in $\mathbf{H}(\mathcal{C}')$ with $\kappa > 0$, which is a contradiction. On the other hand, if $\{x \in \mathcal{V} : Ax = b\}$ is nonempty, there exists a hyperplane properly seperating $\mathcal{A}_p := \{x \in \mathcal{V} : Ax = b\}$ from the relative interior of $\mathcal{C}'$. That is, there exists $\hat{s} \in \mathcal{C}'^*$ for which

$$\begin{aligned} \langle \hat{s}, x\rangle \leq 0 \quad \forall x \in x_0 + \operatorname{null} A, \\ \langle \hat{s}, z\rangle \neq 0 \text{ for some } z \in (x_0 + \operatorname{null} A) \cup \mathcal{C}', \end{aligned}$$

where $x_0 \in \mathcal{A}_p$ and $\mathcal{A}_p = x_0 + \operatorname{null} A$. This implies that $\hat{s} \in (\operatorname{null} A)^\perp = \operatorname{range} A^*$. Hence, $\hat{s} = -A^*\hat{y}$ for some $\hat{y}$, where, evidently, $\langle \hat{s}, x\rangle = -\langle b, \hat{y}\rangle \leq 0$ for all $x \in \mathcal{A}_p$. If $\langle b, \hat{y}\rangle = 0$, then $\langle \hat{s}, z\rangle \neq 0$ for some $z \in \mathcal{C}'$ by proper separation of the sets. Hence, $\hat{s}$ is a facial reduction certificate which, as mentioned above, cannot exist. On the other hand, if $\langle b, \hat{y}\rangle > 0$, then $(0, -A^*\hat{y}, \hat{y}, 0, \langle b, \hat{y}\rangle)$ is a point in $\mathbf{H}(\mathcal{C}')$ with $\kappa > 0$, which is a contradiction. Hence, $\mathbf{P}(\mathcal{C}')$ must be feasible and either has a finite optimal value or an optimal value that is unbounded below.

We have established that $\mathbf{P}(\mathcal{C}')$ is feasible and that no facial reduction certificate for $\mathbf{P}(\mathcal{C}')$ exists. Hence, by Corollary 1.3.2, $\mathbf{P}(\mathcal{C}')$ is strictly feasible. Now suppose that

$\mathbf{P}(\mathcal{C}')$ has a finite optimal value. Then, by Slater's condition, the dual $\mathbf{D}(\mathcal{C}')$ of $\mathbf{P}(\mathcal{C}')$ has an equal optimal value that is attained. Hence, if $\mathbf{P}(\mathcal{C}')$ attains its optimal value, $\mathbf{P}(\mathcal{C}')$ has a complementary solution $(x, s, y)$ where $(x, s, y, 1, 0)$ is a point in $\mathbf{H}(\mathcal{C}')$ with $\tau > 0$; a contradiction. Suppose next that the optimal value equals $-\infty$. If an improving ray $\hat{x}_{ray}$ exists, then $(\hat{x}_{ray}, 0, 0, 0, -\langle c, \hat{x}\rangle_{ray})$ is a point in $\mathbf{H}(\mathcal{C}')$ with $\kappa > 0$, a contradiction. Hence, $\mathbf{P}(\mathcal{C}')$ satisfies (4a) or (4b).

We now show $\mathbf{P}(\mathcal{K})$ also satisfies (4a) or (4b). Consider the first time the dual regularization step executes. Since the feasible sets of $\mathbf{P}(\mathcal{C}')$ and $\mathbf{P}(\mathcal{K})$ are equal, it trivially follows that $\mathbf{P}(\mathcal{K})$ satisfies (4a) if $\mathbf{P}(\mathcal{C}')$ does. If $\mathbf{P}(\mathcal{C}')$ satisfies (4b), then $\mathbf{P}(\mathcal{K})$ is clearly unbounded. Suppose then that $\mathbf{P}(\mathcal{K})$ has an improving ray $x_{ray}$. Then, for any feasible point $x_0$ and facial reduction certificate $s$ used by the primal regularization step,

$$0 = \langle s, x_0 + x_{ray}\rangle = \langle s, x\rangle_{ray}.$$

Hence, $x_{ray} \in \mathcal{C}'$ and is therefore an improving ray for $\mathbf{P}(\mathcal{C}')$, a contradiction.

For the converse direction, we will argue $\tau = \kappa = 0$ holds at each iteration unless the dual regularization step executes. Since the primal regularization step can execute only finitely many times (since $\mathcal{K}$ is finite-dimensional), the converse direction therefore follows. To begin, suppose the optimal value of $\mathbf{P}(\mathcal{K})$ is finite and unattained, i.e., suppose (4a) holds. Then $\tau = \kappa = 0$ for all $(x, s, y, \tau, \kappa) \in \operatorname{relint} \mathbf{H}(\mathcal{K})$; otherwise, either an improving ray would exist, contradicting finiteness, or a complementary solution would exist, contradicting unattainment. Since the feasible sets of $\mathbf{P}(\mathcal{C}')$ and $\mathbf{P}(\mathcal{K})$ are equal unless the dual regularization step executes, repeating this argument shows that $\tau = \kappa = 0$ unless the dual regularization step executes. A similar argument shows the claim assuming (4b).

*Statement 3a.* Strict feasibility of $\mathbf{P}(\mathcal{C}')$ was established in the proof of statement 4. This implies that $\mathbf{P}(\mathcal{C}'')$ is strictly feasible as shown by

$$\mathcal{A}_p \cap \operatorname{relint} \mathcal{C}' \subseteq \mathcal{A}_p \cap \operatorname{relint}(\mathcal{C}' + \operatorname{span} x) = \mathcal{A}_p \cap \operatorname{relint}(\mathcal{C}'^* \cap x^\perp)^* = \mathcal{A}_p \cap \operatorname{relint} \mathcal{C}''.$$

*Statement 3b.* By (3a), if dual regularization is performed, then $\mathbf{P}(\mathcal{C}')$ satisfies Slater's condition, and continues to satisfy Slater's condition at each ensuing iteration. Hence, a facial reduction certificate $s$ cannot exist at any ensuing iteration by Corollary 1.3.2.

*Statement 1.* Since $\mathcal{K}$ is finite dimensional, it trivially follows, using (3b), that both regularization steps can execute only finitely many times. Hence, the algorithm must terminate. Corollary 4.2.1 implies termination in one iteration when complementary solutions or improving rays exist for $\mathbf{P}(\mathcal{K})$ and $\mathbf{D}(\mathcal{K})$.

*Statement 2.* The optimal values of $\mathbf{P}(\mathcal{C}')$ and $\mathbf{P}(\mathcal{C}'')$ are equal when the primal regularization step executes. Similarly, the optimal values of $\mathbf{D}(\mathcal{C}')$ and $\mathbf{D}(\mathcal{C}'')$ are equal when the dual regularization step executes. Moreover, (3a) and Slater's condition imply the optimal values of $\mathbf{P}(\mathcal{C}')$ and $\mathbf{D}(\mathcal{C}')$ are equal when the dual regularization step executes. Combining these facts with (3b) shows that the optimal value of $\mathbf{P}(\mathcal{K})$ equals the optimal value of $\mathbf{P}(\mathcal{C})$ at every iteration.

*Statement 5.* When the primal regularization step executes, $\mathbf{P}(\mathcal{C}')$ and $\mathbf{P}(\mathcal{C}'')$ have equal feasible sets since $s$ is a facial reduction certificate for $\mathbf{P}(\mathcal{C}')$. When the dual regularization step executes, the feasible set $\mathcal{A}_p \cap \mathcal{C}'$ of $\mathbf{P}(\mathcal{C}')$ and the feasible set $\mathcal{A}_p \cap \mathcal{C}''$ of $\mathbf{P}(\mathcal{C}'')$ satisfy

$$\mathcal{A}_p \cap \mathcal{C}' \subseteq \mathcal{A}_p \cap \overline{\mathcal{C}' + \operatorname{span} x} = \mathcal{A}_p \cap (\mathcal{C}'^* \cap x^{\perp})^* = \mathcal{A}_p \cap \mathcal{C}''.$$

Combining these facts with (3b) shows that $\mathcal{A}_p \cap \mathcal{K} \subseteq \mathcal{A}_p \cap \mathcal{C}$. Since dual regularization is performed if and only if (4a) or (4b) hold, the claim follows. □

An illustration of the basic steps is given next on an infeasible problem with no improving ray.

**Example 4.3.1** (Weak infeasibility)**.** *Consider the following primal-dual pair, where* $\mathcal{Q}_r^{2+k} = \{(r, x) \in \mathbb{R}^2_+ \times \mathbb{R}^k : 2r_1r_2 \geq x^T x\}$ *denotes a* rotated Lorentz cone*:*

$$\begin{array}{ll} \text{minimize} & 0 \\ \text{subject to} & x_1 = 0, \\ & x_3 = 1, \\ & x_4 = 1, \\ & x \in \mathcal{Q}_r^3 \times \mathbb{R}_+ \end{array} \qquad \begin{array}{ll} \text{maximize} & y_2 + y_3 \\ \text{subject to} & \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} y_1 \\ 0 \\ y_2 \\ y_3 \end{pmatrix} = s \\ & s \in \mathcal{Q}_r^3 \times \mathbb{R}_+. \end{array} \tag{4.9}$$

*Let $\mathcal{K} = \mathcal{Q}_r^3 \times \mathbb{R}_+$. The primal problem is infeasible, but the dual problem has no improving ray.*

*The first iteration. At the first iteration, Algorithm 4.1 finds a point in* $\operatorname{relint} \mathbf{H}(\mathcal{K})$ *satisfying $\tau = \kappa = 0$, e.g., $(\hat{x}, \hat{s}, \hat{y}, \hat{\tau}, \hat{\kappa})$, where*

$$\hat{x} = (0, 1, 0, 0), \qquad \hat{s} = (1, 0, 0, 0), \qquad \hat{y} = (-1, 0, 0), \qquad \hat{\tau} = \hat{\kappa} = 0.$$

*Since $\hat{s} \notin \mathcal{K}^{\perp}$, Algorithm 4.1 regularizes the primal problem, reformulating it over $\mathcal{K} \cap \hat{s}^{\perp} = \{0\} \times \mathbb{R}_+ \times \{0\} \times \mathbb{R}_+$. This yields a new primal-dual pair $\mathbf{P}(\mathcal{K} \cap \hat{s}^{\perp})$ and*

$\mathbf{D}(\mathcal{K} \cap \hat{s}^{\perp})$ :

$$\begin{array}{ll} \text{minimize} & 0 \\ \text{subject to} & x_1 = 0, \\ & x_3 = 1, \\ & x_4 = 1, \\ & x \in \{0\} \times \mathbb{R}_+ \times \{0\} \times \mathbb{R}_+, \end{array} \qquad \begin{array}{ll} \text{maximize} & y_2 + y_3 \\ \text{subject to} & \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} y_1 \\ 0 \\ y_2 \\ y_3 \end{pmatrix} = s, \\ & s \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+. \end{array}$$

*The second iteration. At the next iteration, Algorithm 4.1 finds a point in* $\operatorname{relint} \mathbf{H}(\mathcal{K} \cap \hat{s}^{\perp})$ *satisfying* $\kappa > 0$, *e.g.,* $(\tilde{x}, \tilde{s}, \tilde{y}, \tilde{\tau}, \tilde{\kappa})$, *where*

$$\tilde{x} = (0,0,0,0), \quad \tilde{s} = (0,0,-1,0), \quad \tilde{y} = (0,1,0), \quad \tilde{\tau} = 0, \quad \tilde{\kappa} = 1.$$

*Since* $\kappa > 0$, *the algorithm terminates. Note that* $(\tilde{s}, \tilde{y})$ *is an improving ray for* $\mathbf{D}(\mathcal{K} \cap \hat{s}^{\perp})$ *showing infeasibility of* $\mathbf{P}(\mathcal{K} \cap \hat{s}^{\perp})$ *and hence of* $\mathbf{P}(\mathcal{K})$ *by Theorem 4.3.1, statement 5.*

Theorem 4.3.1 states an iteration of dual regularization will be performed by Algorithm 4.1 if and only if the primal optimal value is finite but unattained or the primal problem is unbounded but has no improving ray. The next example illustrates this latter scenario.

**Example 4.3.2** (An unbounded problem with no improving ray)**.** *Consider the following primal-dual pair also over a rotated Lorentz cone:*

$$\begin{array}{ll} \text{minimize} & x_3 \\ \text{subject to} & x_1 = 1 \\ & x \in \mathcal{Q}_r^3 \end{array} \qquad \begin{array}{ll} \text{maximize} & y \\ \text{subject to} & \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} y \\ 0 \\ 0 \end{pmatrix} = s \\ & s \in \mathcal{Q}_r^3. \end{array} \tag{4.10}$$

*The primal is unbounded but has no improving ray: if* $x$ *and* $x + d$ *are feasible, then* $d_1 = 0$*; further, if* $d \in \mathcal{Q}_r^3$, *then* $d_3 = 0$.

*The first iteration. Algorithm 4.1 finds a point in* $\operatorname{relint} \mathbf{H}(\mathcal{Q}_r^3)$ *satisfying* $\tau = \kappa = 0$, *e.g.,* $(\hat{x}, \hat{s}, \hat{y}, \hat{\tau}, \hat{\kappa})$, *where*

$$\hat{x} = (0,1,0), \;\; \hat{s} = (0,0,0), \;\; \hat{y} = 0, \;\; \hat{\tau} = \hat{\kappa} = 0.$$

*Since* $\hat{s} \in (\mathcal{Q}_r^3)^{\perp} = \{0\}^3$, *the dual problem is regularized by replacing* $\mathcal{Q}_r^3$ *with* $\mathcal{Q}_r^3 \cap \hat{x}^{\perp}$*:*

$$\begin{array}{ll} \text{minimize} & x_3 \\ \text{subject to} & x_1 = 1 \\ & x \in \mathbb{R}_+ \times \mathbb{R}^2 \end{array} \qquad \begin{array}{ll} \text{maximize} & y \\ \text{subject to} & \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} y \\ 0 \\ 0 \end{pmatrix} = s \\ & s \in \mathbb{R}_+ \times \{0\}^2. \end{array} \tag{4.11}$$

*The second iteration. A point in* $\mathbf{H}(\mathbb{R}_+ \times \mathbb{R}^2)$ *with* $\kappa > 0$ *is obtained, yielding a primal improving ray. Since* $\mathbf{P}(\mathbb{R}_+ \times \mathbb{R}^2)$ *is also feasible, we conclude it and* $\mathbf{P}(\mathcal{Q}_r^3)$ *are unbounded.*

## 4.3.2 Interpretation of outputs

Algorithm 4.1 returns a point $(x, s, y, \tau, \kappa) \in \operatorname{relint} \mathbf{H}(\mathcal{C})$ satisfying $\tau > 0$ or $\kappa > 0$, where the optimal value of the primal problem $\mathbf{P}(\mathcal{C})$ equals the optimal value of $\mathbf{P}(\mathcal{K})$, the problem of interest. By Corollary 4.2.1, we obtain a complementary solution or an improving ray for $\mathbf{P}(\mathcal{C})$ and $\mathbf{D}(\mathcal{C})$ from the output. How, then, does one interpret this output in terms of $\mathbf{P}(\mathcal{K})$? We answer this with the following corollary of Theorem 4.3.1.

**Corollary 4.3.1.** *The following statements hold about the output $(x, s, y, \tau, \kappa)$ of Algorithm 4.1, where $\mathcal{A}_p := \{x \in \mathcal{V} : Ax = b\}$ and $\theta_p \in \mathbb{R} \cup \{\pm\infty\}$ denotes the optimal value of $\mathbf{P}(\mathcal{K})$, i.e.,*

$$\theta_p = \inf\{\langle c, x\rangle : x \in \mathcal{A}_p \cap \mathcal{K}\}.$$

1. *$\tau > 0$ holds if and only if $\theta_p$ is finite. Further, $\theta_p$ is finite and attained if and only if $\tau > 0$ and Algorithm 4.1 does not execute dual regularization steps.*
2. *$\kappa > 0$ holds if and only if $\mathbf{P}(\mathcal{K})$ is infeasible ($\theta_p = \infty$) or unbounded ($\theta_p = -\infty$).*
3. *If $\tau > 0$, then $\theta_p = \frac{1}{\tau}\langle c, x\rangle = \frac{1}{\tau}\langle b, y\rangle$. If $\theta_p$ is attained, then $\frac{1}{\tau}x$ is a solution of $\mathbf{P}(\mathcal{K})$.*
4. *If $\kappa > 0$ and $\langle b, y\rangle > 0$, then $\mathbf{P}(\mathcal{K})$ is infeasible ($\theta_p = \infty$).*

Note that from this corollary, we can only conclude the optimal value of $\mathbf{P}(\mathcal{K})$ is not finite when $\kappa > 0$ and $\langle b, y\rangle \leq 0$; that is, in this situation, $\mathbf{P}(\mathcal{K})$ could be infeasible or it could be unbounded. Though it is tempting to assume that $\langle b, y\rangle > 0$ whenever $\mathbf{P}(\mathcal{K})$ is infeasible, i.e., that $(s, y)$ is a dual improving ray, Example 4.2.2 of Section 4.2.1 illustrated this is not necessarily the case. Hence, more must be done to distinguish unboundedness from infeasibility.

## 4.3.3 A complete algorithm

As just indicated, one cannot determine the optimal value of $\mathbf{P}(\mathcal{K})$ from the output $(x, s, y, \tau, \kappa)$ of Algorithm 4.1 if $\kappa > 0$ and the dual objective satisfies $\langle b, y\rangle \leq 0$; that is, $\mathbf{P}(\mathcal{K})$ could be unbounded or it could be infeasible. Fortunately, there is a simple remedy: reexecuting Algorithm 4.1 after setting the cost vector of $\mathbf{P}(\mathcal{K})$ to zero. That is, one can set $c = 0$ to obtain the following primal-dual pair:

$$\begin{array}{llcll} \text{minimize} & 0 & \qquad & \text{maximize} & \langle b, y\rangle \\ \text{subject to} & Ax = b, & & \text{subject to} & 0 - A^*y = s, \\ & x \in \mathcal{K}, & & & s \in \mathcal{K}^*,\ y \in \mathbb{R}^m, \end{array} \tag{4.12}$$

and then reexecute Algorithm 4.1. The returned solution $(\hat{x}, \hat{s}, \hat{y}, \hat{\tau}, \hat{\kappa})$ satisfies $\hat{\kappa} > 0$ if and only if the primal problem of (4.12) is infeasible or unbounded (Corollary 4.3.1). However, as this problem cannot be unbounded (since $c = 0$), it follows $\hat{\kappa} > 0$ if and only if $\mathcal{K}$ contains no solution to $Ax = b$—i.e., if and only if the original problem $\mathbf{P}(\mathcal{K})$ is infeasible. On the other hand, if $\hat{\kappa} = 0$, then $\mathbf{P}(\mathcal{K})$ must have been unbounded.

Putting everything together, we get a complete algorithm for solving $\mathbf{P}(\mathcal{K})$:

**Algorithm 4.2:** Find optimal value of $\mathbf{P}(\mathcal{K})$ and a point attaining it if one exists.

Execute Algorithm 4.1 and let $(x, s, y, \tau, \kappa)$ denote the output
**if** $\tau > 0$ **or** **(**$\kappa > 0$ **and** $\langle b, y\rangle > 0$**) then**
  **return** optimal value $\theta_p$ and point attaining $\theta_p$ (if one exists) using Corollary 4.3.1.
**else**
  Re-execute Algorithm 4.1 with $c = 0$ and let $(\hat{x}, \hat{s}, \hat{y}, \hat{\tau}, \hat{\kappa})$ denote the output.
  **return** $\begin{cases} \infty & \text{if } \hat{\kappa} > 0 \text{ and } \hat{\tau} = 0; \mathbf{P}(\mathcal{K}) \text{ is infeasible.} \\ -\infty & \text{if } \hat{\kappa} = 0 \text{ and } \hat{\tau} > 0; \mathbf{P}(\mathcal{K}) \text{ is unbounded.} \end{cases}$
**end**

To conclude, we illustrate the execution of Algorithm 4.2 with a simple example.

**Example 4.3.3.** *Consider the following primal-dual pair, where* $\mathcal{Q}_r^{2+k} = \{(r, x) \in \mathbb{R}_+^2 \times \mathbb{R}^k \mid 2r_1 r_2 \geq x^T x\}$ *denotes a* rotated Lorentz cone*:*

$$\begin{array}{ll} \text{minimize} & -x_2 \\ \text{subject to} & x_1 = 0, \\ & x_3 = 1, \\ & x_4 = 1, \\ & x \in \mathcal{Q}_r^3 \times \mathbb{R}_+, \end{array} \qquad \begin{array}{ll} \text{maximize} & y_2 + y_3 \\ \text{subject to} & \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} y_1 \\ 0 \\ y_2 \\ y_3 \end{pmatrix} = s, \\ & s \in \mathcal{Q}_r^3 \times \mathbb{R}_+. \end{array} \tag{4.13}$$

*Here, the primal and dual problems are both infeasible, but no dual improving ray exists. Indeed, all dual rays satisfy* $s_2 = 0$ *and* $y_3 \leq 0$*; since, in addition,* $s_2 = 0$ *implies* $y_2 = 0$*, we conclude that* $y_2 + y_3 > 0$ *cannot hold. On the other hand,* $x = (0, 1, 0, 0)$ *is a primal improving ray.*

*We now illustrate the execution of Algorithm 4.2. Since a primal improving ray exists and no dual improving ray exists, Algorithm 4.1 executes one iteration and returns* $(x, s, y, \tau, \kappa)$ *satisfying* $\kappa > 0$ *and* $\langle b, y\rangle = 0$*. Since* $\langle b, y\rangle = 0$*, the feasibility problem is* *constructed:*

$$\begin{array}{ll} \text{minimize} & 0 \\ \text{subject to} & x_1 = 0, \\ & x_3 = 1, \\ & x_4 = 1, \\ & x \in \mathcal{Q}_r^3 \times \mathbb{R}_+, \end{array} \qquad \begin{array}{ll} \text{maximize} & y_2 + y_3 \\ \text{subject to} & \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} y_1 \\ 0 \\ y_2 \\ y_3 \end{pmatrix} = s, \\ & s \in \mathcal{Q}_r^3 \times \mathbb{R}_+, \end{array}$$

*which matches the primal-dual pair of Example 4.3.1. Algorithm 4.1 reexecutes on this primal-dual pair (performing the steps illustrated in Example 4.3.1) and returns a point satisfying $\kappa > 0$. Since $\kappa > 0$, Algorithm 4.2 reports the primal of* (4.13) *is infeasible.*

## ■ 4.4 Numerical experiments

We conclude with numerical experiments. To be precise, we evaluate tracking of the central path (of the extended-embedding) to its limit point, which yields an element of relint $\mathbf{H}(\mathcal{K})$ by Theorem 4.2.4 and results of [66, 116]. We then verify the classification of $(x, s, y, \tau, \kappa) \in \operatorname{relint} \mathbf{H}(\mathcal{K})$ given by Corollary 4.2.2: if $\tau = \kappa = 0$ and the primal (resp., dual) fails Slater's condition, then $s$ (resp. $x$) is a facial reduction certificate; we'll also confirm there certificates are optimal on a subset of examples. On the negative side, we show regularizing without significantly changing optimal values (a key step of an Algorithm 4.1 implementation) can be difficult. We also illustrate tracking the central path of the extended-embedding is hard for SDPs described in Waki et al. [140]; we attribute this to high singularity degree and significant failure of strict complementarity. Finally, an approximate certificate is used to (informally) argue an instance of the DIMACS library [104] has an unattained optimal value. This illustrates approximate facial reduction certificates can provide useful insight, even if they cannot be used reliably for regularization.

To perform experiments, we have made trivial modifications to the interface of SeDuMi [129], which already tracks the central path of the extended-embedding (though without maintaining the slack variable $s$). We modify the interface to return a tuple $(x, y, \tau, \kappa)$; the point $(x, c\tau - A^*y, y, \tau, \kappa)$ then lies in relint $\mathbf{H}(\mathcal{K})$ to within tolerances achieved by SeDuMi. Each test case considered is an SDP, i.e., $\mathcal{K}$ equals $\mathbb{S}^n_+$, the cone of psd matrices of order $n$. Finally, in the reported results, $\|\cdot\|$ denotes the Frobenius norm when the argument is a matrix and the 2-norm otherwise; $\lambda_{\min}(\cdot)$ denotes the smallest eigenvalue of its argument.

### ■ 4.4.1 Approximate facial reduction certificates

**Weak infeasibility library** The first test cases are SDPs taken from the URL documented in [83]. For each instance, $c$ is strictly feasible for the dual and the primal is weakly infeasible (i.e., infeasible but with no dual improving ray). It follows $\tau = \kappa = 0$ for $(x, s, y, \tau, \kappa) \in \operatorname{relint} \mathbf{H}(\mathcal{K})$. Further, Corollary 4.2.2 implies that $s$ is a facial reduction certificate and that $x = 0$. The tables below confirm that these implications hold.

| set of instances | $\tau$ | $\kappa$ | $\lambda_{\min}(s)$ | $\lvert\langle b, y\rangle\rvert$ | $\lVert s\rVert$ | $\lVert x\rVert$ |
|---|---|---|---|---|---|---|
| weak_clean_10_10 | 4.8e−9 | 2.0e−8 | 7.3e−11 | 3.0e−7 | 5.6 | 1.9e−7 |
| weak_messy_10_10 | 7.5e−8 | 6.7e−8 | 7.2e−10 | 1.1e−6 | 7.6 | 8.6e−7 |
| weak_clean_20_10 | 1.3e−9 | 1.6e−8 | 1.9e−11 | 2.6e−7 | 5.3 | 1.7e−7 |
| weak_messy_20_10 | 7.0e−8 | 4.0e−7 | 7.9e−10 | 4.5e−6 | 8.8 | 3.6e−6 |

**(a)** means

| set of instances | $\tau$ | $\kappa$ | $\lambda_{\min}(s)$ | $\lvert\langle b, y\rangle\rvert$ | $\lVert s\rVert$ | $\lVert x\rVert$ |
|---|---|---|---|---|---|---|
| weak_clean_10_10 | 4.4e−9 | 2.0e−8 | 9.1e−11 | 2.5e−7 | 3.7 | 1.6e−7 |
| weak_messy_10_10 | 8.0e−8 | 7.0e−8 | 7.6e−10 | 1.3e−6 | 4.6 | 1.1e−6 |
| weak_clean_20_10 | 1.8e−9 | 1.8e−8 | 3.1e−11 | 2.8e−7 | 4.9 | 1.9e−7 |
| weak_messy_20_10 | 8.4e−8 | 4.5e−7 | 8.8e−10 | 3.7e−6 | 4.8 | 3.0e−6 |

**(b)** standard deviations

Note that in these tables, we report the mean and standard deviation taken over each set of instances, which each contain 100 SDPs.

**Finite but non-zero duality gaps** Using Algorithm 12.3 of [31], we generated instances with duality gap equal to 100. The other inputs to this algorithm are the number of equations $m$, the order of the matrices $n$, the rank $r_1$ of an optimal facial reduction certificate for the dual problem, and an additional parameter $p$ specifying structure of the constraint matrices. If $(x, s, y, \tau, \kappa) \in \operatorname{relint} \mathbf{H}(\mathcal{K})$, then $\tau = \kappa = 0$ necessarily holds given the duality gap. Further, by Corollary 4.2.2, both $x$ and $s$ are optimal facial reduction certificates, since both primal and dual problems fail Slater's condition. The results below show that (approximate) facial reduction certificates are indeed found:

| $(n, m, p, r_1)$ | $\tau$ | $\kappa$ | $\lambda_{\min}(s)$ | $\lvert\langle b, y\rangle\rvert$ | $\lVert s\rVert$ | $\lambda_{\min}(x)$ | $\lvert\langle c, x\rangle\rvert$ | $\lVert Ax\rVert$ | $\lVert x\rVert$ |
|---|---|---|---|---|---|---|---|---|---|
| $(10, 10, 5, 5)$ | 1.2e−6 | 2.3e−6 | −8.1e−14 | 4.4e−6 | 5.8e2 | −8.6e−14 | 2.2e−6 | 3.4e−4 | 1.7e3 |
| $(20, 20, 10, 5)$ | 1.2e−6 | 1.3e−6 | −2.6e−14 | 2.0e−6 | 7.8e2 | −2.2e−13 | 6.9e−7 | 4.6e−4 | 1.3e3 |
| $(40, 40, 10, 5)$ | 5.9e−7 | 5.2e−8 | −4.0e−15 | 1.0e−7 | 6.3e1 | −1.1e−13 | 4.9e−8 | 3.5e−4 | 8.0e2 |
| $(40, 40, 10, 10)$ | 8.9e−7 | 3.1e−7 | −7.4e−15 | 5.0e−7 | 3.0e2 | −9.0e−14 | 1.9e−7 | 6.5e−4 | 1.0e3 |

Estimating rank as the number of eigenvalues larger than 1e−4 yields the following table, indicating $s$ and $x$ are (approximate) optimal certificates, i.e., $\operatorname{rank} s = n - r_1 - 1$ and $\operatorname{rank} x = r_1$:
Also note that strict complementarity fails modestly for these instances: $\operatorname{rank} s + \operatorname{rank} x = n - 1$. Section 4.4.3 will illustrate large numerical error on instances failing this condition more severely.

| $(n, m, p, r_1)$ | rank $s$ | rank $x$ |
|---|---|---|
| $(10, 10, 5, 5)$ | 4 | 5 |
| $(20, 20, 10, 5)$ | 14 | 5 |
| $(40, 40, 10, 5)$ | 34 | 5 |
| $(40, 40, 10, 10)$ | 29 | 10 |

## ■ 4.4.2 Regularization using approximate certificates

Given an approximate point $(x, s, y, \tau, \kappa) \in \operatorname{relint} \mathbf{H}(\mathcal{K})$, an implementation of Algorithm 4.1 must decide if $\tau > 0$, $\kappa > 0$ or $\tau = \kappa = 0$ for an actual point in $\operatorname{relint} \mathbf{H}(\mathcal{K})$. For the case of SDP ($\mathcal{K} = \mathbb{S}^n_+$), it also must estimate the kernel of an actual facial reduction certificate; indeed, for $s \in \mathcal{K}$, the face $\mathcal{K} \cap s^\perp$ equals the set of psd matrices with range contained in the kernel of $s$ by [10].

Example 4.4.1 below shows that the computed optimal value can be sensitive to errors in these estimates. In this example, we estimate that $\tau$ and $\kappa$ are actually zero if they are less than a threshold $T_{\tau,\kappa}$ and estimate that eigenvalues of $s$ are actually zero if they are less than a threshold $T_\lambda$. Results for different thresholds $T_{\tau,\kappa}$ and $T_\lambda$ follow:

| thresholds | computed optimal value | |
|---|---|---|
| $T_{\tau,\kappa} = 5\text{e}{-5}$ | 3.148e−1 | |
| $T_{\tau,\kappa} = 5\text{e}{-9}$, $T_\lambda = 5\text{e}{-4}$ | 1.0000 | (agrees with actual value) |
| $T_{\tau,\kappa} = 5\text{e}{-9}$, $T_\lambda = 5\text{e}{-6}$ | $\infty$ | (i.e., regularized problem is infeasible) |

The mentioned example used to generate this table is now given.

**Example 4.4.1.** *We consider an SDP with optimal value 1 and a finite nonzero duality gap, given by*

$$\begin{array}{ll} \text{minimize} & \langle c, x\rangle \\ \text{subject to} & \langle a_1, x\rangle = 1, \\ & \langle a_2, x\rangle = 0, \\ & x \in \mathbb{S}^3_+, \end{array} \tag{4.14}$$

*where the data matrices $c$, $a_1$ and $a_2$ are defined in terms of an orthogonal matrix*

*$q \in \mathbb{R}^{3\times 3}$ via:*

$$q := \begin{bmatrix} 7/11 & 6/11 & 6/11 \\ 6/11 & -9/11 & 2/11 \\ -6/11 & -2/11 & 9/11 \end{bmatrix} c := q \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} q^T,$$

$$a_1 := q \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} q^T, \quad a_2 := q \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} q^T.$$

*The solver SeDuMi finds an approximate point in* $\operatorname{relint} \mathbf{H}(\mathbb{S}^3_+)$ *given by* $\tau = 7.945\text{e}{-5}$, $\kappa = 1.028\text{e}{-5}$,

$$s = \begin{bmatrix} 1.995 & -2.993 & -6.652\text{e}{-1} \\ -2.993 & 4.490 & 9.977\text{e}{-1} \\ -6.652\text{e}{-1} & 9.977\text{e}{-1} & 2.217\text{e}{-1} \end{bmatrix}, \quad x = \begin{bmatrix} 1.518 & 5.060\text{e}{-1} & 2.277 \\ 5.060\text{e}{-1} & 1.687\text{e}{-1} & 7.590\text{e}{-1} \\ 2.277 & 7.590\text{e}{-1} & 3.416 \end{bmatrix}.$$

*Picking a threshold* $T_{\tau,\kappa} = 5\text{e}{-5}$, *we estimate* $\tau > 0$ *and* $\kappa = 0$ *and the optimal value*

$$\frac{1}{\tau}\langle c, x\rangle = 3.148\text{e}{-1},$$

*which significantly differs from the actual optimal value of* (4.14).

*On the other hand, taking* $T_{\tau,\kappa} = 5\text{e}{-9}$, *we estimate* $\tau = \kappa = 0$ *and interpret* $s$ *and* $x$ *as facial reduction certificates. To regularize, we compute an eigenvalue decomposition* $\sum_{i=1}^{3} \lambda_i v_i v_i^T$ *of* $s$*:*

$$\lambda = \begin{bmatrix} 4.416\text{e}{-5} \\ -1.857\text{e}{-10} \\ 6.707 \end{bmatrix}, \quad v_1 = \begin{bmatrix} 6.364\text{e}{-1} \\ 5.455\text{e}{-1} \\ -5.455\text{e}{-1} \end{bmatrix}, v_2 = \begin{bmatrix} 5.455\text{e}{-1} \\ 1.818\text{e}{-1} \\ 8.182\text{e}{-1} \end{bmatrix}, v_3 = \begin{bmatrix} -5.455\text{e}{-1} \\ 8.182\text{e}{-1} \\ 1.818\text{e}{-1} \end{bmatrix}.$$

*We then estimate a basis for the kernel of* $s$ *by interpreting eigenvalues below a threshold* $T_\lambda$ *as zero. For* $T_\lambda = 5\text{e}{-4}$, *an estimated basis is* $u = (v_1, v_2)$, *leading to the regularized SDP*

$$\begin{array}{ll} \text{minimize} & \langle c, uxu^T\rangle \\ \text{subject to} & \langle a_1, uxu^T\rangle = 1, \\ & \langle a_2, uxu^T\rangle = 0, \\ & x \in \mathbb{S}^2_+. \end{array}$$

*Solving with SeDuMi, we compute an optimal value of* 1.0000, *which agrees with the actual optimal value of* (4.14). *On the other hand, taking* $T_\lambda = 5\text{e}{-6}$ *yields* $u = (v_2)$

*and the SDP*

$$\begin{array}{ll} \text{minimize} & \langle c, uxu^T \rangle \\ \text{subject to} & \langle a_1, uxu^T \rangle = 1, \\ & \langle a_2, uxu^T \rangle = 0, \\ & x \in \mathbb{S}^1_+. \end{array}$$

*Solving with SeDuMi, we obtain a dual improving ray; hence, we incorrectly conclude that* (4.14) *is infeasible.*

## 4.4.3 Error, singularity degree and strict complementarity

For the instances of Section 4.4.1, we were able to obtain 'good' approximations of facial reduction certificates. The next set of instances, taken from [140], illustrate this is not always the case. For the reported instances, the dual optimal value is finite but unattained. Hence, $\tau = \kappa = 0$ holds for all points in $\mathbf{H}(\mathbb{S}^n_+)$. As reported in Table 4.4, significant error in $\tau$ is observed on several instances:

| Instance | $\tau$ | $\kappa$ | $k_p$ | $n$ | rank $x$ | rank $s$ |
|---|---|---|---|---|---|---|
| `unboundDim1R2` | 1.2e−5 | 6.0e−9 | 3 | 7 | 0 | 2 |
| `unboundDim1R3` | 4.5e−7 | 3.4e−9 | 5 | 10 | 0 | 2 |
| `unboundDim1R4` | 4.3e−7 | 4.5e−9 | 7 | 13 | 0 | 2 |
| `unboundDim1R5` | 1.1† | 1.8e−9 | 9 | 16 | 0 | 2 |
| `unboundDim1R6` | 9.8e−1† | 2.5e−9 | 11 | 19 | 0 | 2 |
| `unboundDim1R7` | 7.9e−1† | 5.2e−9 | 13 | 22 | 0 | 2 |
| `unboundDim1R8` | 7.2e−1† | 3.6e−9 | 15 | 25 | 0 | 2 |
| `unboundDim1R9` | 6.9e−1† | 3.6e−9 | 17 | 28 | 0 | 2 |
| `unboundDim1R10` | 6.6e−1† | 9.2e−10 | 19 | 31 | 0 | 2 |

**Table 4.4:** $k_p$ denotes the singularity degree of the primal, $n$ the order of the semidefinite constraint, and rank $x$ and rank $s$ are actual values for any $(x, s, y, \tau, \kappa) \in$ relint $\mathbf{H}(\mathbb{S}^n_+)$. Large error is marked †.

We offer two explanations for this error based on this table. The first is the singularity degree $k_p$ (see Section 1.3.4) of the primal problem (reported in [138]) is high. Note when singularity degree is high, distance to feasibility (forward error) can be large even when residuals (backwards errors) are small [130]. The other (related) explanation is the extent to which strict complementarity fails, that is, the extent to which rank $x$ + rank $s$ is less than $n$. For $(x, s, y, \tau, \kappa) \in$ relint $\mathbf{H}(\mathbb{S}^n_+)$, we see rank $x$ + rank $s = 2 < n$ for each instance. Indeed, $x$ is the zero matrix by Corollary 4.2.2 and strict feasibility of the dual problem [140, Section 2] and the matrix $s$ has rank two by [138, Section 5.2].

We remark that singularity degree can be high and strict complementarity can fail even if $\mathbf{P}(\mathbb{S}^n_+)$ and $\mathbf{D}(\mathbb{S}^n_+)$ have complementary solutions or improving rays. In other words, the suspected causes of error can also occur when $\tau > 0$ or $\kappa > 0$ holds in the relative interior of $\mathbf{H}(\mathbb{S}^n_+)$.

### 4.4.4 Difficult instances from the DIMACS library

Mittelmann [93] makes the following remark about two instances of the DIMACS SDP library [104]:

> *In the case of the* `hinf12` *and* `hinf13` *instances the results obtained by the various codes are so different, that we cannot be sure whether these problems have in fact ever been solved.*

We will compile evidence the dual optimal value of `hinf12` is finite and unattained, explaining the difficulty of solving this instance. To make a more convincing argument, we obtain approximate points in relint $\mathbf{H}(\mathcal{K})$ using both the self-dual solver of SDPT3v4.0 [131] and SeDuMi. Properties of these points are summarized below:

| Instance (solver) | $\tau$ | $\kappa$ | $\lambda_{\min}(s)$ | $\lvert\langle b, y\rangle\rvert$ | $\lVert s\rVert$ | $\lVert x\rVert$ |
|---|---|---|---|---|---|---|
| `hinf12` (SeDuMi) | 1.5e−6 | 2.9e−8 | 4.1e−9 | −1.5e−7 | 5.5e2 | 1.2e−6 |
| `hinf12` (SDPT3) | 4.4e−11 | 1.4e−15 | 2.0e−15 | −2.0e−15 | 2.5e1 | 3.6e−11 |

These approximate points suggest the following: by Corollary 4.2.1, no complementary solution or improving ray exists (since $\tau \approx 0$ and $\kappa \approx 0$); by Corollary 4.2.2, the primal fails Slater's condition (since $\|s\| \neq 0$); by Corollary 4.2.2, the dual satisfies Slater's condition (since $\|x\| \approx 0$). Hence, these approximate points suggest either the primal is weakly infeasible or the dual optimal value is finite and unattained. But one can verify that $e_6 e_6^T \in \mathbb{S}^{24}_+$ is primal feasible (where $e_i$ is a standard basis vector of $\mathbb{R}^{24}$). Hence, we suspect that the dual optimal value is finite and unattained.

## 4.5 Conclusion

We have unified the facial reduction algorithm of Borwein and Wolkowicz with the self-dual embedding of Goldman and Tucker, bringing together both techniques to, in principle, solve arbitrary conic optimization problems. Implementing a suggested algorithm involves only conceptually simple modifications to solvers that track the central path of extended-embeddings (such as SeDuMi [129]), and these modifications only affect solver execution when both complementary solutions and improving rays do not exist. Nevertheless, numerical experiments illustrate significant practical barriers: in practice, one obtains only approximate facial reduction certificates; optimal values can be infinitely sensitive to inexact regularization; and tracking the central path can be difficult due to loss of strict complementarity. Further work is needed to better understand these barriers. We also used approximate facial reduction certificates to make an informed conjecture about an SDP instance from [104]. Clarifying the usefulness of such certificates is another topic for future research; we note that approximate certificates are used in [31] with backwards stability guarantees.

# Part II

# Jordan reduction

Chapter 5

# Minimal subspaces in Jordan reduction

We propose a new reduction method for cone programs formulated over the cone-of-squares of a Euclidean Jordan algebra that generalizes symmetry reduction and *-algebra techniques (Section 1.4). Recall in symmetry reduction, one uses group theory to construct an orthogonal projection map that satisfies the Constraint Set Invariance Conditions (Definition 1.4.1). These conditions imply the range of the projection contains solutions (if they exist). Further, in the case of semidefinite programming (SDP), the range of the projection intersected with the psd cone $\mathbb{S}^n_+$ is isomorphic to a symmetric cone. For SDP, *-algebra techniques also find projections with the same properties (Section 1.5).

Both symmetry reduction and *-algebra techniques lack algorithms for finding *optimal* projections (Section 1.5.5). This chapter resolves this issue. Specifically, we show how to find the minimum rank projection satisfying the Constraint Set Invariance Conditions. We also consider a variant that restricts to projections whose ranges contain units for Jordan multiplication. Under this unitality condition, we show the range of the minimum rank projection satisfying the Constraint Set Invariance Conditions is always a *subalgebra*. This in turn implies its intersection with the cone-of-squares is isomorphic to a symmetric cone of lower complexity. (Chapter 6 gives techniques for finding the isomorphism.) Finally, we show that minimizing rank optimizes the decomposition of the range into simple algebras.

Symmetry reduction (Chapter 1.4), applied to semidefinite programs, imposes the Constraint Set Invariance Conditions and, if the underlying group is a subgroup of the orthogonal group, the unitality condition (Chapter 1.4.4). The same is true of *-algebra techniques (Chapter 1.5). In addition, the former requires the projection equal the *Reynolds operator* of some group and the latter requires *complete positivity* of the projection. Neither of these additional conditions are implied by Constraint Set Invariance or unitality; hence, our framework is strictly more general. We, therefore, give this framework a name—*Jordan reduction*—which reflects the connection with

Euclidean Jordan Algebras.

We organize this chapter as follows. Section 5.1 reviews preliminaries. Section 5.2 characterizes projections satisfying the Constraint Set Invariance Conditions and the unitality condition. Section 5.3 shows how to minimize rank subject to these conditions. Finally, under the unitality condition, Section 5.4 shows minimizing rank optimizes the direct-sum decomposition of the range into simple algebras. (Chapter 7 presents computational results comparing techniques of this chapter to techniques developed there.)

## 5.1 Preliminaries

### 5.1.1 Constraint Set Invariance

We consider a primal-dual pair of optimization problems formulated over a symmetric cone, or, equivalently, the cone-of-squares $\mathcal{K} \subseteq \mathbf{J}$ of a Euclidean Jordan algebra $\mathbf{J}$. We let $x \circ y$ denote the product of this algebra and let $\langle \cdot, \cdot \rangle$ denote an associative inner product, i.e., an inner product satisfying $\langle x \circ y, z \rangle = \langle y, x \circ z \rangle$ for all $x, y, z \in \mathbf{J}$. The specific primal-dual pair of interest has the following form:

$$\begin{array}{llll} \text{minimize} & \langle c, x \rangle & \text{minimize} & \langle x_0, s \rangle \\ \text{subject to} & x \in x_0 + \mathcal{L}, & \text{subject to} & s \in c + \mathcal{L}^\perp, \\ & x \in \mathcal{K}, & & s \in \mathcal{K}, \end{array} \tag{5.1}$$

where $x \in \mathbf{J}$ and $s \in \mathbf{J}$ are the primal and dual decision variables, points $c \in \mathbf{J}$ and $x_0 \in \mathbf{J}$ are fixed, $\mathcal{L} \subseteq \mathbf{J}$ is a linear subspace with orthogonal complement $\mathcal{L}^\perp \subseteq \mathbf{J}$, and $x_0 + \mathcal{L}$ and $c + \mathcal{L}^\perp$ are affine sets.

Our goal is to find a subspace $\mathcal{S} \subseteq \mathbf{J}$ that contains solutions to *both* problems (if they exist). For this, we search over orthogonal projections that satisfy the Constraint Set Invariance Conditions (Definition 1.4.1) which we reproduce below:

**Definition 5.1.1** (Constraint Set Invariance Conditions)**.** *Let $P_{\mathcal{S}} : \mathbf{J} \to \mathbf{J}$ be an orthogonal projection with range equal to $\mathcal{S}$. Then, $P_{\mathcal{S}}$ satisfies the* Constraint Set Invariance Conditions *for the primal-dual pair* (5.1) *if*

*(a)* $P_{\mathcal{S}} \cdot \mathcal{K} \subseteq \mathcal{K}$, *i.e., $P_{\mathcal{S}}$ is a* positive *map,*

*(b)* $P_{\mathcal{S}} \cdot (x_0 + \mathcal{L}) \subseteq x_0 + \mathcal{L}$,

*(c)* $P_{\mathcal{S}} \cdot (c + \mathcal{L}^\perp) \subseteq c + \mathcal{L}^\perp$,

*where* $P_{\mathcal{S}} \cdot \mathcal{C} := \{P_{\mathcal{S}} x : x \in \mathcal{C}\}$ *for any subset* $\mathcal{C} \subseteq \mathbf{J}$.

Under these conditions, intersecting the primal and the dual feasible sets with $\mathcal{S}$ does not change the primal and dual optimal values (Proposition 1.4.1). Hence, one can replace the cone $\mathcal{K}$ with $\mathcal{K} \cap \mathcal{S}$ and the affine sets $x_0 + \mathcal{L}$ and $c + \mathcal{L}^\perp$ with

$$(x_0 + \mathcal{L}) \cap \mathcal{S} = P_{\mathcal{S}} x_0 + \mathcal{L} \cap \mathcal{S} \qquad \text{and} \qquad (c + \mathcal{L}^\perp) \cap \mathcal{S} = P_{\mathcal{S}} c + \mathcal{L}^\perp \cap \mathcal{S}, \tag{5.2}$$

where (b)-(c) imply the equations (5.2). It turns out this replacement yields a primal-dual pair viewing $\mathcal{S}$ as the ambient space. The next proposition—which is elementary to prove—states this formally.

**Proposition 5.1.1.** *Suppose $P_{\mathcal{S}} : \mathbf{J} \to \mathbf{J}$ satisfies the Constraint Set Invariance Conditions (Definition 5.1.1). Then, treating the subspace $\mathcal{S} \subseteq \mathbf{J}$ as the ambient space, the pair of optimization problems*

$$\begin{array}{lll} \text{minimize} & \langle P_{\mathcal{S}} c, x \rangle \\ \text{subject to} & x \in P_{\mathcal{S}} x_0 + \mathcal{L} \cap \mathcal{S}, \\ & x \in \mathcal{K} \cap \mathcal{S}, \end{array} \qquad \begin{array}{lll} \text{minimize} & \langle P_{\mathcal{S}} x_0, s \rangle \\ \text{subject to} & s \in P_{\mathcal{S}} c + \mathcal{L}^\perp \cap \mathcal{S}, \\ & s \in \mathcal{K} \cap \mathcal{S}, \end{array} \tag{5.3}$$

*is a primal-dual pair, i.e.,*

$$(\mathcal{K} \cap \mathcal{S})^* \cap \mathcal{S} = \mathcal{K} \cap \mathcal{S},$$
$$(\mathcal{L} \cap \mathcal{S})^\perp \cap \mathcal{S} = \mathcal{L}^\perp \cap \mathcal{S}.$$

*Moreover,*

1. *Primal (resp. dual) feasible points, improving rays, and optimal solutions are primal (resp. dual) feasible points, improving rays, and optimal solutions of* (5.1)*;*

2. *The primal (resp. dual) is feasible if and only the primal (resp. dual) of* (5.1) *is feasible;*

3. *The primal (resp. dual) optimal value equals the primal (resp. dual) optimal value of* (5.1)*.*

Note variants of this proposition have appeared (e.g., [45, Proposition 2] or [37, Theorem 2]). We point out the most significant difference: the assumptions made elsewhere lack the primal-dual symmetry of the Constraint Set Invariance Conditions (Definition 5.1.1). For instance, [45] breaks this symmetry by assuming $\mathcal{S}$ contains $c$. The next example illustrates the primal-dual pair (5.1) and its restriction (5.3).

**Example 5.1.1.** *Consider the following primal-dual pair of semidefinite programs:*

$$\begin{array}{ll} \text{min.} & x_1 + x_2 \\ \text{subj. to} & \\ & \begin{pmatrix} x_1 & 1 & x_3 & x_4 \\ 1 & x_2 & x_4 & -x_3 \\ x_3 & x_4 & 1 & x_5 \\ x_4 & -x_3 & x_5 & 0 \end{pmatrix} \in \mathbb{S}^4_+ \end{array} \qquad \begin{array}{ll} \text{max.} & -(s_5 + 2s_1) \\ \text{subj. to} & \\ & \begin{pmatrix} 1 & s_1 & s_2 & s_3 \\ s_1 & 1 & -s_3 & s_2 \\ s_2 & -s_3 & s_5 & 0 \\ s_3 & s_2 & 0 & s_6 \end{pmatrix} \in \mathbb{S}^4_+ \end{array}$$

*If $\mathcal{S}$ is the subspace spanned by the set $\{E_{21}+E_{12}\} \cup \{E_{ii}\}_{i=1}^3$, then $P_\mathcal{S} : \mathbb{S}^4 \to \mathbb{S}^4$ satisfies the Constraint Set Invariance Conditions (Definition 5.1.1). Hence, one obtains primal and dual optimal solutions by solving the following restrictions to $\mathcal{S}$:*

$$\begin{array}{ll} \text{minimize} & x_1 + x_2 \\ \text{subject to} & \\ & \begin{pmatrix} x_1 & 1 & 0 & 0 \\ 1 & x_2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \in \mathbb{S}^4_+, \end{array} \qquad \begin{array}{ll} \text{maximize} & -(s_5 + 2s_1) \\ \text{subject to} & \\ & \begin{pmatrix} 1 & s_1 & 0 & 0 \\ s_1 & 1 & 0 & 0 \\ 0 & 0 & s_5 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \in \mathbb{S}^4_+, \end{array}$$

*which are primal-dual pairs viewing $\mathcal{S}$ as the ambient space.*

## ■ 5.1.2 Projected reformulations

When $\mathcal{S}$ is admissible and unital, we will see that $\mathcal{K} \cap \mathcal{S}$ is isomorphic to a symmetric cone $\mathcal{C}$. In other words, there will exist an injective linear map $\Phi$ and a symmetric cone $\mathcal{C}$ satisfying

$$\mathcal{K} \cap \mathcal{S} = \Phi \cdot \mathcal{C},$$

where $\Phi \cdot \mathcal{C} := \{\Phi z : z \in \mathcal{C}\}$. We can therefore construct a projected reformulation (Section 1.2.5) of the restricted primal-dual pair (5.3) over $\mathcal{C}$. This projected reformulation has additional structure given that $P_\mathcal{S} : \mathbf{J} \to \mathbf{J}$ satisfies the Constraint Set Invariance Conditions (Definition 5.1.1). Specifically, one does not need to solve linear equations to reconstruct solutions to the original dual; one only needs to evaluate the map $\Phi(\Phi^*\Phi)^{-1}$.

**Proposition 5.1.2** (Projected reformulations)**.** *Suppose $P_\mathcal{S} : \mathbf{J} \to \mathbf{J}$ satisfies the Constraint Set Invariance Conditions (Definition 5.1.1). Let $\mathcal{W}$ be an inner product space and $\Phi : \mathcal{W} \to \mathbf{J}$ an injective linear map with adjoint $\Phi^* : \mathbf{J} \to \mathcal{W}$. Finally, suppose the range of $\Phi$ equals $\mathcal{S}$ and that $\Phi \cdot \mathcal{C} = \mathcal{K} \cap \mathcal{S}$ for a self-dual cone $\mathcal{C} \subseteq \mathcal{W}$. Then, if $\hat{x}$ and*

$\hat{s}$ *solve the primal-dual pair*

$$\begin{array}{lll} \text{minimize} & \langle \Phi^* c, \hat{x}\rangle & \text{minimize} & \langle (\Phi^*\Phi)^{-1}\Phi^* x_0, \hat{s}\rangle \\ \text{subject to} & \hat{x} \in (\Phi^*\Phi)^{-1}\Phi^* \cdot (x_0 + \mathcal{L}), & \text{subject to} & \hat{s} \in \Phi^* \cdot (c + \mathcal{L}^{\perp}), \\ & \hat{x} \in \mathcal{C}, & & \hat{s} \in \mathcal{C}, \end{array} \tag{5.4}$$

$\Phi\hat{x}$ *and* $\Phi(\Phi^*\Phi)^{-1}\hat{s}$ *solve the pair* (5.3)*—and hence the pair* (5.1)*.*

*Proof.* Under the assumption $\Phi \cdot \mathcal{C} = \mathcal{K} \cap \mathcal{S}$, a projected reformulation (Section 1.2.5) of (5.3) is

$$\begin{array}{llll} \text{minimize} & \langle \Phi^* c, \hat{x}\rangle & \text{minimize} & \langle t_0, \hat{s}\rangle \\ \text{subject to} & \hat{x} \in t_0 + \Phi^{-1}(\mathcal{L}), & \text{subject to} & \hat{s} \in \Phi^* c + \Phi^* \cdot \mathcal{L}^{\perp}, \\ & \hat{x} \in \mathcal{C}, & & \hat{s} \in \mathcal{C}, \end{array}$$

where $t_0$ is any point in the preimage $\Phi^{-1}(x_0 + \mathcal{L})$. As discussed in Section 1.2.5, the primal and dual optimal values of this projected reformulation are the same as those of (5.3). Further, $\Phi$ and $\Phi(\Phi^*\Phi)^{-1}$ map the primal and dual feasible sets onto those of (5.3) without changing the objective value. The claim follows by showing this projected reformulation is the same as (5.4).

To begin, $(\Phi^*\Phi)^{-1}\Phi^* x_0$ is in the preimage, since $\Phi(\Phi^*\Phi)^{-1}\Phi^*$ equals $P_{\mathcal{S}}$ and $P_{\mathcal{S}} x_0 \in (x_0 + \mathcal{L})$ by the Constraint Set Invariance Conditions. We are done if $(\Phi^*\Phi)^{-1}\Phi^* \cdot \mathcal{L}$ is the preimage $\Phi^{-1}(\mathcal{L})$. Under the Constraint Set Invariance Conditions, $\Phi(\Phi^*\Phi)^{-1}\Phi^* \cdot \mathcal{L} \subseteq \mathcal{L}$ by Lemma 1.4.3; hence, $(\Phi^*\Phi)^{-1}\Phi^* \cdot \mathcal{L}$ is in the preimage. On the other hand, if $\Phi z \in \mathcal{L}$, then

$$\Phi(\Phi^*\Phi)^{-1}\Phi^*\Phi z = \Phi z.$$

But $\Phi$ is injective. Hence, $(\Phi^*\Phi)^{-1}\Phi^*\Phi z = z$, showing that $z \in (\Phi^*\Phi)^{-1}\Phi^* \cdot \mathcal{L}$. □

## ■ 5.2 Admissible subspaces

If an orthogonal projection satisfies the Constraint Set Invariance Conditions (Definition 5.1.1), we say it is *admissible*.

**Definition 5.2.1.** *A subspace* $\mathcal{S}$ *is* admissible *if its orthogonal projection* $P_{\mathcal{S}} : \mathbf{J} \to \mathbf{J}$ *satisfies the Constraint Set Invariance Conditions (Definition 5.1.1).*

Note that finding an admissible subspace of minimum dimension is equivalent to minimizing rank subject to the Constraint Set Invariance Conditions. We also consider subspaces that are *unital*.

**Definition 5.2.2.** *A subspace $\mathcal{S}$ is* unital *if it contains a unit under Jordan multiplication, i.e., if there exists $\hat{e} \in \mathcal{S}$ for which*

$$x \circ \hat{e} = x$$

*for all $x \in \mathcal{S}$.*

Note that $\hat{e}$ is not necessarily the unit of the algebra $\mathbf{J}$. It is, however, always idempotent, i.e., $\hat{e} \circ \hat{e} = \hat{e}$. The next lemma shows admissible subspaces are not necessarily unital.

**Example 5.2.1** (A nonunital admissible subspace)**.** Let $x_0$ be any point in the cone-of-squares $\mathcal{K}$ with two distinct nonzero eigenvalues. Let $c = 0$ and $\mathcal{L} = \{0\}$. Then the subspace spanned by $x_0$ is admissible but not unital.

*Proof.* The orthogonal projection onto the span of $x_0$ is

$$z \mapsto \frac{1}{\langle x_0, x_0 \rangle} \langle x_0, z \rangle x_0.$$

Since $x_0 \in \mathcal{K}$, it holds that $\langle x_0, z \rangle \geq 0$ for all $z \in \mathcal{K}$ since $\mathcal{K}$ is self-dual. Hence, the projection is a positive map. The affine sets $\{x_0\}$ and $c + \mathcal{L}^\perp$ (which equals $\mathbf{J}$) are obviously invariant under this projection. Hence, the subspace spanned by $x_0$ is admissible. It is not unital since it contains no idempotent. □

We now characterize admissible subspaces and the subsets that are unital. Notably, the unital subset consists only of subalgebras. The stated results immediately lead to algorithms for finding subspaces of minimum dimension. We give these algorithms in the next section. The following theorem gives our characterizations. (We prove it quoting results which we state and prove later.)

**Theorem 5.2.1** (Main Result)**.** *Let $\mathcal{L}$, $x_0$ and $c$ be the problem data of the primal-dual pair* $\mathbf{P}$ *and* $\mathbf{D}$*. Let $P_\mathcal{L} : \mathbf{J} \to \mathbf{J}$ denote the orthogonal projection map onto the subspace $\mathcal{L}$, and let $c_\mathcal{L} = P_\mathcal{L} c$ and $x_{0,\mathcal{L}^\perp} = x_0 - P_\mathcal{L} x_0$. Finally, let $\operatorname{proj}_\mathcal{K}(x)$ denote the metric projection of $x$ onto the cone-of-squares $\mathcal{K}$, i.e.,*

$$\operatorname*{proj}_{\mathcal{K}}(x) := \operatorname*{arg\,min}_{w \in \mathcal{K}} \langle x - w, x - w \rangle.$$

*The following statements hold.*

1. *A subspace* $\mathcal{S} \subseteq \mathbf{J}$ *is admissible if and only if*

$$\mathcal{S} \ni c_{\mathcal{L}}, x_{0,\mathcal{L}^\perp},$$
$$\mathcal{S} \supseteq \{P_{\mathcal{L}} x : x \in \mathcal{S}\},$$
$$\mathcal{S} \supseteq \left\{ \operatorname*{proj}_{\mathcal{K}}(x) : x \in \mathcal{S} \right\}.$$

2. *A unital subspace* $\mathcal{S} \subseteq \mathbf{J}$ *is admissible if*

$$\mathcal{S} \ni c_{\mathcal{L}}, x_{0,\mathcal{L}^\perp},$$
$$\mathcal{S} \supseteq \{P_{\mathcal{L}} x : x \in \mathcal{S}\},$$
$$\mathcal{S} \supseteq \{x^2 : x \in \mathcal{S}\}.$$

   *The converse holds if* $\mathbf{J}$ *is special.*

*Proof.* The first statement is an immediate consequence of Theorem 5.2.2 and Theorem 5.2.4. The second is an immediate consequence of Theorem 5.2.3 and Theorem 5.2.4. We prove these theorems at the end of this section. □

For the interested reader, the remainder of this section states and proves the quoted theorems used to characterize admissible subspaces (Theorem 5.2.1).

## ■ 5.2.1 Positive projections

### General subspaces

For *any* closed, convex cone, the condition $P \cdot \mathcal{K} \subseteq \mathcal{K}$ has a characterization in terms of metric projection onto onto $\mathcal{K}$. Consider the following, which appears as [96, Corollary 1]; see also [95]). We offer an elementary proof based on the Moreau decomposition.

**Theorem 5.2.2.** *Let* $P : \mathcal{V} \to \mathcal{V}$ *be an orthogonal projection and* $\mathcal{K} \subseteq \mathcal{V}$ *a closed, convex cone. The following are equivalent.*

- $P \cdot \mathcal{K} \subseteq \mathcal{K}$.
- *The range of* $P$ *is closed under metric projection onto* $\mathcal{K}$*, i.e., it is invariant under the map*

$$x \mapsto \operatorname*{arg\,min}_{w \in \mathcal{K}} \langle x - w, x - w \rangle,$$

*Proof.* We first show the range is closed under metric projection if $P \cdot \mathcal{K} \subseteq \mathcal{K}$. To begin, let $z = \arg\min_{w \in \mathcal{K}} \langle x - w, x - w \rangle$ for $x \in \operatorname{range} P$. By the Moreau decomposition, $x = z + t$, where $-t \in \mathcal{K}^*$ and $\langle t, z \rangle = 0$. Hence,

$$\langle x, x \rangle = \langle z, z \rangle + \langle t, t \rangle$$

Since $Px = x$, we also have that $\langle x, Px\rangle = \langle x, x\rangle$, which implies

$$\begin{aligned}\langle z, z\rangle + \langle t, t\rangle = \langle (z+t), P(z+t)\rangle &= \langle z, Pz\rangle + \langle z, Pt\rangle + \langle t, Pz\rangle + \langle t, Pt\rangle \\ &= \langle z, Pz\rangle + 2\langle t, Pz\rangle + \langle t, Pt\rangle \\ &= \langle Pz, Pz\rangle + 2\langle t, Pz\rangle + \langle Pt, Pt\rangle\end{aligned}$$

This shows that $2\langle t, Pz\rangle = \langle z, z\rangle + \langle t, t\rangle - \langle Pz, Pz\rangle - \langle Pt, Pt\rangle \geq 0$, where the inequality follows because $P$ is a contraction. But since $Pz \in \mathcal{K}$ and $t \in -\mathcal{K}^*$, we also have that $2\langle t, Pz\rangle \leq 0$. We conclude $\langle t, Pz\rangle = 0$

Evidently $x = Pz+Pt$ for $Pz \in \mathcal{K}$ and $P(-t) \in \mathcal{K}^*$ where $Pz$ and $Pt$ are orthogonal. By uniqueness of the Moreau decomposition, it follows $Pz = z$ and $Pt = t$. Hence, the range of $P$ is closed under metric projection.

Now suppose the range is closed under metric projection. For $x \in \mathcal{K}$, let $Px$ have Moreau decomposition $Px = z + t$, i.e., $z \in \mathcal{K}$, $-t \in \mathcal{K}^*$ and $\langle t, z\rangle = 0$. The point $z$ is the metric projection of $Px$ onto $\mathcal{K}$. Hence, $Pz = z$, which implies $Pt = t$. Since $x \in \mathcal{K}$,

$$0 \leq \langle -t, x\rangle = \langle -Pt, x\rangle = \langle -t, Px\rangle = \langle -t, z+t\rangle = -\langle t, t\rangle,$$

showing that $t = 0$. Hence, $Px = z$ and, therefore, $Px \in \mathcal{K}$. □

Recall for cones-of-squares, metric projection operation is easily computed from the spectral decomposition (Proposition 1.6.3).

#### Unital subspaces

The projection onto a unital subspace is always positive if the range is invariant under squaring, or, equivalently, if the range is a subalgebra. The converse holds if $\mathbf{J}$ is special. This section proves these facts. Analogous results for complex Jordan algebras are in [126] [127]; indeed, we will use an identical argument to show the converse direction. See also [96].

We first need the following two lemmas. The first characterizes subalgebras in terms of the spectral decomposition.

**Lemma 5.2.1.** *Let $\mathbf{J}$ be a Euclidean Jordan algebra, and write the spectral decomposition of nonzero $x \in \mathbf{J}$ as*

$$x = \sum_{f \in F_x} \lambda_f f,$$

*where $F_x \subset \mathbf{J}$ is a set of pairwise orthogonal idempotents and each $\lambda_f \in \mathbb{R}$ denotes a distinct nonzero eigenvalue. For a subspace $\mathcal{S} \subseteq \mathbf{J}$, the following are equivalent.*

1. *$\mathcal{S}$ contains the set $F_x$ for all nonzero $x \in \mathcal{S}$.*

*2. $\mathcal{S}$ is a subalgebra, i.e., $\mathcal{S} \supseteq \{x^2 : x \in \mathcal{S}\}$.*

*Proof.* That statement one implies two is immediate given that $x^2 = \sum_{f \in F_x} \lambda_f^2 f$. Conversely, let $\lambda$ denote an eigenvalue of $x$ of maximum magnitude. Then, if statement two holds, the idempotent $\hat{f} = \lim_{n\to\infty}(|\lambda|^{-1}x)^{2n}$ is contained in $\mathcal{S}$. Replacing $x$ with $x - \lambda\hat{f}$ and iterating yields a set of idempotents whose span contains $F_x$; moreover, this set is contained in $\mathcal{S}$. □

**Remark 5.2.1.** *Note that if we had assumed $\mathcal{S}$ contained the identity $e$ of $\mathbf{J}$, the nontrivial direction of this lemma reduces to the fact the subspace spanned by*

$$\{e, x, x^2, \dots, x^q\}$$

*for sufficiently large $q$ contains each idempotent in $F_x$* and *an additional idempotent $e - (\sum_{f \in F_x} f)$, where $e$ is the identity of $\mathbf{J}$; see, e.g., [3, Section 11.4.1] for a proof.*

The next lemma concerns the special case of $\mathbf{J} = \mathbb{S}^n$. It shows if the projection onto a unital subspace is positive, then the subspace is a subalgebra. The proof uses the aforementioned argument from [126, Theorem 2.2.2].

**Lemma 5.2.2.** *Let $P_{\mathcal{S}} : \mathbb{S}^n \to \mathbb{S}^n$ be the orthogonal projection onto a unital subspace $\mathcal{S} \subseteq \mathbb{S}^n$. If $P_{\mathcal{S}}$ is positive, i.e., $P_{\mathcal{S}} \cdot \mathbb{S}^n_+ \subseteq \mathbb{S}^n_+$, then $\mathcal{S}$ is a subalgebra, i.e., $\mathcal{S} \supseteq \{X^2 : X \in \mathcal{S}\}$.*

*Proof.* Since $\mathcal{S}$ is unital, there exists a matrix $Q \in \mathbb{R}^{n\times r}$ (where $r \leq n$) with orthonormal columns and a subspace $\hat{\mathcal{S}} \subseteq \mathbb{S}^r$ for which $QQ^T$ is the unit of $\mathcal{S}$ and

$$\mathcal{S} = \left\{QXQ^T : X \in \hat{\mathcal{S}}\right\}.$$

Note that if $\hat{\mathcal{S}}$ contains $X$ and $X^2$, then $\mathcal{S}$ contains $QXQ^T$ and $(QXQ^T)^2$ given that

$$(QXQ^T)^2 = QXQ^TQXQ^T = QX^2Q^T.$$

Further, if the projection $P_{\mathcal{S}}$ is positive, so is $P_{\hat{\mathcal{S}}} : \mathbb{S}^r \to \mathbb{S}^r$ given that for all $X \in \mathbb{S}^r_+$

$$P_{\mathcal{S}}(QXQ^T) = QP_{\hat{\mathcal{S}}}(X)Q^T.$$

Hence, the result follows by showing $\hat{\mathcal{S}}$ is invariant under squaring. We show this applying the argument from [126, Theorem 2.2.2] and using the fact $\hat{\mathcal{S}}$ contains the identity matrix of order $r$. Dropping the subscript $\hat{\mathcal{S}}$ from $P_{\hat{\mathcal{S}}}$, we first note since $P$ is positive and $P(I) = I$, it satisfies the Kadison inequality, which states $P(X^2) - P(X)P(X) \in \mathbb{S}^r_+$ for all $X \in \mathbb{S}^r$ (e.g., Theorem 2.3.4 of [14]). Hence, for $X$ in the range of $P$

$$P(X^2) - X^2 \in \mathbb{S}^r_+.$$

Letting $Z = P(X^2) - X^2$ and taking the trace shows $\operatorname{Tr} Z = 0$:

$$\operatorname{Tr} Z = \langle I, Z\rangle = \langle P(I), Z\rangle = \langle I, P(Z)\rangle = \operatorname{Tr}\left(P^2(X^2) - P(X^2)\right) = \operatorname{Tr}\left(P(X^2) - P(X^2)\right).$$

Since $Z \in \mathbb{S}^r_+$, it has zero trace only if $Z = 0$. Hence, $P(X^2) = X^2$, showing that $X^2$ is in the range of $P$. □

We can now state and prove the desired result. The following (in part) generalizes the previous lemma from $\mathbb{S}^n$ to any special algebra by using the fact any such algebra is isomorphic to a subalgebra of $\mathbb{S}^n$ (Proposition 1.6.4).

**Theorem 5.2.3.** *Let $\mathbf{J}$ be a Euclidean Jordan algebra. Let $P_{\mathcal{S}} : \mathbf{J} \to \mathbf{J}$ be the orthogonal projection onto a unital subspace $\mathcal{S} \subseteq \mathbf{J}$. Finally, let $\mathcal{K}$ denote the cone-of-squares of $\mathbf{J}$, and consider the following statements.*

1. *The projection $P_{\mathcal{S}}$ is positive, i.e., $P_{\mathcal{S}} \cdot \mathcal{K} \subseteq \mathcal{K}$*
2. *The subspace $\mathcal{S}$ is a subalgebra, i.e., $\mathcal{S} \supseteq \{x^2 : x \in \mathcal{S}\}$.*

*Then, the implication $(2 \Rightarrow 1)$ holds. If $\mathbf{J}$ is special, these statements are equivalent.*

*Proof.* To prove $(2 \Rightarrow 1)$, consider $x \in \mathcal{K}$ and suppose $P_{\mathcal{S}}x$ is nonzero. Further, write the spectral decomposition of $P_{\mathcal{S}}x$ as

$$P_{\mathcal{S}}x = \sum_{f \in E_x} \lambda_f f$$

where $E_x \subset \mathbf{J}$ is a set of idempotents and each $\lambda_f$ is unique and nonzero. If (2) holds, then Lemma 5.2.1 implies $P_{\mathcal{S}}f = f$ for all $f \in E_x$. Hence, for each $f \in E_x$,

$$0 \le \langle f, x\rangle = \langle P_{\mathcal{S}}f, x\rangle = \langle f, P_{\mathcal{S}}x\rangle = \lambda_f \langle f, f\rangle,$$

which shows the eigenvalues of $P_{\mathcal{S}}x$ are nonnegative, i.e., that $P_{\mathcal{S}}x \in \mathcal{K}$. The unitality condition holds because $\mathcal{S}$ is a subalgebra; hence, since it can be viewed as a Euclidean Jordan Algebra, it must have a unit (Chapter 1.6.1).

To prove $(1 \Rightarrow 2)$, we note that since $\mathbf{J}$ is special, there exists a subalgebra $\mathcal{J} \subseteq \mathbb{S}^q$ and a unital subspace $\hat{\mathcal{S}} \subseteq \mathcal{J}$ for which

$$\Phi \cdot \mathbf{J} = \mathbb{S}^q \cap \mathcal{J}, \qquad \Phi \cdot \mathcal{S} = \hat{\mathcal{S}},$$

where $\Phi : \mathbf{J} \to \mathbb{S}^q$ is an injective homomorphism (Proposition 1.6.4). We claim if $\hat{\mathcal{S}}$ is invariant under squaring, so is $\mathcal{S}$. Indeed, if

$$\Phi(x^2) = \Phi(x) \circ \Phi(x) \in \hat{\mathcal{S}},$$

then $x^2 \in \mathcal{S}$ by injectivity. The claim follows by Lemma 5.2.2 if we can show that $P_{\hat{\mathcal{S}}}$ is positive. First note that

$$P_{\hat{\mathcal{S}}} = P_{\hat{\mathcal{S}}} P_{\mathcal{J}} = \Phi P_{\mathcal{S}} \Phi^+ P_{\mathcal{J}},$$

where $\Phi^+ : \mathbb{S}^q \to \mathbf{J}$ denotes the pseudo inverse of $\Phi$. Hence, $P_{\mathcal{J}}$ is positive since it is the projection onto a subalgebra; specifically, it is positive by the same argument used to show $(2 \Rightarrow 1)$. Further, $\Phi P_{\mathcal{S}} \Phi^+$ restricted to the range of $P_{\mathcal{J}}$ is positive: $\Phi^+$ maps squares to squares since $\Phi$ is an injective homomorphism; $P_{\mathcal{S}}$ is positive by assumption; $\Phi$ maps squares to squares since it is a homomorphism.

□

We cannot rule out that statements 1 and 2 are equivalent without the special assumption. In fact, we conjecture this is true.

**Conjecture 5.2.1.** *For any Euclidean Jordan algebra, the orthogonal projection onto a unital subspace is positive if and only if the subspace is a subalgebra.*

We briefly mention why the proof for special algebras fails to prove this conjecture. Specifically, Lemma 5.2.2 relies on Kadison's inequality $P(X^2) - P(X)P(X) \in \mathbb{S}^n_+$ for positive maps $P : \mathbb{S}^n \to \mathbb{S}^n$ satisfying $P(I) = I$. It is unclear if this inequality generalizes to the exceptional algebra.

## ■ 5.2.2 Invariant affine subspaces of projections

The Constraint Set Invariance Conditions (Definition 5.1.1) require that the affine sets $x_0 + \mathcal{L}$ and $c + \mathcal{L}^\perp$ contain their images under the orthogonal projection map $P_{\mathcal{S}} : \mathbf{J} \to \mathbf{J}$. We now characterize these containments in terms of the range $\mathcal{S}$. The first of two lemmas yielding this characterization follows:

**Lemma 5.2.3.** *For affine sets $x_0 + \mathcal{L}$ and $c + \mathcal{L}^\perp$, let $x_{0,\mathcal{L}^\perp} \in \mathbf{J}$ and $c_{\mathcal{L}} \in \mathbf{J}$ denote the projections of $x_0 \in \mathbf{J}$ and $c \in \mathbf{J}$ onto the subspaces $\mathcal{L}^\perp$ and $\mathcal{L}$, respectively. Let $P_{\mathcal{S}} : \mathbf{J} \to \mathbf{J}$ denote the orthogonal projection onto a subspace $\mathcal{S}$ of $\mathbf{J}$. Then,*

1. *$P_{\mathcal{S}} \cdot (x_0 + \mathcal{L}) \subseteq x_0 + \mathcal{L}$ if and only if $P_{\mathcal{S}} x_{0,\mathcal{L}^\perp} = x_{0,\mathcal{L}^\perp}$ and $P_{\mathcal{S}} \cdot \mathcal{L} \subseteq \mathcal{L}$.*
2. *$P_{\mathcal{S}} \cdot (c + \mathcal{L}^\perp) \subseteq c + \mathcal{L}^\perp$ if and only if $P_{\mathcal{S}} c_{\mathcal{L}} = c_{\mathcal{L}}$ and $P_{\mathcal{S}} \cdot (\mathcal{L}^\perp) \subseteq \mathcal{L}^\perp$.*

*Proof.* We show only the first statement, noting the second has identical proof. To begin, first note $P_{\mathcal{S}}$—being an orthogonal projection—is a contraction with respect the norm $\sqrt{\langle x, x \rangle}$; further, $x_{0,\mathcal{L}^\perp}$ is the unique minimizer of this norm over $x_0 + \mathcal{L}$. Hence, if $P_{\mathcal{S}} \cdot (x_0 + \mathcal{L}) \subseteq x_0 + \mathcal{L}$, then $P_{\mathcal{S}} x_{0,\mathcal{L}^\perp} = x_{0,\mathcal{L}^\perp}$; in addition, since $x_0 + \mathcal{L} = x_{0,\mathcal{L}^\perp} + \mathcal{L}$

$$x_{0,\mathcal{L}^\perp} + P_{\mathcal{S}} \cdot (\mathcal{L}) = P_{\mathcal{S}} \cdot (x_{0,\mathcal{L}^\perp} + \mathcal{L}) \subseteq x_{0,\mathcal{L}^\perp} + \mathcal{L},$$

which implies $P_{\mathcal{S}} \cdot (\mathcal{L}) \subseteq \mathcal{L}$. The converse direction follows from the fact that $x_0 + \mathcal{L} = x_{0,\mathcal{L}^\perp} + \mathcal{L}$.

□

We now characterize invariance of $\mathcal{L}$ and $\mathcal{L}^\perp$ under $P_{\mathcal{S}}$. It turns out for an orthogonal projection $P_{\mathcal{S}}$ (or more generally, a self-adjoint linear map), the conditions $P_{\mathcal{S}} \cdot \mathcal{L} \subseteq \mathcal{L}$ and $P_{\mathcal{S}} \cdot (\mathcal{L}^\perp) \subseteq \mathcal{L}^\perp$ are equivalent. Indeed, for projections, even more is true: $\mathcal{L}$ is an invariant subspace of $P_{\mathcal{S}}$ if and only if $\mathcal{S}$ is an invariant subspace of $P_{\mathcal{L}}$. We capture these remarks in the following lemma, which follows from, e.g., Proposition 3.8 of [52].

**Lemma 5.2.4.** *Let $P_{\mathcal{L}} : \mathbf{J} \to \mathbf{J}$ and $P_{\mathcal{S}} : \mathbf{J} \to \mathbf{J}$ denote the orthogonal projections onto subspaces $\mathcal{L}$ and $\mathcal{S}$ of $\mathbf{J}$. The following four*[1] *statements are equivalent.*

- *$\mathcal{L}$ is an invariant subspace of $P_{\mathcal{S}}$*
- *$\mathcal{S}$ is an invariant subspace of $P_{\mathcal{L}}$*
- *$\mathcal{L}^\perp$ is an invariant subspace of $P_{\mathcal{S}}$*
- *$\mathcal{S}^\perp$ is an invariant subspace of $P_{\mathcal{L}}$*

Combining this with the previous lemma gives the desired conditions on $\mathcal{S}$:

**Theorem 5.2.4.** *For affine sets $x_0 + \mathcal{L}$ and $c + \mathcal{L}^\perp$, let $x_{0,\mathcal{L}^\perp} \in \mathbf{J}$ and $c_{\mathcal{L}} \in \mathbf{J}$ denote the projections of $x_0 \in \mathbf{J}$ and $c \in \mathbf{J}$ onto the subspaces $\mathcal{L}^\perp$ and $\mathcal{L}$, respectively. The following are equivalent.*

- *$x_0 + \mathcal{L}$ and $c + \mathcal{L}^\perp$ are invariant under the orthogonal projection $P_{\mathcal{S}} : \mathbf{J} \to \mathbf{J}$.*
- *The subspace $\mathcal{S}$ contains $c_{\mathcal{L}}$ and $x_{0,\mathcal{L}^\perp}$ and is invariant under the orthogonal projection $P_{\mathcal{L}} : \mathbf{J} \to \mathbf{J}$, i.e., $\mathcal{S}$ contains $x_{0,\mathcal{L}^\perp}, c_{\mathcal{L}}$ and $P_{\mathcal{L}} \cdot \mathcal{S}$.*

In summary, $x_0 + \mathcal{L}$ and $c + \mathcal{L}^\perp$ are invariant under $P_{\mathcal{S}}$ precisely when $\mathcal{S}$ is an invariant subspace of $P_{\mathcal{L}}$ that contains the distinguished points $c_{\mathcal{L}}$ and $x_{0,\mathcal{L}^\perp}$. Note these points minimize the norm $\sqrt{\langle x, x\rangle}$ over the affine sets $x_0 + \mathcal{L}$ and $c + \mathcal{L}^\perp$. If, for instance, $x_0 + \mathcal{L}$ is the solution set of linear equations $Ax = b$, then $x_{0,\mathcal{L}^\perp}$ equals $A^*(AA^*)^{-1}b$, the minimum norm solution of $Ax = b$.

## 5.3 Algorithms

By Theorem 5.2.1, arbitrary intersections of admissible subspaces are admissible. For special algebras, the same is true for unital admissible subspaces. This motivates the following definition.

[1]Note a fifth equivalent statement, which we will not use, is that the projections $P_{\mathcal{S}}$ and $P_{\mathcal{L}}$ commute.

**Definition 5.3.1.** *The* minimal admissible subspace $\mathcal{S}_{min}$ *is the intersection of all admissible subspaces:*

$$\mathcal{S}_{min} := \bigcap\{\mathcal{S} \subseteq \mathbf{J} : \mathcal{S} \text{ is admissible}\},$$

*The* minimal admissible-unital subspace $\mathcal{S}_{min,unit}$ *is the intersection of all admissible subspaces that are unital:*

$$\mathcal{S}_{min,unit} := \bigcap\{\mathcal{S} \subseteq \mathbf{J} : \mathcal{S} \text{ is admissible and unital}\},$$

Algorithms for finding the minimal admissible subspace $\mathcal{S}_{min}$ and the minimal admissible-unital subspace $\mathcal{S}_{min,unit}$ are essentially immediate from the characterization of admissibility (Theorem 5.2.1). An algorithm for finding $\mathcal{S}_{min}$ follows.

**Theorem 5.3.1.** *The minimal subspace $\mathcal{S}_{min}$ is the output of the following algorithm:*

$\mathcal{S} \leftarrow \text{span}\{c_{\mathcal{L}}, x_{0,\mathcal{L}^\perp}\}$
**repeat**
  $\mathcal{S} \leftarrow \mathcal{S} + P_{\mathcal{L}}(\mathcal{S})$
  $\mathcal{S} \leftarrow \mathcal{S} + \text{span}\{\text{proj}_{\mathcal{K}}(x) : x \in \mathcal{S}\}$
**until** converged.

*Proof.* By induction and Theorem 5.2.1, $\mathcal{S} \subseteq \mathcal{S}_{min}$ at each iteration. If the algorithm terminates, then $\mathcal{S}$ is admissible (Theorem 5.2.1). Hence, the reverse inclusion holds $\mathcal{S} \supseteq \mathcal{S}_{min}$ by definition of $\mathcal{S}_{min}$. Finally, the algorithm must terminate because it computes an ascending chain of subspaces and the dimension of $\mathbf{J}$ is finite. □

The same basic algorithm finds $\mathcal{S}_{min,unit}$ when $\mathbf{J}$ is special.

**Theorem 5.3.2.** *Suppose* $\mathbf{J}$ *is special. Then the minimal admissible-unital subspace $\mathcal{S}_{min,unit}$ is the output of the following algorithm:*

$\mathcal{S} \leftarrow \text{span}\{c_{\mathcal{L}}, x_{0,\mathcal{L}^\perp}\}$
**repeat**
  $\mathcal{S} \leftarrow \mathcal{S} + P_{\mathcal{L}}(\mathcal{S})$
  $\mathcal{S} \leftarrow \mathcal{S} + \text{span}\{x^2 : x \in \mathcal{S}\}$
**until** converged.

Correctness of this algorithm follows from the same argument that proves Theorem 5.3.1. Note that we need the special assumption to prove that $\mathcal{S} \subseteq \mathcal{S}_{min,unit}$ holds at each iteration. A positive answer to Conjecture 5.2.1 implies this inclusion without this assumption. Also note if $\mathbf{J}$ is *not* special, then the returned subspace is still provably admissible and unital (Theorem 5.2.1); we only lack a minimality proof.

**Remark 5.3.1.** *The Theorem 5.2.1 actually finds the minimal admissible subspace for* any *cone program, i.e., its correctness does not use the fact $\mathcal{K}$ is a cone-of-squares. This fact is used to* evaluate $\text{proj}_{\mathcal{K}}(x)$*, as we review in the next section.*

### 5.3.1 Spectral interpretation

Both algorithms perform a nonlinear operation to a subspace at each iteration. It turns out both these operations are implementable using the spectral decomposition (Section 1.6.4). As the next lemma indicates, these operations apply either the squaring or absolute value function to eigenvalues.

**Lemma 5.3.1.** *For a Euclidean Jordan algebra* $\mathbf{J}$*, consider the following subspaces*

$$\mathcal{S}_{\text{proj}} := \mathcal{S} + \text{span}\left\{\operatorname*{proj}_{\mathcal{K}}(x) : x \in \mathcal{S}\right\}, \qquad \mathcal{S}_{x^2} := \mathcal{S} + \text{span}\{x^2 : x \in \mathcal{S}\}.$$

*For* $x \in \mathbf{J}$*, let* $\sum_{f \in E_x} \lambda_f f$ *denote the spectral decomposition of* $x \in \mathbf{J}$ *and define*

$$\mathcal{S}_{|\lambda|} := \mathcal{S} + \text{span}\left\{\sum_{f \in E_x} |\lambda_f| f : x \in \mathcal{S}\right\}, \qquad \mathcal{S}_{\lambda^2} := \mathcal{S} + \text{span}\left\{\sum_{f \in E_x} \lambda_f^2 f : x \in \mathcal{S}\right\}.$$

*Then,* $\mathcal{S}_{\text{proj}} = \mathcal{S}_{|\lambda|}$ *and* $\mathcal{S}_{x^2} = \mathcal{S}_{\lambda^2}$*.*

*Proof.* To see $\mathcal{S}_{\text{proj}} = \mathcal{S}_{|\lambda|}$ , note that the subspace spanned by $x$ and $\text{proj}_{\mathcal{K}}(x)$ is spanned by $x_+$ and $x_- = x - x_+$, where

$$x_+ = \sum_{f : \lambda_f \geq 0} \lambda_f f, \qquad x_- = \sum_{f : \lambda_f < 0} \lambda_f f,$$

given that $\text{proj}_{\mathcal{K}}(x) = x_+$; see Section 1.6.4. It is therefore spanned by $x$ and $x_+ - x_-$, where $x_+ - x_- = \sum_{f \in E_x} |\lambda_f| f$.

That $\mathcal{S}_{x^2} = \mathcal{S}_{\lambda^2}$ also follows from properties of the spectral decomposition, specifically the identity

$$(\sum_{f \in E_x} \lambda_f f)^2 = \sum_{f \in E_x} \lambda_f^2 f.$$

This holds given that distinct $g, h \in E_x$ are pairwise orthogonal idempotents, i.e., $g \circ h = 0$. □

The subspaces that satisfy $\mathcal{S} = \mathcal{S}_{x^2}$ are precisely the subalgebras of $\mathbf{J}$. On the other hand, the structure of subspaces that satisfy $\mathcal{S} = \mathcal{S}_{|\lambda|}$ is not fully understood. We therefore pose the following open question.

**Problem 5.3.1.** *A subspace* $\mathcal{S} \subseteq \mathbf{J}$ *satisfies* $\mathcal{S} = \mathcal{S}_{|\lambda|}$ *if* $\mathcal{S}$ *is a subalgebra*[2] *or if* $\mathcal{S}$ *equals the span of a square* $x^2$*. Are these the only cases?*

[2] That $\mathcal{S} = \mathcal{S}_{|\lambda|}$ holds for subalgebras ($\mathcal{S} = \mathcal{S}_{x^2}$) follows from the inclusions $\mathcal{S}_{x^2} \supseteq \mathcal{S}_{|\lambda|} \supseteq \mathcal{S}$, a consequence of Lemma 5.2.1.

## ■ 5.3.2 Lattice interpretation

By Theorem 5.2.1, arbitrary intersections of admissible subspaces are admissible. The whole space $\mathbf{J}$ is also admissible. Hence, admissible subspaces are the fixed points of a *closure operator*—an idempotent, extensive, and increasing operator into the subspace lattice of $\mathbf{J}$. When $\mathbf{J}$ is special, similar statements apply to admissible subspaces that are unital. Hence, one can interpret the Theorem 5.3.1 and Theorem 5.3.2 algorithms as evaluation of a closure operator at the subspace spanned by $c_{\mathcal{L}}$ and $x_{0,\mathcal{L}^\perp}$.

## ■ 5.4 Optimal decompositions

Suppose $\mathbf{J}$ is special. The characterization of unital, admissible subspaces (Theorem 5.2.1) indicates that these subspaces are subalgebras of $\mathbf{J}$. As a consequence, each subspace $\mathcal{S}$ has an orthogonal direct sum decomposition into simple ideals $\mathcal{S} = \oplus_{i=1}^{s} \mathcal{S}_i$ (Chapter 1.6.2). Further, each ideal has a complexity measure called *rank* (which equals the number of distinct eigenvalues of a generic element; see Chapter 1.6.4).

In this section we prove this decomposition is optimal in a precise sense for $\mathcal{S}_{min,unit}$. Our statement is in terms of the *rank vector* of an algebra $\mathcal{W} = \oplus_{i=1}^{w} \mathcal{W}_i$

$$r_{\mathcal{W}} := (\operatorname{rank} \mathcal{W}_1, \operatorname{rank} \mathcal{W}_2, \ldots, \operatorname{rank} \mathcal{W}_w),$$

where $\mathcal{W}_i$ are the simple ideals. Specifically, we show the rank vector of $\mathcal{S}_{min,unit}$ and the rank vector of any other admissible, unital subspace satisfies a family of *majorization inequalities*. These inequalities, among other things, imply that $\mathcal{S}_{min,unit}$ minimizes the maximum rank and the sum-of-ranks. The precise definition of majorization and the statement of our result follow.

**Definition 5.4.1.** *The vector $x \in \mathbb{Z}^m$* weakly majorizes *$y \in \mathbb{Z}^n$ if*

$$\sum_{i=1}^{\min\{l,m\}} [x^{\downarrow}]_i \geq \sum_{i=1}^{\min\{l,n\}} [y^{\downarrow}]_i \qquad \forall l \in \{1, \ldots, \max\{m,n\}\},$$

*where $x^{\downarrow}$ and $y^{\downarrow}$ denote $x$ and $y$ with entries sorted in descending order.*

**Theorem 5.4.1** (Main result)**.** *Let $\mathbf{J}$ be Euclidean Jordan algebra that is special. Let $\mathcal{W}$ be any admissible, unital subspace. Finally, let the minimal admissible-unital subspace $\mathcal{S}_{min,unit}$ and $\mathcal{W}$ have the following decompositions into simple ideals:*

$$\mathcal{S}_{min,unit} = \oplus_{i=1}^{s} \mathcal{S}_i, \qquad \mathcal{W} = \oplus_{k=1}^{w} \mathcal{W}_k.$$

*Then, $r_{\mathcal{W}} := (\operatorname{rank} \mathcal{W}_1, \ldots, \operatorname{rank} \mathcal{W}_w)$ weakly majorizes $r_{\mathcal{S}} := (\operatorname{rank} \mathcal{S}_1, \ldots, \operatorname{rank} \mathcal{S}_s)$.*

*Proof.* By definition, $\mathcal{S}_{min,unit}$ is the intersection of all admissible, unital subspaces.

Hence, $\mathcal{W} \supseteq \mathcal{S}_{min,unit}$. The result then follows from Theorem 5.4.2 (to be stated and proven shortly).

□

Proving this theorem exploits the following fact: $\mathcal{S}_{min,unit}$ is a subalgebra of all other unital, admissible subspaces. We therefore study the rank vectors of subalgebras. We first give examples that illustrate the majorization inequalities.

## 5.4.1 Majorization examples

The following subalgebras—each parametrized by a set of variables $\{t_i\}_{i=1}^q$—satisfy $\mathcal{U}_i \supseteq \mathcal{U}_{i+1}$ and the vector of ranks $r_{\mathcal{U}_i}$ weakly majorizes $r_{\mathcal{U}_{i+1}}$:

$$\mathcal{U}_1 := \begin{pmatrix} t_1 & t_2 & 0 & 0 & 0 \\ t_2 & t_3 & 0 & 0 & 0 \\ 0 & 0 & t_4 & t_5 & t_6 \\ 0 & 0 & t_5 & t_7 & t_8 \\ 0 & 0 & t_6 & t_8 & t_9 \end{pmatrix} \qquad \mathcal{U}_2 := \begin{pmatrix} t_1 & t_2 & 0 & 0 & 0 \\ t_2 & t_3 & 0 & 0 & 0 \\ 0 & 0 & t_4 & t_5 & 0 \\ 0 & 0 & t_5 & t_7 & 0 \\ 0 & 0 & 0 & 0 & t_9 \end{pmatrix}$$

$$r_{\mathcal{U}_1} = (2,3) \qquad r_{\mathcal{U}_2} = (2,2,1)$$

$$\mathcal{U}_3 := \begin{pmatrix} t_1 & t_2 & 0 & 0 & 0 \\ t_2 & t_3 & 0 & 0 & 0 \\ 0 & 0 & t_1 & t_2 & 0 \\ 0 & 0 & t_2 & t_3 & 0 \\ 0 & 0 & 0 & 0 & t_4 \end{pmatrix} \qquad \mathcal{U}_4 := \begin{pmatrix} t_1 & 0 & 0 & 0 & 0 \\ 0 & t_1 & 0 & 0 & 0 \\ 0 & 0 & t_1 & 0 & 0 \\ 0 & 0 & 0 & t_1 & 0 \\ 0 & 0 & 0 & 0 & t_2 \end{pmatrix}$$

$$r_{\mathcal{U}_3} = (2,1) \qquad r_{\mathcal{U}_4} = (1,1)$$

Also of note are the subalgebras $\mathcal{U}_3$ and $\mathcal{U}_4$; despite having *three* nonzero blocks, $\mathcal{U}_3$ is isomorphic to a product of *two* simple algebras since its second two-by-two block is a copy of the first; similar remarks apply to $\mathcal{U}_4$.

## 5.4.2 Rank vectors of subalgebras

We now establish basic properties about the rank vectors of subalgebras. For this, we need the following technical results.

**Lemma 5.4.1.** *Let $\mathbf{J}$ be a Euclidean Jordan algebra and let $\mathcal{V} \subseteq \mathbf{J}$ be a subalgebra that is simple (viewed as an algebra). Let $\mathbf{J} = \oplus_{k=1}^{w} \mathbf{J}_k$ denote the orthogonal direct-sum decomposition of $\mathbf{J}$ into simple ideals. Finally, let $\Phi_k : \mathbf{J} \to \mathbf{J}$ denote the orthogonal projection onto $\mathbf{J}_k$. The following statements hold for all $k \in [w]$, where $[w] := \{1, \ldots, w\}$:*

1. *If $e \in \mathbf{J}$ is idempotent, then $\Phi_k e$ is idempotent.*

2. *If $e, f \in \mathbf{J}$ are idempotent and $\langle e, f \rangle = 0$, then $\langle \Phi_k e, \Phi_k f \rangle = 0$.*

3. *Suppose $e, f \in \mathcal{V}$ are idempotent and nonzero. If $\Phi_k e \neq 0$, then $\Phi_k f \neq 0$.*

*Proof.* Since $\mathbf{J}_k$ is a simple ideal, the projection map $\Phi_k$ from $\mathbf{J}$ onto $\mathbf{J}_k$ is a Jordan homomorphism by [67, Lemma 2.5.6]; hence, $\Phi_k e \circ \Phi_k e = \Phi_k e^2 = \Phi_k e$, showing the first statement.

For the second statement, recall $\mathbf{J} = \oplus_{k=1}^{w} \mathbf{J}_k$ is an orthogonal direct-sum decomposition of $\mathbf{J}$. We conclude

$$e = \sum_{k=1}^{w} \Phi_k e, \qquad f = \sum_{k=1}^{w} \Phi_k e.$$

Since $\langle \Phi_i e, \Phi_j f \rangle \geq 0$ and

$$\langle e, f \rangle = \sum_{i=1}^{w} \sum_{j=1}^{w} \langle \Phi_i e, \Phi_j f \rangle,$$

$\langle \Phi_i e, \Phi_j f \rangle = 0$ if $\langle e, f \rangle = 0$.

For the third statement, view $\mathcal{V}$ as a simple algebra and let $e = \sum_{i=1}^{q} e_i$ and $f = \sum_{j=1}^{r} f_j$ denote the decompositions of $e$ and $f$ into primitive idempotents of $\mathcal{V}$. Then, there exists $t \in \mathcal{V}$ (depending on $i$ and $j$) such that $e_i = 2t \circ (t \circ f_j) - t^2 \circ f_j$ [51, Corollary IV.2.4]. Since $\Phi_k$ is a homomorphism,

$$\begin{aligned} \Phi_k e_i &= \Phi_k (2t \circ (t \circ f_j) - t^2 \circ f_j) \\ &= \Phi_k (2t) \circ (\Phi_k t \circ \Phi_k f_j) - \Phi_k t^2 \circ \Phi_k f_j \end{aligned}$$

showing $\Phi_k f_j \neq 0$ if $\Phi_k e_i \neq 0$. Since

$$\Phi_k e = \sum_{i=1}^{q} \Phi_k e_i, \quad \Phi_k f = \sum_{j=1}^{r} \Phi_k f_j,$$

and $\Phi_k e_i$ and $\Phi_k f_j$ are idempotent and hence in the cone-of-squares, it follows $\Phi_k f \neq 0$ if $\Phi_k e \neq 0$.

□

**Theorem 5.4.2.** *Let $\mathcal{S} = \oplus_{i=1}^{s} \mathcal{S}_i$ and $\mathcal{W} = \oplus_{k=1}^{w} \mathcal{W}_k$ be Jordan subalgebras of $\mathbf{J}$, where $\mathcal{S}_i$ and $\mathcal{W}_k$ are simple ideals of $\mathcal{S}$ and $\mathcal{W}$ (viewed as algebras), respectively. Suppose $\mathcal{S} \subseteq \mathcal{W}$. The following statements hold:*

1. *For each* $k \in [w]$*, let* $I_k := \{i \in [s] : \mathcal{S}_i \not\subseteq (\mathcal{W}_k)^\perp\}$*. Then, for all* $k \in [w]$*,*

$$\operatorname{rank} \mathcal{W}_k \geq \sum_{i \in I_k} \operatorname{rank} \mathcal{S}_i.$$

2. *The vector* $r_\mathcal{W}$ *weakly majorizes* $r_\mathcal{S}$*, where*

$$r_\mathcal{W} := (\operatorname{rank} \mathcal{W}_1, \dots, \operatorname{rank} \mathcal{W}_w), \qquad r_\mathcal{S} := (\operatorname{rank} \mathcal{S}_1, \dots, \operatorname{rank} \mathcal{S}_s).$$

*Proof.* First note $\mathcal{S}_i$ contains a set $\mathcal{E}_i := \{e^i_j\}_{j=1}^{\operatorname{rank} \mathcal{S}_i}$ of pairwise-orthogonal idempotents (i.e., a Jordan frame). Further, if $i \in I_k$, then $\Phi_k(e) \neq 0$ for a nonzero idempotent $e$ in $\mathcal{S}_i$. We conclude all elements of $\{\Phi_k(e) : e \in \cup_{i \in I_k} \mathcal{E}_i\}$ are nonzero (Lemma 5.4.1-3); moreover, they are idempotent (Lemma 5.4.1-1) and pairwise orthogonal (Lemma 5.4.1-2). It follows $\mathcal{W}_k$ contains at least $\sum_{i \in I_k} \operatorname{rank} \mathcal{S}_i$ nonzero idempotents that are pairwise orthogonal. Hence, $\operatorname{rank} \mathcal{W}_k \geq \sum_{i \in I_k} \operatorname{rank} \mathcal{S}_i$.

For the second statement, we note the first implies the following: for each $l \in \max\{s, w\}$, there is a subset $T \subseteq [w]$ with $|T| \leq \min\{l, w\}$ for which

$$\sum_{k \in T} \operatorname{rank} \mathcal{W}_k \geq \sum_{k \in T} \sum_{i \in I_k} \operatorname{rank} \mathcal{S}_i \geq \sum_{i=1}^{\min\{l,s\}} [r^\downarrow_\mathcal{S}]_i.$$

Specifically, letting $\pi$ be a permutation of $[s]$ satisfying $[r^\downarrow_\mathcal{S}]_i = [r_\mathcal{S}]_{\pi(i)}$, we can pick $T$ such that $\pi(i) \in I_k$ for some $k \in T$ for all $i \in \{1, \dots, \min\{l, s\}\}$. It follows by definition of $r_\mathcal{W}$ that $\sum_{i=1}^{\min\{l,w\}} [r^\downarrow_\mathcal{W}]_i \geq \sum_{k \in T} \operatorname{rank} \mathcal{W}_k$; hence, the claim follows. □

## ■ 5.5 Conclusion

We proposed a new reduction method that finds the minimum rank orthogonal projection satisfying the Constraint Set Invariance Conditions for symmetric cone optimization problems, i.e., optimization problems formulated over the cone-of-squares of a Euclidean Jordan algebra. When an additional condition of unitality is imposed, we showed that the range of the projection is a Jordan subalgebra. Finally, we showed that minimizing rank of the projection also optimizes the direct-sum decomposition of this subalgebra into simple ideals.

Chapter 6

# Constructing isomorphisms between Euclidean Jordan algebras

In Chapter 5, we found a subalgebra $\mathcal{S} \subseteq \mathbf{J}$ containing solutions to a given cone program formulated over the cone-of-squares $\mathcal{K}$ of a Euclidean Jordan algebra $\mathbf{J}$. This enables one to select an algebra $\hat{\mathbf{J}}$ isomorphic to $\mathcal{S}$ (represented in a computationally convenient basis) whose cone-of-squares $\mathcal{C}$ satisfies

$$\mathcal{K} \cap \mathcal{S} = \{\Phi z : z \in \mathcal{C}\}$$

for some *Jordan isomorphism* $\Phi : \hat{\mathbf{J}} \to \mathcal{S}$, i.e., for some invertible linear map $\Phi$ satisfying

$$\Phi(x^2) = (\Phi x)^2 \qquad \forall x \in \hat{\mathbf{J}}.$$

In this chapter, we show how to find $\Phi$. Combined with techniques from Chapter 5, this allows one to construct a projected reformulation of a given cone program (Chapter 1.2.5) by first finding $\mathcal{S}$ and then $\Phi$.

Specifically, this chapter addresses the following fundamental topic: given bases for two isomorphic algebras, algorithmically construct an explicit isomorphism. (Note testing isomorphism between Euclidean Jordan algebras is easily done with linear algebra; see, e.g., Section 6.1.3.) The structure of isomorphisms arises from basic Jordan algebra theory and early work of Jacobson [73]. To elaborate, an isomorphism is a direct-sum of maps between simple algebras, since any Euclidean Jordan algebra equals a direct-sum of its simple ideals. In turn, isomorphisms between simple algebras are isometries whose form depends on the algebras' common *rank*. For algebras of rank two or less, they are simply isometries that map the identity to the identity—a basic result we will show with elementary arguments. For algebras of rank three or more, they arise from the composition of two types of isometries, which we show building on results of Jacobson [73]. An isometry of the first type is constructed from the multiplication operators of so-called *Jordan matrix units* [91]. An isometry of the second type is an isomorphism between *coordinate algebras*—Euclidean Hurwitz algebras that arise by equipping a subspace of each Jordan algebra with a new product (e.g., Chapter V of [51]). Note coordinate algebras and Jordan matrix units were introduced by Jacobson [73, 75] to prove his *coordinization theorem*. Indeed, we directly generalize an isomorphism constructed in this proof, removing a particular matrix algebra assumption.

After establishing the structure of Jordan isomorphisms, we'll give an explicit algorithm whose implementation requires only a basis for each algebra. Subroutines include a novel algorithm for finding Jordan matrix units and a subroutine for finding coordinate-algebra isomorphisms. This latter subroutine, essentially due to Jacobson [74, Section 3], iteratively constructs an isomorphism via the *Cayley-Dickson construction*. It makes no reference to the isomorphism class of the coordinate algebra (which is either the real numbers, the complex numbers, the quaternions, or the octonions), leading to simpler implementations.

From an applications point of view, our algorithm serves the same purpose as the *-algebra methods [89, 42] used for block diagonalizing semidefinite programs within the framework of symmetry reduction. These algorithms are insufficient for finding $\Phi$ in general. On the other hand, the algorithm we present can be seen as an alternative to these methods with a few appealing properties. One, we make no assumption on the algebras other than the availability of a basis, whereas these algorithms assume canonical matrix representations for one of the algebras. Further, unlike [89], no 'case' statements are employed for the isomorphism class; as mentioned above, we treat every class in a unified way using the Cayley-Dickson construction. Finally, we show how to find sparse isomorphisms assuming the availability of diagonal idempotents when $\mathbf{J}$ is a matrix algebra.

The chapter is organized as follows. Section 6.1 gives preliminaries. Section 6.2 states the complete algorithm, and the next three sections fill in missing details: Section 6.3 shows how to decompose an algebra into simple ideals and Sections 6.4-6.5 show how to construct isomorphisms between simple algebras. Section 6.6 addresses sparsity. Section 6.7 gives conclusions. Technical lemmas appear in an appendix (Section 6.8).

## ■ 6.1 Preliminaries

As in Section 1.6, $\mathbf{J}$ denotes a Euclidean Jordan algebra with product $x \circ y$ and inner product $\langle x, y\rangle_{\mathbf{J}}$ satisfying

$$\langle z \circ x, y\rangle_{\mathbf{J}} = \langle x, z \circ y\rangle_{\mathbf{J}}$$

for all $x, y, z \in \mathbf{J}$. Materials needed by this chapter from Section 1.6 include the decomposition into simple ideals and the classification of simple algebras, which include $\mathbf{H}_n(\mathbb{R})$, $\mathbf{H}_n(\mathbb{C})$, and $\mathbf{H}_n(\mathbb{H})$, i.e., the Hermitian matrices of order $n$ with real, complex or quaternion entries; the spin-factor algebra; and the exceptional algebra $\mathbf{H}_3(\mathbb{O})$, where

$\mathbb{O}$ denotes the octonions. We also need material not yet discussed. Specifically, we need the Peirce decomposition and the isomorphism test that it affords, which we overview next.

### 6.1.1 Peirce decomposition

Recall an element $x \in \mathbf{J}$ is *idempotent* if $x \circ x = x$. A set of pairwise orthogonal idempotents induces a canonical decomposition of $\mathbf{J}$ called the *Peirce decomposition*.

**Lemma 6.1.1** (Peirce Decomposition and Multiplication Rules)**.** *[67, 2.6.5] Let* $\mathbf{J}$ *denote a Euclidean Jordan algebra with identity* $e$. *Let* $\{e_i\}_{i=1}^n$ *be a set of pairwise orthogonal idempotents satisfying* $\sum_{i=1}^n e_i = e$. *For all* $(i,j) \in [n] \times [n]$*, define the* Peirce space $\mathbf{J}_{ij}$ *via*

$$\mathbf{J}_{ij} := \left\{ x \in \mathbf{J} : e_i \circ x = \frac{1}{2}x,\ \ e_j \circ x = \frac{1}{2}x \right\} \text{ if } i \neq j, \qquad \mathbf{J}_{ii} := \{x \in \mathbf{J} : e_i \circ x = x\}.$$

*Then,* $\mathbf{J}_{ij} = \mathbf{J}_{ji}$. *Further,* $\mathbf{J}$ *equals an orthogonal direct-sum* $\bigoplus_{1 \le i \le j \le n} \mathbf{J}_{ij}$. *Finally, the following* Peirce Multiplication Rules *hold:*

- $\mathbf{J}_{ij} \circ \mathbf{J}_{kl} = \{0\}$ *if* $\{i,j\} \cap \{k,l\} = \emptyset$,
- $\mathbf{J}_{ij} \circ \mathbf{J}_{jk} \subseteq \mathbf{J}_{ik}$ *if* $i, j, k$ *are all distinct,*
- $\mathbf{J}_{ij} \circ \mathbf{J}_{ij} \subseteq \mathbf{J}_{ii} + \mathbf{J}_{jj}$,
- $\mathbf{J}_{ii} \circ \mathbf{J}_{ij} \subseteq \mathbf{J}_{ij}$,

*where* $\mathbf{J}_{ij} \circ \mathbf{J}_{kl} := \{x \circ y : x \in \mathbf{J}_{ij}, y \in \mathbf{J}_{kl}\}$.

As the following example illustrates, the Peirce decomposition is analogous to a partition of a matrix into diagonal and off-diagonal blocks. Indeed, we will refer to $\mathbf{J}_{ij}$ as an *off-diagonal* Peirce space when $i \neq j$.

**Example 6.1.1.** *Let* $\mathbf{J}$ *denote* $\mathbf{H}_3(\mathbb{R})$*, the Jordan algebra of real symmetric matrices of order three with product* $x \circ y = \frac{1}{2}(xy + yx)$. *Orthogonal idempotents* $\{e_1, e_2\}$ *summing to* $e$ *and the induced Peirce spaces are*

$$e_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \mathbf{J}_{11} = \begin{bmatrix} * & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{J}_{22} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{bmatrix} \mathbf{J}_{12} = \begin{bmatrix} 0 & * & * \\ * & 0 & 0 \\ * & 0 & 0 \end{bmatrix},$$

*where* $\mathbf{J}_{ij}$ *is the subspace of matrices with indicated sparsity pattern.*

The Peirce decomposition also allows one to break multiplication into pieces. Writing $\sum_{ij} x_{ij}$ to mean $\sum_{i=1}^{n}\sum_{j=i}^{n} x_{ij}$, we have by bilinearity of $\circ$ that

$$x \circ y = \left(\sum_{ij} x_{ij}\right) \circ \left(\sum_{kl} y_{kl}\right) = \sum_{ij}\sum_{kl} x_{ij} \circ y_{kl},$$

where $x_{ij}$ denotes the $ij$-component of $x$ with respect to the direct-sum decomposition $\bigoplus_{1\leq j\leq i\leq n} \mathbf{J}_{ij}$. This will be convenient for constructing isomorphisms, since we can consider the products $x_{ij} \circ y_{kl}$ between Peirce components separately and make use of the Peirce Multiplication Rules (Lemma 6.1.1). Finally, the Peirce component $x_{ij}$ satisfies

$$x_{ij} = \begin{cases} 2e_i \circ (e_i \circ x) - e_i \circ x, & i = j \\ 2e_i \circ (x \circ e_j) + 2(e_i \circ x) \circ e_j & i \neq j. \end{cases}$$

Note the maps yielding the Peirce components have names in the literature. Specifically, $x_{ii}$ equals the *quadratic representation* of $e_i$ evaluated at $x$. Similarly, $x_{ij}$ is proportional to the *Jordan-triple-product* of $x$ with the set of orthogonal idempotents $\{e_i, e_j\}$.

## ■ 6.1.2 Peirce decompositions from Jordan frames

Recall a *Jordan frame* is a set of pairwise orthogonal idempotents that are each primitive, meaning each cannot be written as the sum of two distinct idempotents (Section 1.6). The Peirce decomposition induced by a Jordan frame has additional properties.

**Lemma 6.1.2.** *Let $\mathbf{J}$ be a Euclidean Jordan algebra and let $\{\mathbf{J}_{ij}\}_{i,j=1}^{n}$ denote the set of Peirce spaces induced by a Jordan frame $\{e_i\}_{i=1}^{n}$. For all distinct $i, j \in [n]$, the following statements hold.*

*(a) [67, 2.9.4]. The subspace $\mathbf{J}_{ii}$ is one-dimensional, specifically, $\mathbf{J}_{ii} = \{\lambda e_i : \lambda \in \mathbb{R}\}$.*

*(b) [51, Corollary IV.2.6]. If $\mathbf{J}$ is simple, there is an constant $d_{\mathbf{J}}$ for which $\dim \mathbf{J}_{ij} = d_{\mathbf{J}}$ for all $i \neq j$. The constant $d_{\mathbf{J}}$ does not depend on $\{i, j\}$ or the Jordan frame.*

The following example illustrates Lemma 6.1.2.

**Example 6.1.2.** *Let $\mathbf{J} = \mathbf{H}_3(\mathbb{R})$, the Jordan algebra of real symmetric matrices of order three with product $x \circ y = \frac{1}{2}(xy + yx)$. A Jordan frame $\{e_1, e_2, e_3\}$ and the induced Peirce*

| $\dim \mathbf{J}_{ij}$ | Iso. class |
|---|---|
| 1 | $\mathbf{H}_n(\mathbb{R})$ |
| 2 | $\mathbf{H}_n(\mathbb{C})$ |
| 4 | $\mathbf{H}_n(\mathbb{H})$ |
| 8 | $\mathbf{H}_3(\mathbb{O})$ |

| | algebra |
|---|---|
| $\mathbb{R}$ | real numbers |
| $\mathbb{C}$ | complex numbers |
| $\mathbb{H}$ | quaternions |
| $\mathbb{O}$ | octonions |

**Table 6.1:** Isomorphism class as function of off-diagonal ($i \neq j$) Peirce-space dimension for simple algebras of rank $n \geq 3$. Rank-one and rank-two algebras are isomorphic to the real numbers and the spin-factors, respectively. Hermitian matrices of order $n$ with entries from $\mathbb{T}$ are denoted $\mathbf{H}_n(\mathbb{T})$.

*spaces are*

$$e_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} e_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \mathbf{J}_{11} = \begin{bmatrix} * & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{J}_{22} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & * & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{J}_{33} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & * \end{bmatrix}$$

$$e_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{J}_{12} = \begin{bmatrix} 0 & * & 0 \\ * & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{J}_{13} = \begin{bmatrix} 0 & 0 & * \\ 0 & 0 & 0 \\ * & 0 & 0 \end{bmatrix} \mathbf{J}_{23} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & * \\ 0 & * & 0 \end{bmatrix}$$

*where $\mathbf{J}_{ij}$ is the subspace of matrices with indicated sparsity pattern. Note each $\mathbf{J}_{ii}$ has dimension one and all $\mathbf{J}_{ij}$ have the same dimension, consistent with Lemma 6.1.2.*

### ■ 6.1.3 Isomorphic simple algebras

Recall the rank of a simple algebra equals the cardinality of a Jordan frame. For simple algebras, rank and the dimension of off-diagonal Peirce spaces yield an isomorphism test ([51, p.97]):

**Proposition 6.1.1.** *Let $\mathbf{J}^A$ and $\mathbf{J}^B$ be simple Euclidean Jordan algebras with ranks at least two. Let $\mathbf{J}^A_{ij} \subset \mathbf{J}^A$ be any off-diagonal Peirce space induced by a Jordan frame, and similarly define $\mathbf{J}^B_{kl} \subset \mathbf{J}^B$. Then, $\mathbf{J}^A$ and $\mathbf{J}^B$ are isomorphic if and only if they have the same rank and* $\dim \mathbf{J}^A_{ij} = \dim \mathbf{J}^B_{kl}$.

Note that rank and off-diagonal Peirce space dimension also indicate isomorphism class (Table 6.1).

## ■ 6.2 Algorithm

A complete procedure for constructing an isomorphism appears in Algorithm 6.1.
The remainder of this chapter explains the steps, which are readily implemented using linear algebra. In particular, one obtains a Jordan frame from the spectral decomposi-

**Algorithm 6.1:** Finds isomorphism between Euclidean Jordan Algebras

```
Input: two isomorphic algebras J^A and J^B
Output: an isomorphism Φ : J^A → J^B
begin
    for each K ∈ {A, B} do
        Find Jordan frame for J^K and its simple partition (Definition 6.3.1)
        Get simple ideals {J_i^K}_{i=1}^m of J^K from simple partition (Theorem 6.3.1)
    end
    Find a matching {(J_i^A, J_{σ(i)}^B)}_{i=1}^m of isomorphic ideals using Proposition 6.1.1.
    For each pair (J_i^A, J_{σ(i)}^B), find isomorphism Φ_i : J_i^A → J_{σ(i)}^B
    return Φ = ⊕_{i=1}^m Φ_i
end
```

tion of a *regular* element, where a random (e.g., uniform) combination of basis elements is regular. The *simple partition* of a Jordan frame is induced by an equivalence relation defined by the nonzero Peirce spaces (Section 6.3). A matching of isomorphic ideals is just a matching in a bipartite graph $G$ with node sets $\{\mathbf{J}_i^A\}_{i=1}^m$ and $\{\mathbf{J}_j^B\}_{j=1}^m$, where $\mathbf{J}_i^A$ and $\mathbf{J}_j^B$ are adjacent if and only if they are isomorphic; i.e., if and only if they have the same rank and off-diagonal Peirce-space dimension (Proposition 6.1.1, Section 6.1.3). Finally, isomorphisms between simple algebras are isometries induced by special orthogonal bases (Section 6.4) found (in part) by subroutines from Section 6.5. We now fill in missing details, beginning with identification of simple ideals.

## 6.3 Decomposition into simple ideals

Any Euclidean Jordan algebra $\mathbf{J}$ equals an orthogonal direct-sum of its simple ideals (Proposition 1.6.1). If $\mathbf{J}$ were an associative algebra, we could find these ideals using methods from [50, 72, 89, 42]. It turns out a Jordan analogue of [89] holds. It involves finding a partition of *any* Jordan frame and then constructing a new set of idempotents by summing over each partition class. The nonzero Peirce spaces associated with this new set are precisely the simple ideals.

The mentioned partition of a Jordan frame $\{e_i\}_{i=1}^n$ is defined by the nonzero Peirce spaces it induces. Formally, write $e_i \sim e_j$ if and only if $\mathbf{J}_{ij} \neq \{0\}$. The relation $\sim$ is is an equivalence relation [67, 2.9.4iv and 2.9.5]. It follows the equivalence classes of $\sim$ form a partition of $\{e_i\}_{i=1}^n$:

**Definition 6.3.1.** *The* simple partition *of a Jordan frame $\{e_i\}_{i=1}^n$ is the set of equivalence classes $P_1, \ldots, P_m$ induced by the equivalence relation $\sim$.*

From the simple partition, we can directly construct the decomposition into simple ideals:

**Theorem 6.3.1.** *Let $\{e_i\}_{i=1}^n$ be a Jordan frame of* $\mathbf{J}$ *with simple partition $P_1, \dots, P_m$. For each $P_k$, define the new idempotent $v_k = \sum_{e_i \in P_k} e_i$ and the Peirce space*

$$\mathbf{J}_k = \{x \in \mathbf{J} : v_k \circ x = x\}.$$

*Then, $\mathbf{J}_k$ is ideal of $\mathbf{J}$ and the orthogonal direct-sum decomposition holds:*

$$\mathbf{J} = \bigoplus_{k=1}^m \mathbf{J}_k.$$

*Further, viewed as an algebra, $\mathbf{J}_k$ is simple with identity element $v_k$ and Jordan frame $P_k$.*

*Proof.* Let $\oplus_{1 \le i \le j \le n} \mathbf{J}_{ij}$ denote the Peirce decomposition induced by the Jordan frame. That $\mathbf{J} = \oplus_{k=1}^m \mathbf{J}_k$ is immediate, since it equals $\oplus_{1 \le i \le j \le n} \mathbf{J}_{ij}$ with all trivial subspaces removed from the summation by definition of the simple partition.

That $\mathbf{J}_k$ is an ideal follows by applying the Peirce Multiplication Rules (Lemma 6.1.1) to the Peirce decomposition induced by $\{v_k\}_{k=1}^m$; it also follows from the combination of [67, 2.5.7], [67, 2.9.4,iv] and [67, 2.9.5]. To show $\mathbf{J}_k$ is simple, first note that $\mathbf{J}_{ij} \neq \{0\}$ if $e_i, e_j \in P_k$ by definition of $P_k$. We will show this leads to a contradiction if $\mathbf{J}_k$ is not simple. To begin, if $\mathbf{J}_k$ is not simple, then $\mathbf{J}_k = V_1 \oplus V_2$ for nontrivial ideals $V_1$ and $V_2$ of $\mathbf{J}_k$. Further $|P_k| > 1$, otherwise $\mathbf{J}_k$ is spanned by a single idempotent and has no nontrivial ideal. We claim there exists $e_i, e_j \in P_k$ satisfying $e_i \in V_1$ and $e_j \in V_2$. To see this, note $e_\ell \in V_1$ or $e_\ell \in V_2$ for each $\ell \in P_k$, otherwise $e_\ell = e_{\ell,1} + e_{\ell,2}$ for nonzero idempotents $e_{\ell,i} \in V_i$, contradicting the assumption $e_\ell$ is primitive. Further, $v_k$ cannot be contained in $V_1$ or $V_2$ since (by definition) it is the identity of $\mathbf{J}_k$; hence, there exists $e_i \in V_1$ and $e_j \in V_2$ as claimed. Picking arbitrary $x \in \mathbf{J}_{ij}$, we have from the definition of $\mathbf{J}_{ij}$ that

$$e_i \circ x = \frac{1}{2}x, \qquad e_j \circ x = \frac{1}{2}x$$

and, since $V_i$ are ideals, that $\frac{1}{2}x \in V_1 \cap V_2$. But since $V_1 \cap V_2 = \{0\}$ and since $x$ was arbitrary, we conclude $\mathbf{J}_{ij} = \{0\}$, a contradiction. Hence, $\mathbf{J}_k$ must be simple. That $v_k$ is the identity is (as mentioned) by definition. That $P_k$ is a Jordan frame is again immediate: its elements sum to $v_k$ and are primitive and pairwise-orthogonal.

□

As noted in the proof, that $\mathbf{J}_k$ is an ideal follows immediately from results of [67, Chapter 2]; hence, the contribution of Theorem 6.3.1 is that these ideals are simple. We also note that each idempotent $v_k$ is *central*—that is, $a \circ (v_k \circ b) = (a \circ v_k) \circ b$ for all $a, b \in \mathbf{J}$. (See [67, 2.9.5] and [67, 2.9.4,iv].) Finding idempotents that are central (in a related sense) is the basis of decomposition methods [50, 42] for associative algebras.

## 6.4 Isomorphisms between simple algebras

As mentioned, an isomorphism between two Jordan algebras equals a direct-sum of isomorphisms between ideals that, when viewed as algebras, are simple. Hence, this section will focus on the construction of isomorphisms between simple algebras. We let $\|x\|_{\mathbf{J}}$ denote the norm induced by $\langle x, x\rangle_{\mathbf{J}}$ with a carefully chosen rescaling:

**Definition 6.4.1.** *For a rank $n$ simple algebra $\mathbf{J}$ with identity $e$, define the norm*

$$\|x\|_{\mathbf{J}} := \sqrt{\frac{\langle x, x\rangle_{\mathbf{J}}}{\frac{2}{n}\langle e, e\rangle_{\mathbf{J}}}}.$$

The scaling $\frac{2}{n}\langle e, e\rangle_{\mathbf{J}}$ ensures that $x^2$ is idempotent if $x$ has unit norm and $x \in \mathbf{J}_{ij}$ for $i \neq j$. (Note if $\mathbf{J} = \mathbb{S}^n$, then $\|\cdot\|_{\mathbf{J}}$ is just the Frobenius norm rescaled such that $E_{ij} + E_{ji}$ has unit norm when $i \neq j$.) This and other properties of $\|\cdot\|_{\mathbf{J}}$ follow:

**Lemma 6.4.1** (Properties of $\|\cdot\|_{\mathbf{J}}$)**.** *Let $\mathbf{J}$ be a rank $n$ simple algebra and let $\{\mathbf{J}_{ij}\}_{i,j=1}^n$ be the Peirce spaces induced by a Jordan frame $\{e_i\}_{i=1}^n$. Then,*

*(a) For all $x \in \mathbf{J}_{ij}$ with $i \neq j$, the identity $x^2 = \|x\|_{\mathbf{J}}^2(e_i + e_j)$ holds;*

*(b) For all $x \in \mathbf{J}_{ii}$, the identity $x^2 = 2(\|x\|_{\mathbf{J}}^2)e_i$ holds;*

*(c) For $x, y \in \mathbf{J}$,*

$$\frac{\langle x, y\rangle_{\mathbf{J}}}{\frac{1}{n}\langle e, e\rangle_{\mathbf{J}}} = \|x + y\|_{\mathbf{J}}^2 - \|x\|_{\mathbf{J}}^2 - \|y\|_{\mathbf{J}}^2.$$

*Proof.* Statement (a) is [51, Proposition IV.1.4], statement (b) follows from the fact $e_i$ spans $\mathbf{J}_{ii}$ and has norm $\|e_i\| = \frac{1}{\sqrt{2}}$, and statement (c) is immediate from definitions. □

We will find isomorphisms between algebras $\mathbf{J}^A$ and $\mathbf{J}^B$ are always isometries with respect to $\|\cdot\|_{\mathbf{J}^A}$ and $\|\cdot\|_{\mathbf{J}^B}$, which is essentially immediate from (a)-(b). Note statement (c) indicates isometries preserve orthogonality; hence, they are mappings between appropriately scaled, orthogonal bases.

We break the description of isomorphisms into three sections. The first concerns simple algebras with rank two and the next simple algebras with rank at least three. The third gives simplifications.

### 6.4.1 Rank two algebras

As just noted, isomorphisms will necessarily be isometries. In the rank two case, this is almost sufficient. Indeed, any isometry that also maps a Jordan frame of $\mathbf{J}^A$ to a Jordan frame of $\mathbf{J}^B$ is an isomorphism:

**Theorem 6.4.1** (Isomorphisms for rank two)**.** *Suppose $\mathbf{J}^A$ and $\mathbf{J}^B$ are simple algebras of rank two. Let $\{e_1^A, e_2^A\} \subset \mathbf{J}^A$ and $\{e_1^B, e_2^B\} \subset \mathbf{J}^B$ be Jordan frames, and let $\Phi : \mathbf{J}^A \to \mathbf{J}^B$ be an invertible linear map with the following properties:*

- $\Phi e_1^A = e_1^B$ *and* $\Phi e_2^A = e_2^B$.
- $\Phi$ *is an isometry with respect to the norms* $\|\cdot\|_{\mathbf{J}^A}$ *and* $\|\cdot\|_{\mathbf{J}^B}$*, i.e.,*

$$\|\Phi x\|_{\mathbf{J}^B} = \|x\|_{\mathbf{J}^A} \qquad \forall x \in \mathbf{J}^A.$$

*Then, $\Phi(\mathbf{J}_{ij}^A) = \mathbf{J}_{ij}^B$ for all $i, j \in \{1, 2\}$. Further, $\Phi$ is a Jordan isomorphism.*

*Proof.* That $\Phi(\mathbf{J}_{ii}^A) = \mathbf{J}_{ii}^B$ follows since $\mathbf{J}_{ii}^A$ and $\mathbf{J}_{ii}^B$ are one dimensional (Lemma 6.1.2-(a)). Further, if $\Phi$ is an isometry, it preserves orthogonality by Lemma 6.4.1-(c). Hence, $\Phi(\mathbf{J}_{12}^A) \subseteq (\mathbf{J}_{11}^B)^\perp \cap (\mathbf{J}_{22}^B)^\perp = \mathbf{J}_{12}^B$. Finally, since $\Phi$ is invertible, $\Phi(\mathbf{J}_{12}^A) = \mathbf{J}_{12}^B$, otherwise $\Phi$ is not surjective.

To show that $\Phi$ is an isomorphism, we verify $\Phi(x \circ x) = \Phi(x) \circ \Phi(x)$ by checking the products between each Peirce component of $x$.

**Case** $x_{ii} \circ x_{ii}$ Since the Peirce space $\mathbf{J}_{ii}^A$ is spanned by $e_i^A$ (Lemma 6.1.2-(a)), there exists $\lambda \in \mathbb{R}$ for which $x_{ii} = \lambda e_i^A$. Hence,

$$\Phi(\lambda e_i^A \circ \lambda e_i^A) = \lambda^2 \Phi(e_i^A \circ e_i^A) = \lambda^2 \Phi e_i^A = \lambda^2 e_i^B = \lambda^2 e_i^B \circ e_i^B = \Phi(\lambda e_i^A) \circ \Phi(\lambda e_i^A).$$

**Case** $x_{ii} \circ x_{12}$ We let $x_{ii} = \lambda e_i^A$ and use the fact $e_i \circ z_{ij} = \frac{1}{2} z$ by definition of $\mathbf{J}_{ij}$:

$$\lambda \Phi(e_i^A \circ x_{12}) = \lambda \Phi(\frac{1}{2} x_{12}) = \lambda \frac{1}{2} \Phi(x_{12}) = \lambda e_1^B \circ \Phi(x_{12}) = \lambda \Phi(e_i^A) \circ \Phi(x_{12}).$$

**Case** $x_{12} \circ x_{12}$ We use Lemma 6.4.1-(a) to express $x_{12} \circ x_{12}$ in terms of $\|\cdot\|_{\mathbf{J}}$ and the assumption $\Phi$ is an isometry mapping $e_i^A$ to $e_i^B$:

$$\begin{aligned}
\Phi(x_{12} \circ x_{12}) &= \|x\|_{\mathbf{J}^A}^2 \Phi(e_1^A + e_2^A) && \text{Lemma 6.4.1-(a)} \\
&= \|x\|_{\mathbf{J}^A}^2 (e_1^B + e_2^B) && \Phi e_i^A = e_i^B \\
&= \|\Phi x\|_{\mathbf{J}^B}^2 (e_1^B + e_2^B) && \Phi \text{ is an isometry} \\
&= (\Phi x_{12}) \circ (\Phi x_{12}) && \text{Lemma 6.4.1-(a).}
\end{aligned}$$

□

Note since an inner product induces $\|\cdot\|_{\mathbf{J}^A}$ and $\|\cdot\|_{\mathbf{J}^B}$, constructing an isometry is trivial: one simply maps a suitably-scaled orthogonal basis of $\mathbf{J}^A$ (containing a Jordan frame) onto an orthogonal basis of $\mathbf{J}^B$ (containing a Jordan frame). The following example illustrates this construction.

**Example 6.4.1.** *Let $\mathbf{J}^A$ denote the rank two Jordan algebra spanned by the symmetric matrices*

$$e_1^A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, e_2^A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, t_1^A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, t_2^A = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

*with product $x \circ y = \frac{1}{2}(xy + yx)$. (Note $\mathbf{J}^A$ can be viewed as a proper subalgebra of $\mathbf{H}_4(\mathbb{R})$.) Let $\mathbf{J}^B$ denote the spin-factor algebra $\mathbb{R} \oplus \mathbb{R}^3$ spanned by*

$$e_1^B = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \end{bmatrix}^T, e_2^B = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & 0 \end{bmatrix}^T, t_1^B = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}^T, t_2^B = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}^T$$

*with product $(x_0 \oplus x) \circ (y_0 \oplus y) := (x_0 y_0 + x^T y) \oplus (x_0 y + y_0 x)$. The linear map $\Phi : \mathbf{J}^A \to \mathbf{J}^B$ satisfying*

$$\Phi e_i^A = e_i^B, \qquad \Phi t_i^A = t_i^B \qquad i \in \{1, 2\}$$

*maps the primitive idempotent $e_i^A$ to $e_i^B$ and is an isometry with respect to the norms*

$$\|x\|_{\mathbf{J}^A}^2 = \frac{1}{4} \operatorname{Tr} x^2, \qquad \|x\|_{\mathbf{J}^B}^2 = x^T x.$$

*Hence, $\Phi$ is a Jordan isomorphism.*

## ■ 6.4.2 Rank $n \geq 3$ algebras

Like the rank-two case, an isomorphism $\Phi$ between algebras with rank larger than two will still be an isometry that maps one Jordan frame to another. Unfortunately, $\Phi$ must satisfy additional conditions arising from the fact there are distinct off-diagonal Peirce spaces (i.e., $\mathbf{J}_{ij}$ and $\mathbf{J}_{kl}$ for which $\{i, j\} \neq \{k, l\}$). Nevertheless, $\Phi$ still has straightforward structure that builds upon [76]. Specifically, the mapping between $\mathbf{J}_{ij}^A$ and $\mathbf{J}_{ij}^B$ will equal the composition of isometries

$$W_{ij}^B T W_{ij}^A,$$

where $T$ is an isomorphism between *coordinate algebras* and $W_{ij}^{A/B}$ maps the Peirce space $\mathbf{J}_{ij}^{A/B}$ onto the coordinate algebra of $\mathbf{J}^{A/B}$ (and vice versa). Both $W_{ij}^{A/B}$ and the coordinate algebra are defined by *Jordan matrix units*, structured samples of each

Peirce space defined next.

#### Jordan matrix units

A set of Jordan matrix units is a Jordan frame combined with elements of each off-diagonal Peirce space that are both normalized and have consistent 'orientation.' Formally:

**Definition 6.4.2.** *[91, Chapter 6] Let $\{e_i\}_{i=1}^n$ be a Jordan frame for a simple algebra $\mathbf{J}$. Let $\{\mathbf{J}_{ij}\}_{i,j=1}^n$ be the set of Peirce spaces induced by $\{e_i\}_{i=1}^n$. Then, $\{u_{ij} \in \mathbf{J}_{ij}\}_{i,j=1}^n$ is a set of* Jordan matrix units *for $\{\mathbf{J}_{ij}\}_{i,j=1}^n$ if for all distinct $i, j, k \in [n]$*

- $u_{ii} = e_i$, $u_{ij} = u_{ji}$
- $u_{ij} \circ u_{ij} = u_{ii} + u_{jj}$ *(normalization constraint)*
- $u_{ij} \circ u_{jk} = \frac{1}{2} u_{ik}$ *(orientation constraint)*

The condition $u_{ij} \circ u_{ij} = u_{ii} + u_{jj}$ holds if and only if $\|u_{ij}\|_{\mathbf{J}} = 1$. The orientation condition $u_{ij} \circ u_{jk} = \frac{1}{2} u_{ik}$, however, is more complicated and is not satisfied simply by sampling each Peirce space and normalizing. Nevertheless, matrix units always exist [51, Lemma V.3.3] and, indeed, are not unique. The following example illustrates non-uniqueness.

**Example 6.4.2.** *Let $\mathbf{J} = \mathbf{H}_3(\mathbb{C})$, the Jordan algebra of complex Hermitian matrices of order three with product $x \circ y = \frac{1}{2}(xy + yx)$. The Jordan frame $\{e_1, e_2, e_3\}$ given by*

$$e_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad e_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

*is completed to a set of Jordan matrix units by*

$$u_{12} = \begin{bmatrix} 0 & a & 0 \\ a^* & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad u_{13} = \begin{bmatrix} 0 & 0 & b \\ 0 & 0 & 0 \\ b^* & 0 & 0 \end{bmatrix}, \quad u_{23} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & a^*b \\ 0 & b^*a & 0 \end{bmatrix}$$

*for any complex numbers $a$ and $b$ satisfying $a^*a = b^*b = 1$.*

Note this example freely sets $u_{12}$ and $u_{13}$ (via $a$ and $b$) and *generates* $u_{23}$—taking $u_{23} = 2u_{12} \circ u_{13}$. Section 6.5.1 generalizes this construction to arbitrary simple algebras.

#### Coordinate algebras

Matrix units $\{u_{ij}\}_{i,j=1}^n$ equip any off-diagonal Peirce space $\mathbf{J}_{ij}$ with identity $u_{ij}$ and product

$$(x, y) \to 8(u_{ik} \circ x) \circ (u_{jk} \circ y),$$

where $k$ is any element of $[n]$ not equal to $i$ or $j$. All choices of $k$ define the same product [51, Proposition V.3.4]. Further, off-diagonal Peirce spaces are isomorphic algebras under this product. (The $2^3 = 8$ factor appears because the Jordan product $\circ$ is applied three times.)

For our purposes, we only need to equip one Peirce space with a product, which we can choose arbitrarily. Following McCrimmon [91], we equip $\mathbf{J}_{12}$ and refer to it as the *coordinate algebra*. We also need a set of structured maps $W_{ij} : \mathbf{J} \to \mathbf{J}$ from Jacobson [73] induced by matrix units; see also [51, Chapter V]. We name these maps *ij-Peirce space transformations* for reasons that will become clear (Lemma 6.4.2).

**Definition 6.4.3.** *Let $U = \{u_{ij} \in \mathbf{J}_{ij}\}_{i,j=1}^n$ be Jordan matrix units for the set of Peirce spaces $\{\mathbf{J}_{ij}\}_{i,j=1}^n$.*

- *[91, Chapter 6]. The* coordinate algebra $\tilde{\mathbf{J}}_{12}$ *induced by $U$ is the Peirce space $\mathbf{J}_{12}$ equipped with product*

  $$x \times y := 8(u_{13} \circ x) \circ (u_{23} \circ y)$$

  *and identity $u_{12}$.*

- *For $i \neq j$, the $ij$*-Peirce-space-transformation $W_{ij} : \mathbf{J} \to \mathbf{J}$ *induced by $U$ is the map given by*

  $$W_{ij} := \begin{cases} x \mapsto x & \{i,j\} = \{1,2\} \\ x \mapsto 2u_{2k} \circ x & \{i,j\} = \{1,k\}, \ \ k \geq 3 \\ x \mapsto 2u_{1k} \circ x & \{i,j\} = \{2,k\}, \ \ k \geq 3 \\ x \mapsto 4u_{1i} \circ (u_{2j} \circ x) & \{i,j\} \cap \{1,2\} = \emptyset. \end{cases}$$

The norm $\| \cdot \|_{\mathbf{J}}$ is *multiplicative* with respect to coordinate-algebra multiplication, i.e., $\|x \times y\|_{\mathbf{J}} = \|x\|_{\mathbf{J}} \|y\|_{\mathbf{J}}$; see, e.g., [51, Proposition V.3.4]. This property implies coordinate-algebra isomorphisms are isometries (e.g., Lemma 6.8.3). The $ij$-Peirce-space-transformation is also an isometry with extremely special structure:

**Lemma 6.4.2** (Peirce space transformations)**.** *For $\{i,j\} \neq \{1,2\}$, let $W_{ij} : \mathbf{J} \to \mathbf{J}$ denote the ij-Peirce-space-transformation induced by a set of Jordan matrix units*

$\{u_{ij}\}_{i,j=1}^n$. *Then,*

$$W_{ij}(\mathbf{J}_{12}) = \mathbf{J}_{ij}, \qquad W_{ij}(\mathbf{J}_{ij}) = \mathbf{J}_{12}. \tag{6.1}$$

*Further, the restriction of* $W_{ij}$ *to* $\mathbf{J}_{12} \oplus \mathbf{J}_{ij}$ *is self-adjoint, isometric, and an involution, i.e., for any* $x, y \in \mathbf{J}_{12} \oplus \mathbf{J}_{ij}$*,*

- $\langle W_{ij}x, y\rangle = \langle x, W_{ij}y\rangle$;
- $\|W_{ij}x\|_{\mathbf{J}} = \|x\|_{\mathbf{J}}$;
- $W_{ij}W_{ij}x = x$.

*Finally,* $u_{ij} = W_{ij}u_{12}$ *and* $u_{12} = W_{ij}u_{ij}$.

*Proof.* We first show the self-adjoint property. That $W_{ij}$ is self-adjoint when $W_{ij} = 2u_{2j} \circ x$ or $W_{ij} = 2u_{1j} \circ x$ is immediate since, by definition, multiplication is a self-adjoint operation in a Euclidean Jordan algebra (Definition 1.6.1). When $W_{ij}$ equals $x \mapsto 4u_{1i} \circ (u_{2j} \circ x)$, it is the composition of self-adjoint maps that, by Lemma 6.8.1, commute on $\mathbf{J}_{12} \oplus \mathbf{J}_{ij}$. Hence, it is self-adjoint.

We next show the involution property. Consider the case $i = 1$ and $j > 3$ and let $x = x_{12} + x_{1j}$. Then,

$$W_{1j}W_{1j}x = 2u_{2j} \circ (2u_{2j} \circ x) \overset{1}{=} 2(u_{2j} \circ u_{2j}) \circ x \overset{2}{=} 2(u_{22} + u_{jj}) \circ x \overset{3}{=} x.$$

Here, the equality $\overset{1}{=}$ follows from Lemma 6.8.1, the equality $\overset{2}{=}$ from the definition of matrix units (Definition 6.4.2) and the equality $\overset{3}{=}$ from the Peirce Multiplication Rules (Lemma 6.1.1). The case where $W_{ij}$ equals $x \mapsto 2u_{1j} \circ x$ has identical proof. For the remaining case, consider distinct $p, q, r, s \in [n]$. For $x_{rs} \in \mathbf{J}_{rs}$, set $z_{pq} = u_{pr} \circ (u_{qs} \circ x_{rs})$ and note by the associativity identity (Lemma 6.8.1)

$$16u_{pr} \circ (u_{qs} \circ z_{pq}) = 16u_{qs} \circ (u_{pr} \circ z_{pq}). \tag{6.2}$$

Setting $w_{rq} = u_{qs} \circ x_{rs} \in \mathbf{J}_{rq}$, we have $z_{pq} = u_{pr} \circ w_{rq}$ and

$$u_{pr} \circ z_{pq} = u_{pr} \circ (u_{pr} \circ w_{rq}) = \frac{1}{2}(u_{pr} \circ u_{pr}) \circ w_{rq} = \frac{1}{2}(u_{rr} + u_{pp}) \circ w_{rq} = \frac{1}{4}w_{rq}.$$

Which shows

$$16u_{qs} \circ (u_{pr} \circ z_{pq}) = 4u_{qs} \circ w_{rq} = 4u_{qs} \circ (u_{qs} \circ x_{rs}) = 2(u_{qs} \circ u_{qs}) \circ x_{rs} = x_{rs}.$$

Taking $(p, q, r, s) = (i, j, 1, 2)$ shows that $W_{ij}W_{ij}x = x$ for all $x \in \mathbf{J}_{12}$. Taking $(p, q, r, s) = (1, 2, i, j)$ shows that $W_{ij}W_{ij}x = x$ for all $x \in \mathbf{J}_{ij}$.

The isometry property holds by combining the self-adjoint and involution property. Specifically, since

$$\langle W_{ij}x, W_{ij}x\rangle_{\mathbf{J}} = \langle W_{ij}W_{ij}x, x\rangle_{\mathbf{J}} = \langle x, x\rangle_{\mathbf{J}},$$

we must have that $\|x\|_{\mathbf{J}} = \|W_{ij}x\|_{\mathbf{J}}$ given that $\|z\|_{\mathbf{J}}^2 = k\langle z, z\rangle_{\mathbf{J}}$ for a constant independent of $z$. (See Definition 6.4.1.)

Finally, that $u_{ij} = W_{ij}u_{12}$ and $u_{12} = W_{ij}u_{ij}$ holds by the definitions of $W_{ij}$ and Jordan matrix units. For the relations (6.1), note the inclusion $\subseteq$ holds by the Peirce Multiplication rules; equality holds since $W_{ij}$ is an involution on $\mathbf{J}_{12} \oplus \mathbf{J}_{ij}$ and is hence surjective.

□

One can also express Jordan multiplication between elements of distinct off-diagonal Peirce spaces using coordinate algebra multiplication. Formally:

**Lemma 6.4.3.** *Let $W_{ij} : \mathbf{J} \to \mathbf{J}$ and $W_{jk} : \mathbf{J} \to \mathbf{J}$ be the $ij$- and $jk$-Peirce-space-transformations induced by a set of Jordan matrix units $U$. Then, for all $x, y \in \mathbf{J}$ and distinct $i, j, k \in [n]$,*

$$x_{ij} \circ y_{jk} = W_{ik}\left(W_{ij}x_{ij} \times W_{jk}y_{jk}\right)$$

*Proof.* We verify this directly:

$$\begin{aligned} x_{ij} \circ y_{jk} &= (W_{ij}W_{ij}x_{ij}) \circ (W_{jk}W_{jk}y_{jk}) \\ &= W_{ik}\left(W_{ij}x_{ij} \times W_{jk}y_{jk}\right). \end{aligned}$$

Here, the first line uses the fact $W_{jk}$ and $W_{ik}$ are involutions (Lemma 6.4.2) and the second is a direct application of [73, Lemma 6]. □

Combining results leads to an isomorphism $\Phi$ between Jordan algebras with rank at least three:

**Theorem 6.4.2** (Isomorphisms for rank $n \geq 3$)**.** *Suppose $\mathbf{J}^A$ and $\mathbf{J}^B$ are simple algebras of rank $n \geq 3$ with Jordan frames $\{e_i^A\}_{i=1}^n$ and $\{e_i^B\}_{i=1}^n$. For $K \in \{A, B\}$, let*

- *$U^K$ be Jordan matrix units for the Peirce spaces $\{\mathbf{J}_{ij}^K\}_{i,j=1}^n$ induced by $\{e_i^K\}_{i=1}^n$;*
- *$\tilde{\mathbf{J}}_{12}^K$ and $W_{ij}^K$ be the coordinate algebra and $ij$-Peirce-space-transformation induced by $U^K$;*
- *$T : \tilde{\mathbf{J}}_{12}^A \to \tilde{\mathbf{J}}_{12}^B$ denote an isomorphism between the coordinate algebras $\tilde{\mathbf{J}}_{12}^A$ and $\tilde{\mathbf{J}}_{12}^B$, i.e.,*

$$T(x \times y) = (Tx) \times (Ty).$$

*Then, the unique linear map $\Phi : \mathbf{J}^A \to \mathbf{J}^B$ satisfying*

- $\Phi e_i^A = e_i^B$ *for all* $i \in [n]$,
- $\Phi x = W_{ij}^B T W_{ij}^A$ *for all* $x \in \mathbf{J}_{ij}^A$ *with* $i \neq j$,

*is an isomorphism between* $\mathbf{J}^A$ *and* $\mathbf{J}^B$.

*Proof.* We verify $\Phi(x^2) = (\Phi x)^2$ by considering products between Peirce components seperately. For this, we let $i, j, k, l$ denote distinct integers.

**Cases** $x_{ij} \circ x_{kl}$**,** $x_{ii} \circ x_{kl}$**.** Under the assumption on $i, j, k, l$, we have that both $\Phi(x_{ij} \circ x_{kl})$ and $\Phi(x_{ij} \circ x_{kl})$ equal $\Phi(0)$ by the Peirce Multiplication Rules (Lemma 6.1.1). Further, $\Phi$ maps $\mathbf{J}_{ij}^A$, $\mathbf{J}_{ii}^A$, and $\mathbf{J}_{kl}^A$ into $\mathbf{J}_{ij}^B$, $\mathbf{J}_{ii}^B$, and $\mathbf{J}_{kl}^B$, respectively. Hence, the Peirce Multiplication Rules (Lemma 6.1.1) also show

$$\Phi(x_{ij}) \circ (\Phi x_{kl}) = 0, \qquad \Phi(x_{ii}) \circ (\Phi x_{kl}) = 0.$$

We conclude $\Phi(x_{ij}) \circ (\Phi x_{kl}) = \Phi(x_{ij} \circ x_{kl})$ and $\Phi(x_{ii}) \circ (\Phi x_{kl}) = \Phi(x_{ii} \circ x_{kl})$ as desired.

**Cases** $x_{ii} \circ x_{ii}$**,** $x_{ii} \circ x_{ij}$**,** $x_{ij} \circ x_{ij}$**.** We apply Theorem 6.4.1 to show the map $\tilde{\Phi} := \Phi_{ii} + \Phi_{jj} + \Phi_{ij}$ is a Jordan isomorphism between the subalgebras $\mathbf{J}_{ii}^A + \mathbf{J}_{ij}^A + \mathbf{J}_{jj}^A$ and $\mathbf{J}_{ii}^B + \mathbf{J}_{ij}^B + \mathbf{J}_{jj}^B$. By construction, $\tilde{\Phi}$ is an invertible transformation mapping $e_i^A$ to $e_i^B$ and $e_j^A$ to $e_j^B$. Further, by Lemmas 6.4.2 and 6.8.3 it is an isometry. Hence, it satisfies the hypothesis of Theorem 6.4.1 and is a Jordan isomorphism.

**Case** $x_{ij} \circ x_{jk}$**.** The prove the remaining case, we use the multiplication identity given by Lemma 6.4.3 and the fact $W_{ml}^K$ is an involution (Lemma 6.4.2); the line marked $(\star)$ uses the assumption $T$ is an isomorphism:

$$\begin{aligned}
\Phi_{ik}(x_{ij} \circ x_{jk}) &= W_{ik}^B T W_{ik}^A (x_{ij} \circ x_{jk}) \\
&= W_{ik}^B T W_{ik}^A W_{ik}^A \left( W_{ij}^A x_{ij} \times W_{jk}^A x_{jk} \right) && \text{(Lemma 6.4.3)} \\
&= W_{ik}^B T \left( W_{ij}^A x_{ij} \times W_{jk}^A x_{jk} \right) && \text{(Lemma 6.4.2)} \\
&= W_{ik}^B \left( (T W_{ij}^A x_{ij}) \times (T W_{jk}^A x_{jk}) \right) && (\star) \\
&= W_{ik}^B \left( (W_{ij}^B W_{ij}^B T W_{ij}^A x_{ij}) \times (W_{jk}^B W_{jk}^B T W_{jk}^A x_{jk}) \right) && \text{(Lemma 6.4.2)} \\
&= \left( W_{ij}^B T W_{ij}^A x_{ij} \right) \circ \left( W_{jk}^B T W_{jk}^A x_{jk} ) \right) && \text{(Lemma 6.4.3)} \\
&= (\Phi x_{ij}) \circ (\Phi x_{jk})
\end{aligned}$$

□

The isomorphism of Theorem 6.4.2 strongly relates to one used to prove Jacobson's *coordinization theorem* [76, Theorem 9.1]; see also [75] and [91, Chapter 6]. Specifically, Jacobson constructs an isomorphism between a rank-three algebra $\mathbf{J}^A$ and an algebra of matrices with elements in $\tilde{\mathbf{J}}_{12}$. His construction involves the maps $W_{ij}^A$. One can also use coordinate-algebra isomorphisms and $W_{ij}^A$ to prove the Jordan-von Neumann-Wigner classification of simple algebras (Proposition 1.6.1); see, e.g., [51, Chapter V] for such a proof.

### 6.4.3 Simplifications

**Rank two algebras** The description of isomorphisms between rank-two algebras (Theorem 6.4.1), though crucial for proving Theorem 6.4.2, is actually more complicated than necessary. The following simplification removes explicit reference to Jordan frames, using instead the identity elements of $\mathbf{J}^A$ and $\mathbf{J}^B$:

**Corollary 6.4.1.** *Let $\mathbf{J}^A$ and $\mathbf{J}^B$ be rank two algebras with identities $e^A$ and $e^B$. Let $\Phi : \mathbf{J}^A \to \mathbf{J}^B$ be an invertible map that is both an isometry and satisfies $\Phi e^A = e^B$. Then, $\Phi$ is a Jordan isomorphism.*

*Proof.* By Theorem 6.4.1, the claim follows if we can construct a Jordan frame for which $\Phi e_i^A = e_i^B$. Towards this, let $v^A$ be *any* element orthogonal to $e^A$ with $\|v_A\|_{\mathbf{J}^A} = \|e_A\|_{\mathbf{J}^A}$. Define $v^B = \Phi v^A$. Since $\Phi$ is an isometry, $v^B$ and $e^B$ are orthogonal and $\|v^B\|_{\mathbf{J}^B} = \|e^B\|_{\mathbf{J}^B}$. Further, $\Phi(e^A \pm v^A) = e^B \pm v^B$. We now show $\left\{\frac{1}{2}(e^A \pm v^A)\right\}$ and $\left\{\frac{1}{2}(e^B \pm v^B)\right\}$ are Jordan frames.

To begin, $\frac{1}{2}(e^A \pm v^A)$ are clearly pairwise orthogonal. We need to show they are also idempotent. Towards this, note

$$\frac{1}{2}(e^A + v^A) \circ \frac{1}{2}(e^A + v^A) = \frac{1}{4}(e^A + 2v^A + v^A \circ v^A).$$

Idempotency follows if $v^A \circ v^A = e^A$. In any Peirce decomposition, $v_{11}^A = -v_{22}^A$ since $v^A$ is orthogonal to $e^A$. Hence, $v^A \circ v^A = \lambda e^A$ by the Peirce Multiplication Rules and Lemma 6.4.1-(a). But

$$\langle e^A, e^A \rangle_{\mathbf{J}^A} = \langle v^A, v^A \rangle_{\mathbf{J}^A} = \langle v^A \circ v^A, e^A \rangle_{\mathbf{J}^A} = \langle \lambda e^A, e^A \rangle_{\mathbf{J}^A},$$

showing $\lambda = 1$ as desired. That $\frac{1}{2}(e^A - v^A)$ is idempotent has identical proof. It follows $\frac{1}{2}(e^A \pm v^A)$ is a Jordan frame; that $\frac{1}{2}(e^B \pm v^B)$ is a Jordan frame follows from an identical argument.

□

Note this corollary generalizes a widely-used transformation between the algebra $\mathbf{H}_2(\mathbb{R})$ of real symmetric matrices and the spin-factor algebra $\mathbb{R} \times \mathbb{R}^2$. Specifically, the map $\Phi$ satisfying

$$\Phi\left(\begin{bmatrix} a+b & c \\ c & a-b \end{bmatrix}\right) = (a,b,c)^T$$

maps the identity matrix to $(1,0,0)^T$ and is an isometry; hence, it is a Jordan isomorphism. As a consequence, $\Phi$ maps $\mathbb{S}^2_+$ onto the Lorentz cone $\mathcal{Q}^3$, a fact we exploited in Chapter 2.

**One-dimensional coordinate algebras** Recall isomorphisms between rank $n \geq 3$ algebras involve the composition of maps $W_{ij}TW_{ij}$, where $T$ was an isomorphism between coordinate algebras $\tilde{\mathbf{J}}^{A/B}_{12}$ (Theorem 6.4.2). This composition simplifies when the coordinate algebras are one-dimensional. Indeed, we can restate Theorem 6.4.2 using only matrix units:

**Corollary 6.4.2.** *Let $\mathbf{J}^A$ and $\mathbf{J}^B$ denote rank $n$, simple algebras. Let $\{u^A_{ij}\}^n_{i,j=1} \subset \mathbf{J}^A$ and $\{u^B_{ij}\}^n_{i,j=1} \subset \mathbf{J}^B$ denote Jordan matrix units. Finally, suppose the coordinate algebras $\tilde{\mathbf{J}}^A_{12}$ and $\tilde{\mathbf{J}}^B_{12}$ induced by these units are one-dimensional. Then, the linear map $\Phi : \mathbf{J}^A \to \mathbf{J}^B$ satisfying*

$$\Phi u^A_{ij} = u^B_{ij}$$

*is uniquely defined and invertible. Further, it is a Jordan isomorphism.*

*Proof.* Under this assumption, every Peirce space $\mathbf{J}^{A/B}_{ij}$ is spanned by $u^{A/B}_{ij}$, giving uniqueness. Letting $T$ denote $u^A_{12} \mapsto u^B_{12}$ (extended via linearity) gives a coordinate-algebra isomorphism. By Lemma 6.4.2, $W^A_{ij}u^A_{ij} = u^A_{12}$ and $W^B_{ij}u^B_{12} = u^B_{ij}$. Hence, $u^B_{ij} = W^B_{ij}TW^A_{ij}u^A_{ij}$, showing $\Phi$ is a Jordan isomorphism (Theorem 6.4.2). □

Recall $\tilde{\mathbf{J}}^{A/B}_{12}$ is one-dimensional precisely when $\mathbf{J}^{A/B}$ is isomorphic to $\mathbf{H}_n(\mathbb{R})$, the Jordan algebra of real symmetric matrices of order $n$. Such algebras arise frequently in the preprocessing of a semidefinite optimization problems; see, e.g., [108, Section 6].

## ■ 6.5 Subroutines

Isomorphism between simple algebras of rank $n \geq 3$ require a set of matrix units and a coordinate-algebra isomorphism. We now give subroutines that find these objects. These routines assume Peirce decompositions of each ideal (induced by a Jordan frame), which can be obtained from the simple partition (Theorem 6.3.1).

### ■ 6.5.1 Finding Jordan matrix units

Suppose $\mathbf{J}$ is a simple Euclidean Jordan algebra. Recall Jordan matrix units (Definition 6.4.2) for $\mathbf{J}$ consist of a Jordan frame $\{u_{ii}\}^n_{i=1}$ and $u_{ij}$ from each off-diagonal Peirce space $\mathbf{J}_{ij}$ satisfying

$$u_{ij}^2 = u_{ii} + u_{jj} \ \ \forall i \neq j \quad \text{(normalization)}, \qquad u_{ij} \circ u_{jk} = \frac{1}{2} u_{ik} \quad \text{(orientation)}.$$

We will give a procedure for generating $u_{ij}$ satisfying these conditions. The input will be $x \in \mathbf{J}$ with structured *off-diagonal Peirce-support*:

**Definition 6.5.1** (Off-diagonal Peirce-support). *Let $\{\mathbf{J}_{ij}\}_{i,j=1}^n$ be the Peirce spaces induced by a Jordan frame of a simple algebra $\mathbf{J}$. The* off-diagonal Peirce-support *of $x \in \mathbf{J}$ is the undirected graph $G = ([n], E)$ for which $\{i, j\} \in E$ iff $x_{ij}$ (the $ij$-th Peirce component of $x$) is nonzero (recalling $x_{ij} = x_{ji}$).*

Suppose $x \in \mathbf{J}$ has off-diagonal Peirce-support $G$ equal to a *tree*—i.e., a connected graph with no cycles. Each pair of nodes is uniquely connected by a path. If one labels each edge in $G$ by its corresponding sample $\frac{1}{\|x_{ij}\|_{\mathbf{J}}} x_{ij}$, each path in $G$ induces a product of samples. The following shows the set of such products equals a set of matrix units.

**Theorem 6.5.1** (Tree-induced matrix units). *Let $\{e_i\}_{i=1}^n$ be a Jordan frame for a rank $n$ simple algebra $\mathbf{J}$. Let $\{\mathbf{J}_{ij}\}_{i,j=1}^n$ be the set of Peirce spaces induced by $\{e_i\}_{i=1}^n$. Suppose the off-diagonal Peirce-support of $x \in \mathbf{J}$ equals a tree $G = ([n], E)$. For each node $i \in [n]$ and edge $\{i, j\} \in E$, define*

$$u_{ii} := e_i, \qquad u_{ij} := \frac{x_{ij}}{\|x_{ij}\|_{\mathbf{J}}}, \qquad u_{ji} := u_{ij}.$$

*For each edge $\{i, j\} \in E$ define the linear maps $U_{ij} : \mathbf{J} \to \mathbf{J}$ and $U_{ji} : \mathbf{J} \to \mathbf{J}$ via*

$$U_{ij} x = 2 u_{ij} \circ x, \qquad U_{ji} x = 2 u_{ji} \circ x.$$

*Finally, if $\{i, j\} \notin E$, let $p_1, p_2, \dots, p_m \in [n] \times [n]$ denote the unique path from $i$ to $j$ and define*

$$u_{ij} := U_{p_m} U_{p_{m-1}} \cdots U_{p_2} u_{p_1}, \qquad u_{ji} := U_{p_1} U_{p_2} \cdots U_{p_{m-1}} u_{p_m}. \tag{6.3}$$

*Then, $u_{ij} \in \mathbf{J}_{ij}$. Further, $\{u_{ij}\}_{i,j=1}^n$ is a set of Jordan matrix units, i.e., for all distinct $i, j, k \in [n]$:*

1. $u_{ii} = e_i$
2. $u_{ij} = u_{ji}$
3. $u_{ij} \circ u_{jk} = \frac{1}{2} u_{ik}$
4. $u_{ij} \circ u_{ij} = u_{ii} + u_{jj}$

*Proof.* That $u_{ij} \in \mathbf{J}_{ij}$ follows by the Peirce Multiplication Rules (Lemma 6.1.1). To show the $u_{ij}$ form a set of matrix units, we verify Statements 1-4. Statement 1 is immediate from definition of $u_{ii}$. For Statement 2, we note $u_{ij} = u_{ji}$ holds by definition if $\{i, j\}$ is an edge. Otherwise, letting $p_1, \ldots, p_m$ denote the unique path from $i$ to $j$ and using the definition (6.3) shows

$$\underbrace{U_{p_m} U_{p_{m-1}} \cdots U_{p_2} u_{p_1}}_{u_{ij}} \stackrel{a}{=} L_{u_{p_1}} U_{p_2} \cdots U_{p_{m-1}} 2u_{p_m} \stackrel{b}{=} \underbrace{U_{p_1} U_{p_2} \cdots U_{p_{m-1}} u_{p_m}}_{u_{ji}},$$

where $L_{u_{p_1}}$ denotes $x \mapsto u_{p_1} \circ x$. Here, the equality $\stackrel{a}{=}$ follows from Lemma 6.8.2 and $\stackrel{b}{=}$ by replacing $2u_{p_m}$ and $L_{u_{p_1}}$ with $u_{p_m}$ and $U_{p_1}$.

For Statement 3, suppose both $\{i, j\}$ and $\{j, k\}$ are non-edges and let $p_1, \ldots, p_m$ and $q_1, \ldots, q_\ell$ denote the unique paths from $i$ to $j$ and from $j$ to $k$, respectively. Using (6.3), we have

$$u_{jk} \circ u_{ij} = (U_{q_\ell} U_{q_{\ell-1}} \cdots U_{q_2} u_{q_1}) \circ (U_{p_m} U_{p_{m-1}} \cdots U_{p_2} u_{p_1}).$$

We can now iteratively pull out each $U_{q_r}$ (for $r = \ell, \ell-1, \cdots, 1$) term by using the identity $(x \circ y) \circ z = x \circ (y \circ z)$ with $x = U_{q_r}$ and $z = U_{p_m} U_{p_{m-1}} \cdots U_{p_2} u_{p_1}$:

$$\begin{aligned}
u_{ij} \circ u_{jk} &= (U_{q_\ell} U_{q_{\ell-1}} U_{q_{\ell-2}} \cdots U_{q_2} u_{q_1}) \circ (U_{p_m} U_{p_{m-1}} \cdots U_{p_2} u_{p_1}) \\
&= U_{q_\ell} \left( (U_{q_{\ell-1}} U_{q_{\ell-2}} \cdots U_{q_2} u_{q_1}) \circ (U_{p_m} U_{p_{m-1}} \cdots U_{p_2} u_{p_1}) \right) \\
&= U_{q_\ell} U_{q_{\ell-1}} \left( (U_{q_{\ell-2}} \cdots U_{q_2} u_{q_1}) \circ (U_{p_m} U_{p_{m-1}} \cdots U_{p_2} u_{p_1}) \right) \\
&\;\;\vdots \\
&= U_{q_\ell} U_{q_{\ell-1}} U_{q_{\ell-2}} \cdots U_{q_2} \left( u_{q_1} \circ (U_{p_m} U_{p_{m-1}} \cdots U_{p_2} u_{p_1}) \right) \\
&= U_{q_\ell} U_{q_{\ell-1}} U_{q_{\ell-2}} \cdots U_{q_2} \frac{1}{2} U_{q_1} U_{p_m} U_{p_{m-1}} \cdots U_{p_2} u_{p_1} \\
&= \frac{1}{2} u_{ik},
\end{aligned}$$

where the last line uses (6.3) and the fact $p_1, \ldots, p_m, q_1, \ldots, q_\ell$ is the unique path from $i$ to $k$. The case where only $\{i, j\}$ is an edge has essentially identical proof. Finally, if both $\{i, j\}$ and $\{j, k\}$ are edges, then $\{i, k\}$ is a non-edge—otherwise the tree has a cycle. Hence, applying (6.3) to the path $(i, j), (j, k)$ shows $u_{ik} = U_{jk} u_{ij} = 2u_{jk} \circ u_{ij}$.

We only need to prove Statement 4 if $\{i, j\}$ is a non-edge since it otherwise holds by assumption. Let $p_1, p_2 \ldots, p_m$ denote the path from $i$ to $j$. For $1 \le \ell \le m$, define $z_\ell \in \mathbf{J}_{k_\ell i}$ via

$$z_\ell := \begin{cases} u_{p_1} & \ell = 1, \\ U_{p_\ell} U_{p_{\ell-1}} \cdots U_{p_2} u_{p_1} & 2 \le \ell \le m. \end{cases}$$

Using induction, we will show $z_\ell^2 = u_{k_\ell k_\ell} + u_{ii}$, noting the base case holds by assumption. To ease notation let $(r, s)$ equal $p_{\ell+1}$ (which implies $k_\ell = r$ and $k_{\ell+1} = s$). Since $z_{\ell+1} \in \mathbf{J}_{si}$ and since $z_{\ell+1} = 2u_{rs} \circ z_\ell$,

$$(2u_{rs} \circ z_\ell)^2 = \lambda(u_{ii} + u_{ss})$$

for some $\lambda$. We will show $\lambda = 1$ to prove Statement 4. Since $u_{ii}$ and $u_{ss}$ are orthogonal,

$$\begin{aligned}
\lambda\langle u_{ii}, u_{ii}\rangle &= \langle \lambda(u_{ii} + u_{ss}), u_{ii}\rangle \\
&= \langle (2u_{rs} \circ z_\ell)^2, u_{ii}\rangle \\
&= \langle 2u_{rs} \circ z_\ell, (2u_{rs} \circ z_\ell) \circ u_{ii}\rangle \\
&= \langle 2u_{rs} \circ z_\ell, u_{rs} \circ z_\ell\rangle \\
&= \langle z_\ell, (2u_{rs} \circ (u_{rs} \circ z_\ell))\rangle \\
&= \langle z_\ell, (u_{rs} \circ u_{rs}) \circ z_\ell\rangle \\
&= \langle z_\ell, (u_{rr} + u_{ss}) \circ z_\ell\rangle \\
&= \frac{1}{2}\langle z_\ell, z_\ell\rangle.
\end{aligned}$$

But under the inductive hypothesis that $z_\ell^2 = u_{ii} + u_{rr}$, it follows that

$$\langle z_\ell, z_\ell\rangle = \langle z_\ell^2, e\rangle = \langle u_{ii} + u_{rr}, e\rangle = \langle u_{ii}, e\rangle + \langle u_{rr}, e\rangle = \langle u_{ii}, u_{ii}\rangle + \langle u_{rr}, u_{rr}\rangle = 2\langle u_{ii}, u_{ii}\rangle,$$

where the last equality follows given that primitive idempotents have the same norm by Lemma 6.4.1-(b). Hence, $\lambda\langle u_{ii}, u_{ii}\rangle = \langle u_{ii}, u_{ii}\rangle$, showing $\lambda = 1$ as desired.

□

Theorem 6.5.1 suggests an obvious procedure for constructing matrix units that involves enumerating all paths in tree. It turns out we can state an alternative procedure (Algorithm 6.2) that avoids enumeration of these paths. It arises from the following observation:

**Lemma 6.5.1.** *Let $U \subset \mathbf{J}$ be a set of Jordan matrix units and let $\mathbf{cone}\, U := \{\lambda u : u \in U, \lambda \geq 0\}$ denote the cone generated by $U$. If $x, y \in \mathbf{cone}\, U$, then $x \circ y \in \mathbf{cone}\, U$.*

*Proof.* Suppose $x, y \in \mathbf{cone}\, U$. Then, by definition, $x = \sum_{u \in U} \lambda_u u$ and $y = \sum_{u \in U} \beta_u u$ for some $\lambda_u \geq 0$ and $\beta_u \geq 0$. Hence,

$$x \circ y = \sum_{u \in U} \sum_{w \in U} \lambda_u \beta_w u \circ w.$$

From the definition of matrix units, $u \circ w \in \mathbf{cone}\, U$ if $u, w \in U$. Hence, $x \circ y \in \mathbf{cone}\, U$. □

To see the algorithmic implications of Lemma 6.5.1, consider $x \in \mathbf{J}$ with $x_{ii} \geq 0$ and off-diagonal Peirce-support equal to a tree $G$. Then, $x \in \mathbf{cone}\, U$ for the set of matrix units $U$ of Theorem 6.5.1. By Lemma 6.5.1, the Peirce components of $y = x + x^2 + x^3 + \cdots + x^m$ are also in the cone generated by $U$ for all $m$. Further, if $m$ is sufficient large, each Peirce component is nonzero (since $G$ is a tree). It follows we can recover the full set of matrix units by normalizing the Peirce components of $y$. Example 6.5.1 illustrates this procedure and Algorithm 6.2 states it formally. Note we never need to take $m$ larger than $n$ by the Peirce Multiplication Rules.

**Example 6.5.1.** *Let $\mathbf{J} = \mathbf{H}_4(\mathbb{R})$, the Jordan algebra of symmetric matrices of order four with product $x \circ y = \frac{1}{2}(xy + yx)$. For the Peirce decomposition induced by the Jordan frame $\{e_i\}_{i=1}^4$, given by*

$$e_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad e_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad e_4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

*the following matrix $x \in \mathbf{J}$ has off-diagonal Peirce-support equal to a tree:*

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

*Letting $y = x + x^2 + x^3$ and $u \in \mathbf{J}$ satisfy $u_{ij} = u_{ij} \frac{1}{\|y_{ij}\|_{\mathbf{J}}}$ and $u_{ii} = \frac{1}{\sqrt{2}} \frac{y_{ii}}{\|y_{ii}\|_{\mathbf{J}}}$ gives*

$$y = \begin{bmatrix} 1 & -1 & -1 & 3 \\ -1 & 2 & 3 & -4 \\ -1 & 3 & 1 & -1 \\ 3 & -4 & -1 & 2 \end{bmatrix}, \qquad u = \begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

*The Peirce components of $u$ form a set of Jordan matrix units; the diagonal components*

**Algorithm 6.2:** Finds matrix units for Peirce decomposition $\mathbf{J} = \bigoplus_{1\le i\le j\le n} \mathbf{J}_{ij}$

**Input:** $x \in \mathbf{J}$ with $x_{ii} \geq 0$ and off-diagonal Peirce-support equal to a tree.
**Output:** A set of Jordan matrix units $U = \{u_{ij} \in \mathbf{J}_{ij}\}_{i,j=1}^n$ with $x \in \mathbf{cone}\, U$.

```
begin
    y ← 0, x̂ ← x
    repeat
        y ← y + x̂
        x̂ ← x̂ ∘ x
    until all y_ij ≠ 0;
    for each (i, j) ∈ [n] × [n] do
        if i = j, rescale so u_ii is idempotent: u_ii ← (1/√2) y_ii/‖y_ii‖_J
        if i ≠ j, rescale so u_ij^2 is idempotent: u_ij ← y_ij/‖y_ij‖_J
    end
    return {u_ij ∈ J_ij}_{i,j=1}^n
end
```

*satisfy* $u_{ii} = e_i$ *and the (distinct) off-diagonal components (recall* $u_{ij} = u_{ji}$*) are*

$$u_{12} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad u_{13} = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad u_{14} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

$$u_{24} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \qquad u_{34} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{bmatrix}, \qquad u_{23} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

## 6.5.2 Constructing coordinate-algebra isomorphisms

Recall the Jordan isomorphism of Theorem 6.4.2 is defined in part by a coordinate-algebra isomorphism $T$. It turns out we can iteratively construct $T$ using an idea of Jacobson [74, Section 3]. This idea exploits the fact coordinate algebras are *Euclidean Hurwitz algebras*:

**Definition 6.5.2** (e.g., Chapter V of [51])**.** *An algebra* $\mathbf{E}$ *over* $\mathbb{R}$ *(with product denoted* $\times$*) is a* Euclidean Hurwitz algebra *if there exists an identity* $u$ *and an inner product* $\langle\cdot,\cdot\rangle_{\mathbf{E}}$ *for which the norm* $\|x\|_{\mathbf{E}} := \sqrt{\langle x, x\rangle_{\mathbf{E}}}$ *is multiplicative, i.e.,*

$$\|x \times y\|_{\mathbf{E}} = \|x\|_{\mathbf{E}}\|y\|_{\mathbf{E}} \qquad \forall x, y \in \mathbf{E}.$$

**Lemma 6.5.2** (e.g., Proposition V.3.4 of [51]). *Let $\mathbf{J}$ be a simple, Euclidean Jordan algebra of rank $n \geq 3$ and let $\tilde{\mathbf{J}}_{12}$ be the coordinate algebra induced by a set of Jordan matrix units. Then, $\tilde{\mathbf{J}}_{12}$ is a Euclidean Hurwitz algebra. Specifically,*

$$\|x \times y\|_{\mathbf{J}} = \|x\|_{\mathbf{J}} \|y\|_{\mathbf{J}}$$

*for the norm $\|x\|_{\mathbf{J}}$ of Definition 6.4.1.*

We exploit a key property of Euclidean Hurwitz algebras: they admit construction of an ascending chain of subalgebras via the *Cayley-Dickson construction*. Specifically, one can iteratively extend any proper subalgebra (containing the identity) to a subalgebra of twice the dimension:

**Lemma 6.5.3** (Proposition V.1.4 of [51]). *Let $\mathbf{E}$ be a Euclidean Hurwitz algebra with identity $u$. Let $S \subsetneq \mathbf{E}$ be a proper subalgebra containing $u$, and let $S^{\perp} := \{x \in \mathbf{E} : \langle x, y \rangle_{\mathbf{E}} = 0 \ \ \forall y \in S\}$ denote its orthogonal complement. For $j \in S^{\perp}$ with unit norm (i.e., $\|j\|_{\mathbf{E}} = 1$), define the subspace $S \times j := \{x \times j : x \in S\}$. The following statements hold:*

*(a) The orthogonal complement $S^{\perp}$ contains $S \times j$; hence, $S \oplus (S \times j)$ is a direct-sum of subspaces.*

*(b) The subspace $S \oplus (S \times j)$ is a subalgebra of $\mathbf{E}$, i.e., if $x, y \in S \oplus (S \times j)$, then $x \times y \in S \oplus (S \times j)$.*

*(c) For all $w, x, y, z \in S$, the following holds*

$$(w + x \times j) \times (y + z \times j) = (w \times y - z^* \times x) + (x \times y^* + z \times w) \times j.$$

*where $t^* := 2\langle t, u \rangle_{\mathbf{E}} u - t$.*

As sketched in Jacobson [74, Section 3], an isomorphism between proper subalgebras can also be extended via this construction. Iterating this procedure yields an isomorphism between the full algebras. For completeness, we state these iterations explicitly (Algorithm 6.3). Correctness relies on the following theorem:

**Theorem 6.5.2** (Isomorphism Extensions). *Let $\mathbf{E}^A$ and $\mathbf{E}^B$ be Euclidean Hurtwitz Algebras and let $S^A \subseteq \mathbf{E}^A$ and $S^B \subseteq \mathbf{E}^B$ be proper subalgebras containing the identity elements of $\mathbf{E}^A$ and $\mathbf{E}^B$, respectively. Let $T : S^A \to S^B$ be an isomorphism between $S^A$ and $S^B$, i.e., an invertible linear map satisfying*

$$T(x \times y) = T(x) \times T(y) \ \ \forall x, y \in S^A.$$

*Finally, for* $j^A \in (S^A)^\perp$ *and* $j^B \in (S^B)^\perp$ *of unit norm, define* $U : (S^A \times j^A) \to (S^B \times j^B)$ *via*

$$U(x \times j^A) := (Tx) \times j^B \qquad \forall x \in S^A.$$

*Then, the extension* $T \oplus U$ *is an isomorphism between* $S^A \oplus (S^A \times j^A)$ *and* $S^B \oplus (S^B \times j^B)$*.*

*Proof.* Let $\tilde{T} = T \oplus U$ and consider $p = w + x \times j^A$ and $q = y + z \times j^A$ for $w, x, y, z \in S^A$.

Applying the identity Lemma 6.5.3-(c) to $p \times q$, using the definition of $\tilde{T}$, and using the fact $T$ is an isomorphism gives:

$$\begin{aligned} \tilde{T}(p \times q) &= \tilde{T}\left((w \times y) - z^* \times x + (x \times y^* + z \times w) \times j^A\right) \\ &= T\left(w \times y - z^* \times x\right) + \left(T\left(x \times y^* + z \times w\right)\right) \times j^B \\ &= (Tw \times Ty) - (Tz^* \times Tx) + (Tx \times Ty^* + Tz \times Tw) \times j^B \end{aligned} \tag{6.4}$$

By definition of $\tilde{T}$ and Lemma 6.5.3-(c),

$$\begin{aligned} \tilde{T}p \times \tilde{T}q &= \left(Tw + (Tx) \times j^B\right) \times \left(Ty + (Tz) \times j^B\right) \\ &= Tw \times Ty - (Tz)^* \times Tx + (Tx \times (Ty)^* + Tz \times Tw) \times j^B \end{aligned} \tag{6.5}$$

Since $T$ is an isomorphism, $(Tx)^* = Tx^*$; see [74, Section 3]. Hence, the expressions (6.4) and (6.5) are equal, showing $\tilde{T}(p \times q) = \tilde{T}p \times \tilde{T}q$ as desired. □

Note up to isomorphism, the only Euclidean Hurwitz algebras are the real numbers, the complex numbers, the quaternions and the octonions. The number of 'while' loop iterations of Algorithm 6.3 identifies the isomorphism class. Specifically, zero iterations occur if $\mathbf{E}^A$ and $\mathbf{E}^B$ are isomorphic to the real numbers, one iteration occurs if they are isomorphic to the complex numbers, two iterations if the quaternions, and three iterations if the octonions.

## 6.6 Computational considerations and numerical examples

We next discuss ways of optimizing presented algorithms for computational efficiency. (Chapter 7 gives numerical examples.) To begin, when $\mathbf{J}^A$ and $\mathbf{J}^B$ are matrix algebras—i.e., subalgebras of $\mathbf{H}_n(\mathbb{T})$ where $\mathbb{T}$ denotes the real numbers, complexes, quaternions or octonions—the following optimizations are possible:

- Efficient storage of idempotents via factorization;
- Finding Peirce components via congruence transformation;
- Refining non-primitive idempotents by computing in a lower-rank algebra;

**Algorithm 6.3:** Finds isomorphism between Euclidean Hurwitz algebras (e.g., coordinate algebras)

**Inputs:** Isomorphic algebras $\mathbf{E}^A$ and $\mathbf{E}^B$ with identities $u^A$ and $u^B$.
**Output:** An isomorphism $T : \mathbf{E}^A \to \mathbf{E}^B$ between $\mathbf{E}^A$ and $\mathbf{E}^B$.
**begin**
  Set $S^A = \text{span}\{u^A\}$, $S^B = \text{span}\{u^B\}$
  Set $T : S^A \to S^B$ to the linear map satisfying $Tu^A = u^B$.
  **while** $S^A \neq \mathbf{E}^A$ **do**

1. Pick unit vectors $j^A \in (S^A)^\perp$ and $j^B \in (S^B)^\perp$
2. Define the map $U : (S^A \times j^A) \to (S^B \times j^B)$ via
$$U(x \times j^A) = (Tx) \times j^B \qquad \forall x \in S^A$$
3. Update $S^A$, $S^B$ and extend map $T$ via
$$\begin{aligned} S^A &\leftarrow S^A \oplus (S^A \times j^A), \\ S^B &\leftarrow S^B \oplus (S^B \times j^B), \\ T &\leftarrow T \oplus U \end{aligned}$$

  **end**
**end**

- Finding sparse isomorphisms using diagonal idempotents.

We overview each of these ideas and note they are straightforward observations from known results. Indeed, refining idempotents in associative algebras is proposed in [49, Section 3.2]. Sparse isomorphisms arise by exploiting sparsity of Peirce decompositions induced by diagonal idempotents; analogous Peirce spaces of matrix *-algebras (called *cells*) are studied extensively in [142]. For simplicity, we focus only on algebras of real symmetric matrices (i.e., subalgebras of $\mathbf{H}_n(\mathbb{R})$), which are also the most practically-relevant.

**Factorized idempotents** Any idempotent $e_i$ of $\mathbf{H}_n(\mathbb{R})$ is simply a projection matrix. Hence, it has a factorization $e_i = q_i q_i^T$ for $q_i \in \mathbb{R}^{n \times r_i}$ with orthonormal columns. If $\{e_i\}_{i=1}^m$ is a Jordan frame (of a subalgebra), the corresponding set of factors $\{q_i\}_{i=1}^m$ are pairwise orthogonal—i.e., $q_i^T q_j = 0_{r_i \times r_j}$ for $i \neq j$. It follows we can store a Jordan frame by storing the matrix $q = (q_1, q_2, \ldots, q_m) \in \mathbb{R}^{n \times p}$ (where $p \leq n$) and the block-partition of its columns, i.e., the partition $P_1, \ldots P_m$ of $[p]$ for which $j \in P_i$ if the $j^{th}$ column of $q$ is a column of $q_i$. This allows one to store a general Jordan frame using no more than $n^2$ numbers as opposed to $mn^2$.

**Peirce components via congruence transformation** The matrix $q$ and its column partition $P_1, \ldots, P_m$ also admit compact formulas for Peirce components. Letting $\star$ denote Hadamard (i.e., entrywise) multiplication, we have

$$x_{ij} = q\left(J^{ij} \star (q^T x q)\right) q^T,$$

where $J^{ij}$ is the 0/1 symmetric matrix induced by the partition classes of $i$ and $j$:

$$[J^{ij}]_{uv} = \begin{cases} 1 & (u,v) \in (P_i \times P_j) \cup (P_j \times P_i). \\ 0 & \text{otherwise.} \end{cases}$$

This allows identification of *all* Peirce components by a single congruence transformation $q^T x q$ and application of different (non-overlapping) sparsity masks $J^{ij}$. The (equivalent) formula for off-diagonal Peirce components (Section 6.1)—given by $x_{ij} = 2e_i \circ (x \circ e_j) + 2(e_i \circ x) \circ e_j$—suggests an arguably more complicated procedure unlikely to benefit (as much) from highly-optimized matrix-multiplication libraries.

**Refining idempotents** Finding a Jordan frame is crucial for all computation described in this paper. Often non-primitive idempotents are easily available (e.g., from a basis element or from the spectral decomposition of a generator for $\mathbf{J}$.) How, then, can these idempotents be refined into sums of primitive idempotents to yield a Jordan frame?

First observe we can refine a non-primitive idempotent $e_i$ by working in the subalgebra $\mathbf{J}_{ii} = \{x \in \mathbf{J} : x \circ e_i = x\}$—an observation made for associative algebras in [49, Section 3.2]. Indeed, if $\{f_j\}_{j=1}^r$ is a Jordan frame for $\mathbf{J}_{ii}$, then $e_i = \sum_{j=1}^r f_j$, yielding a refinement. Next note if $e_i = q_i q_i^T$ for $q_i \in \mathbb{R}^{n \times r_i}$, then $\mathbf{J}_{ii}$ is isomorphic to

$$\hat{\mathbf{J}}_{ii} := \left\{ q_i^T x q_i : x \in \mathbf{J} \right\},$$

a subalgebra of $\mathbf{H}_{r_i}(\mathbb{R})$ with $(r_i \leq n)$. This allows one to find a Jordan frame $F$ for $\hat{\mathbf{J}}_{ii}$ doing computation with smaller matrices. One can then construct a Jordan frame $\{q_i x q_i^T : x \in F\}$ for $\mathbf{J}_{ii}$—yielding the desired refinement of $e_i$.

**Diagonal idempotents and sparsity** Suppose a subalgebra of $\mathbf{H}_n(\mathbb{R})$ contains a set of idempotents that are *diagonal* matrices. Such diagonal idempotents induce sparse Peirce spaces (e.g., Example 6.1.1), which enables construction of sparse Jordan frames, sparse matrix units and, ultimately, sparse isomorphisms. The next example illustrates this tremendous utility. We then discuss when diagonal idempotents arise.

**Example 6.6.1** (Diagonal idempotents)**.** *The algebra* $\mathbf{J}$ *of symmetric matrices*

$$\mathbf{J} = \left\{ \begin{pmatrix} t_1 & t_5 & t_6 & t_6 \\ t_5 & t_2 & t_4 & t_4 \\ t_6 & t_4 & t_3 & t_7 \\ t_6 & t_4 & t_7 & t_3 \end{pmatrix} : t \in \mathbb{R}^7 \right\}$$

*contains a set of diagonal idempotents* $\{e_1, e_2, \hat{e}_3\}$ *summing to the identity* $e$*:*

$$e_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} e_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \hat{e}_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

*The set of idempotents* $\{e_1, e_2, \hat{e}_3\}$ *induces sparse Peirce spaces (partially-listed):*

$$\mathbf{J}_{12} = \begin{bmatrix} 0 & t_5 & 0 & 0 \\ t_5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \mathbf{J}_{23} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & t_4 & t_4 \\ 0 & t_4 & 0 & 0 \\ 0 & t_4 & 0 & 0 \end{bmatrix}, \mathbf{J}_{13} = \begin{bmatrix} 0 & 0 & t_6 & t_6 \\ 0 & 0 & 0 & 0 \\ t_6 & 0 & 0 & 0 \\ t_6 & 0 & 0 & 0 \end{bmatrix}, \mathbf{J}_{33} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & t_3 & t_7 \\ 0 & 0 & t_7 & t_3 \end{bmatrix}.$$

*The following primitive idempotents* $e_3, e_4 \in \mathbf{J}_{33}$ *refine* $\hat{e}_3$

$$e_3 = \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \qquad e_4 = \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix},$$

*yielding a sparse Jordan frame* $\{e_1, e_2, e_3, e_4\}$*. The sets* $\{e_1, e_2, e_3\}$ *and* $\{e_4\}$ *form the* *simple partition of this frame. Hence, $\{e_1, e_2, e_3\}$ is a Jordan frame for a simple ideal. It is completed to a set of sparse matrix units by*

$$u_{12} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad u_{23} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad u_{13} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

*Each ideal of $\mathbf{J}$ is isomorphic to an algebra of real symmetric matrices. We can construct an isomorphism $\Phi : \mathbf{H}_3(\mathbb{R}) \oplus \mathbb{R} \to \mathbf{J}$ by mapping matrix units to matrix units (Corollary 6.4.2):*

$$\Phi\left( \begin{bmatrix} a & d & e \\ d & b & f \\ e & f & c \end{bmatrix} \oplus g \right) = (ae_1 + be_2 + ce_3 + du_{12} + eu_{13} + fu_{23}) + ge_4, \tag{6.6}$$

*which upon substitution gives*

$$\Phi\left( \begin{bmatrix} a & d & e \\ d & b & f \\ e & f & c \end{bmatrix} \oplus g \right) = \begin{bmatrix} a & d & \frac{e}{\sqrt{2}} & \frac{e}{\sqrt{2}} \\ d & b & \frac{f}{\sqrt{2}} & \frac{f}{\sqrt{2}} \\ \frac{e}{\sqrt{2}} & \frac{f}{\sqrt{2}} & \frac{c+g}{2} & \frac{c-g}{2} \\ \frac{e}{\sqrt{2}} & \frac{f}{\sqrt{2}} & \frac{c-g}{2} & \frac{c+g}{2} \end{bmatrix}.$$

*Clearly a matrix representation of $\Phi$ in the bases $(a, b, c, d, e, f, g)^T$ and $(t_1, \ldots, t_7)^T$ is sparse.*

Note sparsity of the isomorphism $\Phi$ is not guaranteed simply by existence of diagonal idempotents—one must recognize and use these idempotents explicitly. We illustrate this by constructing a different isomorphism using a Jordan frame not obtained from the diagonal idempotents:

**Example 6.6.1** (Continued)**.** *Consider the matrix $q = (q_1, q_2, q_3)$ and $\tilde{e}_i \in \mathbf{J}$ and $\tilde{u}_{ij} \in \mathbf{J}$ defined via*

$$q = \frac{1}{11} \begin{bmatrix} 6 & 7 & 6 \\ 2 & 6 & -9 \\ \frac{9}{\sqrt{2}} & \frac{-6}{\sqrt{2}} & \frac{-6}{\sqrt{2}} \\ \frac{9}{\sqrt{2}} & \frac{-6}{\sqrt{2}} & \frac{-6}{\sqrt{2}} \end{bmatrix} \qquad \tilde{e}_i = q_i q_i^T, \qquad \tilde{u}_{ij} = q_i q_j^T + q_j q_i^T.$$

*One can confirm $\{\tilde{e}_i\}_{i=1}^3 \cup \{e_4\}$ is a Jordan frame for $\mathbf{J}$ and that $\tilde{u}_{ij}$ completes $\{\tilde{e}_i\}_{i=1}^3$ to a set of matrix units. Replacing $\{e_i\}_{i=1}^3$ with $\{\tilde{e}_i\}_{i=1}^3$ and $u_{ij}$ with $\tilde{u}_{ij}$ in* (6.6) *yields a different isomorphism $\tilde{\Phi}$ that is not sparse. For example, $x := (0, 0, 0, d, 0, 0, 0)^T$ and*

$y := (0,0,0,0,f,0,0)^T$ *map to dense matrices* $\tilde{u}_{12}$ *and* $\tilde{u}_{23}$*:*

$$\tilde{\Phi}x = \frac{d}{121}\begin{bmatrix} 84 & 50 & \frac{27}{\sqrt{2}} & \frac{27}{\sqrt{2}} \\ 50 & 24 & \frac{42}{\sqrt{2}} & \frac{42}{\sqrt{2}} \\ \frac{27}{\sqrt{2}} & \frac{42}{\sqrt{2}} & -54 & -54 \\ \frac{27}{\sqrt{2}} & \frac{42}{\sqrt{2}} & -54 & -54 \end{bmatrix} \qquad \tilde{\Phi}y = \frac{f}{121}\begin{bmatrix} 84 & -27 & \frac{-50}{\sqrt{2}} & \frac{-50}{\sqrt{2}} \\ -27 & -108 & \frac{42}{\sqrt{2}} & \frac{42}{\sqrt{2}} \\ \frac{-50}{\sqrt{2}} & \frac{42}{\sqrt{2}} & 12 & 12 \\ \frac{-50}{\sqrt{2}} & \frac{42}{\sqrt{2}} & 12 & 12 \end{bmatrix}.$$

*Indeed, each parameter of* $(a,b,c,d,e,f)$ *maps to a fully dense matrix.*

While diagonal idempotents may seem rare, they actually occur quite naturally. Indeed, any *coherent* algebra [142, 69] (ubiquitous in the symmetry reduction of semidefinite optimization problems) has a set of diagonal idempotents (so-called *fibers* of an underlying *coherent configuration*). These idempotents are orthogonal and sum to the identity; hence, they induce a partition of $[n]$. Frequently, these partitions are naturally interpreted as the orbit partition of a particular group action. Chapter 7 will introduce a wider class of subspaces with diagonal idempotents.

## ■ 6.7 Conclusion

We showed how to construct isomorphisms between Euclidean Jordan algebras and gave explicit algorithms for this task. Via the Cayley-Dickson construction, we avoid explicit reference to the isomorphism classes of the simple ideals, leading to a succinct algorithm with no case statements (cf. [89]). Degrees-of-freedom available in this construction were also illustrated, which include choices of Jordan frame, Jordan matrix units, and coordinate-algebra isomorphism. We discussed optimizations for computational efficiency; notably, we illustrated that diagonal idempotents lead to sparse Jordan frames and ultimately sparse isomorphisms.

## ■ 6.8 Appendix

### ■ 6.8.1 Peirce associativity

Off-diagonal Peirce spaces satisfy the following associativity identities:

**Lemma 6.8.1** (Lemma 1 of [73])**.** *Consider a Peirce decomposition* $\bigoplus_{1\le i\le j\le n}\mathbf{J}_{ij}$ *and let* $i,j,k,l$ *be distinct. Then,*

- $x_{ij}\circ(y_{jk}\circ z_{kl}) = (x_{ij}\circ y_{jk})\circ z_{kl}$
- $(x_{ij}\circ y_{jk})\circ y_{jk} = \frac{1}{2}x_{ij}\circ(y_{jk}\circ y_{jk})$

We make two observations about the first identity. First, the indices $(i,j),(j,k),(k,l)$ form a path of length three in the complete graph with $n$ nodes. Two, this identity is equivalent to the multiplication operators of $x_{ij}$ and $z_{kl}$ commuting on the subspace spanned by $y_{jk}$. A generalization to longer paths stated in terms of commuting multiplication operators follows:

**Lemma 6.8.2.** *Suppose the list $p_1, \ldots, p_m \in [n] \times [n]$ of ordered pairs forms a simple path in the complete graph on $n$ nodes. Fix a Peirce decomposition $\bigoplus_{1 \le i \le j \le n} \mathbf{J}_{ij}$. For all $x_i \in \mathbf{J}_{p_i}$, the following relationship holds*

$$L_{x_m} L_{x_{m-1}} \cdots L_{x_2} x_1 = L_{x_1} L_{x_2} \cdots L_{x_{m-1}} x_m,$$

*where $L_{x_i} : \mathbf{J} \to \mathbf{J}$ denotes the multiplication operator $t \mapsto x_i \circ t$.*

*Proof.* If $m > n$, let $L_{x_{m:n}} = L_{x_m} L_{x_{m-1}} L_{x_{m-2}} \cdots L_{x_n}$, if $m < n$, let

$$L_{x_{m:n}} = L_{x_m} L_{x_{m+1}} L_{x_{m+2}} \cdots L_{x_n},$$

and if $m = n$ let $L_{x_{m:n}} = L_{x_m}$. With this notation, we want to show for $m \ge 3$ that

$$L_{x_{m:2}} x_1 = L_{x_{1:m-1}} x_m.$$

We proceed by induction, noting the base case ($m = 3$) holds by the Lemma 6.8.1 identity $(x_1 \circ x_2) \circ x_3 = x_1 \circ (x_2 \circ x_3)$. For $m > 3$, we have

$$\begin{aligned} L_{x_{m:2}} x_1 &= x_m \circ (x_{m-1} \circ (L_{x_{(m-2):2}} x_1)) \\ &= (L_{x_{(m-2):2}} x_1) \circ (x_{m-1} \circ x_m) \qquad (6.7) \\ &= (L_{x_{1:m-3}} x_{m-2}) \circ (x_{m-1} \circ x_m) \qquad (6.8) \end{aligned}$$

where line (6.7) uses Lemma 6.8.1 and line (6.8) uses the inductive hypothesis. If $m = 4$, line (6.8) equals $(x_1 \circ x_2) \circ (x_3 \circ x_4)$ which, by Lemma 6.8.1, equals $x_1 \circ (x_2 \circ (x_3 \circ x_4))$ showing the desired result. For $m > 4$, we can apply Lemma 6.8.1 iteratively. For clarity, assume $m > 6$. Then, continuing from (6.8), these iterative applications take the form:

$$\begin{aligned} &= L_{x_1}((L_{x_{2:m-3}} x_{m-2}) \circ L_{x_{m-1}} x_m) \\ &= L_{x_{1:2}}((L_{x_{3:m-3}} x_{m-2}) \circ L_{x_{m-1}} x_m)) \\ &\quad \vdots \\ &= L_{x_{1:m-4}}((L_{x_{m-3}} x_{m-2}) \circ L_{x_{m-1}} x_m) \\ &= L_{x_{1:m-3}}(x_{m-2} \circ L_{x_{m-1}} x_m) \\ &= L_{x_{1:m-2}} L_{x_{m-1}} x_m \\ &= L_{x_{1:m-1}} x_m \end{aligned}$$

□

## ■ 6.8.2 Coordinate-algebra isomorphisms

To prove Theorem 6.4.2, we used the fact coordinate-algebra isomorphisms are isometries. This is a special case of the following (stated without proof in Jacobson [74, Section 3]):

**Lemma 6.8.3.** *Let $T : \mathbf{E}^A \to \mathbf{E}^B$ be an isomorphism between Euclidean Hurwitz algebras $\mathbf{E}^A$ and $\mathbf{E}^B$. Then, $T$ is an isometry with respect to $\|\cdot\|_{\mathbf{E}^A}$ and $\|\cdot\|_{\mathbf{E}^B}$.*

*Proof.* Let $u$ denote the identity of $\mathbf{E}^A$ and suppose $x = \lambda u$. Then $\|x\|_{\mathbf{E}^A} = |\lambda|$. Further, since $T$ is an isomorphism, $Tu$ is the identity of $\mathbf{E}^B$, implying $\|Tu\|_{\mathbf{E}^B} = 1$. This gives

$$\|Tx\|_{\mathbf{E}^B} = T|\lambda u\|_{\mathbf{E}^B} = \lambda\|Tu\|_{\mathbf{E}^B} = |\lambda|$$

showing $\|x\|_{\mathbf{E}^A} = \|Tx\|_{\mathbf{E}^B}$ when $x$ is in the span of $u$. Now suppose $x \in (\text{span}\{u\})^\perp$. Then, by [74, Section 3],

$$x \times x = (-\|x\|^2_{\mathbf{E}^A})u, \qquad Tx \times Tx = (-\|Tx\|^2_{\mathbf{E}^B})Tu.$$

Hence, if $T(x \times x) = Tx \times Tx$, it must hold that

$$\|x\|^2_{\mathbf{E}^A} = \|Tx\|^2_{\mathbf{E}^B}.$$

□

Chapter 7

# Combinatorial variations and computational results

In Chapter 5, we gave algorithms for finding a subspace that provably intersects the solution set of a cone program formulated over the cone-of-squares of a Euclidean Jordan algebra (i.e., a symmetric cone). In this chapter, we modify one of these procedures to address issues of sparsity and efficiency. Specifically, we develop a variant that trades-off the dimension of the identified subspace with the storage complexity of a basis. These variants restrict to subspaces that have certain combinatorial descriptions. Like in Chapter 5, the identified subspace will be a subalgebra of the Euclidean Jordan algebra. Further, it will frequently contain diagonal idempotents, enabling construction of a sparse isomorphism (Chapter 6.6) and, ultimately, a sparse projected reformulation (Chapter 1.2.5) of the cone program. We also include computational results comparing these combinatorial variations to the original algorithm of Chapter 5.

## 7.1 Preliminaries

To enable combinatorial descriptions of subspaces, this chapter restricts to $\mathbb{S}^n$, the Euclidean Jordan algebra of $n \times n$ symmetric matrices equipped with Jordan product $X \circ Y := \frac{1}{2}(XY + YX)$ and inner product $\langle X, Y \rangle := \operatorname{Tr} XY$. In other words, this chapter restricts to semidefinite programs (SDPs). Recall any special Jordan algebra is isomorphic to a subalgebra of $\mathbb{S}^n$ for some $n$ (Proposition 1.6.4). Hence, by finding isomorphisms, one can execute the presented algorithms on cone programs formulated over special algebras.[1]

[1] This approach comes with two caveats. First, execution depends on the isomorphism (which is not unique). Second, the order $n$ of the isomorphic subalgebra may be large; see Section 1.6.5.)

The SDP of interest has decision variable $X \in \mathbb{S}^n$ and takes the following form

$$\begin{aligned} \text{minimize} \;\; & \operatorname{Tr} CX \\ \text{subject to} \;\; & X \in X_0 + \mathcal{L}, \\ & X \in \mathbb{S}^n_+, \end{aligned}$$

where $C \in \mathbb{S}^n$ and $X_0 \in \mathbb{S}^n$ are fixed and $\mathcal{L} \subseteq \mathbb{S}^n$ is a linear subspace. As shown in Chapter 5, any *admissible subspace* $\mathcal{S}$ intersects the set of solutions, where admissible means the orthogonal projection onto $\mathcal{S}$ satisfies the Constraint Set Invariance Conditions (Definition 1.4.1). As also shown, any admissible subspace that is *unital* (meaning it contains a unit element for Jordan multiplication) is a subalgebra. Given these appealing properties, we presented an algorithm for finding the *minimal-unital subspace*

$$\mathcal{S}_{opt} = \bigcap \{\mathcal{S} \subseteq \mathbb{S}^n : \mathcal{S} \text{ is admissible and unital}\},$$

which is also admissible and unital since intersection preserves these properties.[2]

This chapter considers combinatorial analogues of $\mathcal{S}_{opt}$ induced by three families of subspaces: the *coordinate subspaces*, the *partition subspaces*, and the 0/1 *subspaces* (defined shortly). Each family is also closed under intersection. Hence, each leads to a refinement of $\mathcal{S}_{opt}$:

$$\mathcal{S}_{coord} := \bigcap \{\mathcal{S} \subseteq \mathbb{S}^n : \mathcal{S} \text{ is admissible, unital, and a coordinate subspace}\},$$

$$\mathcal{S}_{part} := \bigcap \{\mathcal{S} \subseteq \mathbb{S}^n : \mathcal{S} \text{ is admissible, unital, and a partition subspace}\},$$

$$\mathcal{S}_{0/1} := \bigcap \{\mathcal{S} \subseteq \mathbb{S}^n : \mathcal{S} \text{ is admissible, unital, and a 0/1 subspace}\}.$$

Figure 7.1 summarizes the basic properties of these subspaces.

### ■ 7.1.1 Combinatorial families of subspaces

A coordinate, partition, or 0/1 subspace has an orthogonal basis $\mathcal{B} \subset \{0,1\}^{n \times n}$ of 0/1 matrices. Equivalently, it is the span of 0/1 matrices that have pairwise *disjoint support*, where, letting $[n] := \{1, \ldots, n\}$, the *support* of $X \in \mathbb{S}^n$ is the following subset of $[n] \times [n]$:

$$\operatorname{supp}(X) := \{(i,j) \in [n] \times [n] : X_{ij} \neq 0\}.$$

[2] Note that $\mathcal{S}_{opt}$ was denoted $\mathcal{S}_{min,unit}$ in Chapter 5. To ease notation, we use the shorter subscript $\mathcal{S}_{opt}$, meaning 'optimal.'

The coordinate subspaces, partition subspaces, and 0/1 subspaces are differentiated by the structure of $\mathcal{B}$.

#### Coordinate subspaces

Suppose $\mathcal{B} \subseteq \{0,1\}^{n\times n}$ is a set of pairwise orthogonal 0/1 matrices. Then the span of $\mathcal{B}$ is a *coordinate subspace* if each $B \in \mathcal{B}$ is a standard basis matrix of $\mathbb{S}^n$. That is, each $B \in \mathcal{B}$ is in the span of $E_{ij} + E_{ji}$ for some $(i,j) \in [n] \times [n]$, where $E_{ij} \in \mathbb{R}^{n\times n}$ is the 0/1 matrix with support equal to $(i,j)$. Example coordinate subspaces of $\mathbb{S}^3$ are

$$\mathcal{S}_1 = \begin{pmatrix} a & b & 0 \\ b & c & d \\ 0 & d & e \end{pmatrix}, \qquad \mathcal{S}_2 = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}, \qquad \mathcal{S}_3 = \begin{pmatrix} a & b & 0 \\ b & c & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Coordinate subspaces are closed under intersection. Indeed, a basis for the intersection is just the intersection of bases, i.e., if coordinate subspaces $\mathcal{S}_1$ and $\mathcal{S}_2$ have 0/1 orthogonal bases $\mathcal{B}_1$ and $\mathcal{B}_2$, then a basis for $\mathcal{S}_1 \cap \mathcal{S}_2$ is just $\mathcal{B}_1 \cap \mathcal{B}_2$.

Finally, coordinate subspaces are in one-to-one correspondence with *symmetric relations*, i.e., subsets $\mathcal{R}$ of $[n]\times[n]$ satisfying $(i,j) \in \mathcal{R}$ if and only if $(j,i) \in \mathcal{R}$. Specifically, the support of $\sum_{B\in\mathcal{B}} B$ is a symmetric relation. Conversely, given a symmetric relation, the span of $\{E_{ij} + E_{ji} : (i,j) \in \mathcal{R}\}$ is a coordinate subspace.

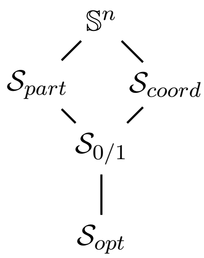
#### Partition subspaces

Suppose $\mathcal{B} \subseteq \{0,1\}^{n\times n}$ is a set of pairwise orthogonal 0/1 matrices. The span of $\mathcal{B}$ is a *partition subspace* if $\sum_{B\in\mathcal{B}} B$ is the matrix of all ones. Since the matrices in $\mathcal{B}$ have disjoint support, the supports of $B \in \mathcal{B}$ form a partition of $[n] \times [n]$—hence, the name partition subspace. Example partition subspaces of $\mathbb{S}^3$ are given by

$$\mathcal{S}_1 = \begin{pmatrix} a & a & b \\ a & a & c \\ b & c & d \end{pmatrix}, \qquad \mathcal{S}_2 = \begin{pmatrix} a & b & b \\ b & c & b \\ b & b & d \end{pmatrix}, \qquad \mathcal{S}_3 = \begin{pmatrix} a & a & b \\ a & a & b \\ b & b & c \end{pmatrix}.$$

Like coordinate subspaces, partition subspaces are closed under intersection. For instance, in the above examples, $\mathcal{S}_1 \cap \mathcal{S}_2 = \mathcal{S}_3$. Generally, if partition subspaces $\mathcal{S}_1$ and $\mathcal{S}_2$ induce partitions $\mathcal{P}_1$ and $\mathcal{P}_2$ of $[n] \times [n]$, then their intersection induces the *join* $\mathcal{P}_1 \bigvee \mathcal{P}_2$, the finest partition refined by both $\mathcal{P}_1$ and $\mathcal{P}_2$.

While partition subspaces seem quite special, they arise naturally in symmetry reduction of semidefinite programs. For example, the subset of $\mathbb{S}^n$ that commutes with a group of permutation matrices is a partition subspace defined by an *orbit partition* of $[n] \times [n]$; see, e.g., [37]. The intersection of any *coherent algebra* with the symmetric matrices is also a partition subspace (Section 1.5.3).




| subspace | sparse basis | sparse isomorphism | data type of basis |
|---|---|---|---|
| $\mathcal{S}_{opt}$ | no | no | dense matrices |
| $\mathcal{S}_{part}$ | yes | w/ diag. idemps. | partition |
| $\mathcal{S}_{coord}$ | yes | always | relation |
| $\mathcal{S}_{0/1}$ | yes | w/ diag. idemps. | partition of relation |

**Figure 7.1:** Hasse diagram of set inclusions, sparsity properties, and a data type (mathematical object) used to represent the basis. Sparse isomorphism means a sparse linear map exists between the subspace (when it is a subalgebra) and a particular representation of an isomorphic algebra; see Section 7.3.2.

#### Zero-one subspaces

Suppose $\mathcal{B} \subseteq \{0,1\}^{n\times n}$ is a set 0/1 matrices. Then, the span of $\mathcal{B}$ is a 0/1 subspace simply when the matrices in $\mathcal{B}$ are pairwise orthogonal. Hence, the family of 0/1 subspaces contains coordinate subspaces, partition subspaces and subspaces that fall into neither family:

$$\mathcal{S}_1 = \begin{pmatrix} a & b & 0 \\ b & c & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \mathcal{S}_2 = \begin{pmatrix} a & a & c \\ a & a & c \\ c & c & d \end{pmatrix}, \qquad \mathcal{S}_3 = \begin{pmatrix} a & a & 0 \\ a & a & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

These subspaces are also closed under intersection. For instance, in the above examples, $\mathcal{S}_3 = \mathcal{S}_1 \cap \mathcal{S}_2$. Note that any 0/1 subspace induces a partition of a relation. Specifically, the support of each $B \in \mathcal{B}$ together form a partition of $\cup_{B\in\mathcal{B}} \operatorname{supp}(B)$. Finally, note that 0/1 subspaces are precisely the subspaces closed under entrywise (i.e., Schur) multiplication.

## ■ 7.2 Algorithms

We now give algorithms for finding the minimal partition, coordinate, and 0/1 subspaces. These algorithms are modifications of a Chapter 5 algorithm that finds $\mathcal{S}_{opt}$, the minimal unital subspace. We reproduce this algorithm below:
Here, $P_{\mathcal{L}} : \mathbb{S}^n \to \mathbb{S}^n$ denotes the orthogonal projection on the subspace $\mathcal{L} \subseteq \mathbb{S}^n$, and $C_{\mathcal{L}} \in \mathbb{S}^n$ and $X_{0,\mathcal{L}^\perp} \in \mathbb{S}^n$ denote the points $P_{\mathcal{L}}(C)$ and $X_0 - P_{\mathcal{L}}(X_0)$, respectively.

**Images as polynomial matrices** Modifications of Algorithm 7.1 will find a coordinate, partition, or 0/1 subspace that contains the image of a subspace under the map $X \mapsto X^2$ or the map $P_{\mathcal{L}} : \mathbb{S}^n \to \mathbb{S}^n$. As we will see, this is conveniently done if we represent the image of a subspace under each map as a polynomial matrix, an idea inspired by [142].

**Algorithm 7.1:**

$\mathcal{S} \leftarrow \text{span}\{C_{\mathcal{L}}, X_{0,\mathcal{L}^\perp}\}$
**repeat**
  $\mathcal{S} \leftarrow \mathcal{S} + P_{\mathcal{L}}(\mathcal{S})$
  $\mathcal{S} \leftarrow \mathcal{S} + \text{span}\{X^2 : X \in \mathcal{S}\}$
**until** converged.

Given a basis $\mathcal{B}$ for a subspace $\mathcal{S}$, we construct this polynomial matrix as follows:

$$f_{X^2}(t_{\mathcal{B}}; \mathcal{B}) := \left(\sum_{B \in \mathcal{B}} t_B B\right)^2 \qquad f_{\mathcal{L}}(t_{\mathcal{B}}; \mathcal{B}) := \sum_{B \in \mathcal{B}} t_B P_{\mathcal{L}}(B),$$

where $t_{\mathcal{B}}$ is a vector of scalar indeterminates indexed by $\mathcal{B}$. Note the set of point evaluations of $f_{X^2}(t_{\mathcal{B}}; \mathcal{B})$ equals $\{X^2 : X \in \mathcal{S}\}$, i.e.,

$$\{X^2 : X \in \mathcal{S}\} = \{f_{X^2}(t_{\mathcal{B}}; \mathcal{B})_{|t_{\mathcal{B}} = t^*} : t^* \in \mathbb{R}^{|\mathcal{B}|}\},$$

and similarly for $f_{\mathcal{L}}(t_{\mathcal{B}}; \mathcal{B})$ and $P_{\mathcal{L}}(\mathcal{S})$. The following example illustrates this notation.

**Example 7.2.1.** *For $\mathcal{B} = \{U, V, W\}$, where*

$$U = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad V = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad W = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

*we have $f_{X^2}(t_{\mathcal{B}}; \mathcal{B}) := (t_U U + t_V V + t_W W)^2$. Expanding then shows*

$$f_{X^2}(t_{\mathcal{B}}; \mathcal{B}) = \begin{pmatrix} {t_U}^2 + {t_W}^2 & 0 & t_U t_W & t_U t_V \\ 0 & {t_U}^2 + {t_W}^2 & t_U t_V & t_U t_W \\ t_U t_W & t_U t_V & {t_U}^2 + {t_V}^2 & 0 \\ t_U t_V & t_U t_W & 0 & {t_U}^2 + {t_V}^2 \end{pmatrix}.$$

To ease notation going forward, we will simply write $f_{X^2}(\mathcal{B})$ to mean $f_{X^2}(t_{\mathcal{B}}; \mathcal{B})$, and similarly for $f_{\mathcal{L}}(\mathcal{B})$.

## ■ 7.2.1 The minimal partition subspace

We now modify Algorithm 7.1 to find the partition subspace $\mathcal{S}_{part}$. For this we must first introduce new notation. Given a partition $\mathcal{P}$ of $[n] \times [n]$, we let $\mathcal{B}_{\mathcal{P}} := \{B_P\}_{P \in \mathcal{P}}$, where $B_P \in \mathbb{R}^{n \times n}$ denotes the zero-one valued *characteristic matrix* of the subset $P$, i.e., for $P \in \mathcal{P}$, we have $(B_P)_{ij} = 1$ if $(i, j) \in P$ and $(B_P)_{ij} = 0$ otherwise. For a matrix

$T \in \mathbb{S}^n$, we let $\operatorname{part}(T)$ denote the partition of $[n] \times [n]$ induced by the unique entries of $T$, i.e., $(i,j)$ and $(k,l)$ are in the same class of $\operatorname{part}(T)$ if and only if $T_{ij} = T_{kl}$. We similarly define $\operatorname{part}(f(\mathcal{B}_{\mathcal{P}}))$ by its unique polynomial entries for $f \in \{f_{\mathcal{L}}, f_{X^2}\}$. We illustrate this latter notation below:

**Example 7.2.1** (continued)**.** *For $\mathcal{B}$ defined previously, the polynomial matrix $f_{X^2}(\mathcal{B})$ is given by*

$$f_{X^2}(\mathcal{B}) = \begin{pmatrix} {t_U}^2 + {t_W}^2 & 0 & t_U t_W & t_U t_V \\ 0 & {t_U}^2 + {t_W}^2 & t_U t_V & t_U t_W \\ t_U t_W & t_U t_V & {t_U}^2 + {t_V}^2 & 0 \\ t_U t_V & t_U t_W & 0 & {t_U}^2 + {t_V}^2 \end{pmatrix}. \tag{7.1}$$

*The partition* $\operatorname{part}(f_{X^2}(\mathcal{B}))$ *has five classes induced by the unique polynomial entries of $f_{X^2}(\mathcal{B})$. The characteristic matrices of these classes and the corresponding polynomials are:*

$$\begin{pmatrix} 1&0&0&0\\0&1&0&0\\0&0&0&0\\0&0&0&0 \end{pmatrix} \quad \begin{pmatrix} 0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0 \end{pmatrix} \quad \begin{pmatrix} 0&0&1&0\\0&0&0&1\\1&0&0&0\\0&1&0&0 \end{pmatrix} \quad \begin{pmatrix} 0&0&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&0 \end{pmatrix} \quad \begin{pmatrix} 0&0&0&0\\0&0&0&0\\0&0&1&0\\0&0&0&1 \end{pmatrix}$$

$$ {t_U}^2 + {t_W}^2 \qquad\qquad 0 \qquad\qquad t_U t_W \qquad\qquad t_U t_V \qquad\qquad {t_U}^2 + {t_V}^2. $$

We can now state an algorithm.

**Theorem 7.2.1.** *For partitions $\mathcal{P}_1$ and $\mathcal{P}_2$ of $[n] \times [n]$, let $\mathcal{P}_1 \bigwedge \mathcal{P}_2$ denote the coarsest partition of $[n] \times [n]$ that refines both $\mathcal{P}_1$ and $\mathcal{P}_2$. Let $\mathcal{P}$ be the partition returned by*

$\mathcal{P} \leftarrow \operatorname{part}(C_{\mathcal{L}}) \bigwedge \operatorname{part}(X_{0,\mathcal{L}^\perp})$
**repeat**
  $\mathcal{P} \leftarrow \mathcal{P} \bigwedge \operatorname{part}(f_{\mathcal{L}}(\mathcal{B}_{\mathcal{P}}))$
  $\mathcal{P} \leftarrow \mathcal{P} \bigwedge \operatorname{part}(f_{X^2}(\mathcal{B}_{\mathcal{P}}))$
**until** converged.

*The minimal partition subspace $\mathcal{S}_{part}$ equals the span of the characteristic matrices $\mathcal{B}_{\mathcal{P}}$.*

Correctness of this algorithm follows from a simple induction argument and the following easily checkable fact: the partition subspace spanned by $\mathcal{B}_{\mathcal{P}}$ contains $X \in \mathbb{S}^n$ if and only if $\mathcal{P}$ refines $\operatorname{part}(X)$.

## ■ 7.2.2 The minimal coordinate subspace

We now show how to find $\mathcal{S}_{coord}$, the minimal coordinate subspace. We can find $\mathcal{S}_{coord}$ using the same basic approach that finds the minimal partition subspace $\mathcal{S}_{part}$. Instead of iteratively refining a partition, we will now iteratively grow a relation.

To state an algorithm, we need more notation. Specifically, for a symmetric relation $\mathcal{R} \subseteq [n] \times [n]$, we let $\mathcal{B}_{\mathcal{R}}$ denote the 0/1 basis for the span of $\{E_{ij} + E_{ji} : (i,j) \in \mathcal{R}\}$. We also define the support of a polynomial matrix $X$ to be the subset of $(i,j) \in [n] \times [n]$ for which $X_{ij}$ is not the zero polynomial. The following example illustrates this latter notation for the polynomial matrix $f(\mathcal{B}_{\mathcal{R}})$.

**Example 7.2.1** (continued)**.** *For $\mathcal{B}$ defined previously, the polynomial matrix $f_{X^2}(\mathcal{B})$ is given by*

$$f_{X^2}(\mathcal{B}) = \begin{pmatrix} {t_U}^2 + {t_W}^2 & 0 & t_U t_W & t_U t_V \\ 0 & {t_U}^2 + {t_W}^2 & t_U t_V & t_U t_W \\ t_U t_W & t_U t_V & {t_U}^2 + {t_V}^2 & 0 \\ t_U t_V & t_U t_W & 0 & {t_U}^2 + {t_V}^2 \end{pmatrix}.$$

*For $n = 4$, the support* $\operatorname{supp}(f_{X^2}(\mathcal{B}))$ *is the complement of* $\{(1,2),(2,1),(3,4),(4,3)\} \subseteq [n] \times [n]$.

An algorithm for finding $\mathcal{S}_{coord}$ now follows.

**Theorem 7.2.2.** *Let $\mathcal{R}$ be the relation returned by*

$$\begin{array}{l} \mathcal{R} \leftarrow \operatorname{supp}(C_{\mathcal{L}}) \cup \operatorname{supp}\left(X_{0,\mathcal{L}^\perp}\right) \\ \textbf{repeat} \\ \quad \mathcal{R} \leftarrow \mathcal{R} \cup \operatorname{supp}(f_{\mathcal{L}}(\mathcal{B}_{\mathcal{R}})) \\ \quad \mathcal{R} \leftarrow \mathcal{R} \cup \operatorname{supp}(f_{X^2}(\mathcal{B}_{\mathcal{R}})) \\ \textbf{until} \ \text{converged.} \end{array}$$

*The optimal coordinate subspace $\mathcal{S}_{coord}$ equals the span of $\mathcal{B}_{\mathcal{R}}$.*

Correctness of this algorithm follows from a simple induction argument and the following easily checkable fact: the coordinate subspace spanned by $\mathcal{B}_{\mathcal{R}}$ contains $X \in \mathbb{S}^n$ if and only if the relation $\mathcal{R}$ contains the support of $X$.

## ■ 7.2.3 The minimal $0/1$ subspace

A procedure for finding $\mathcal{S}_{0/1}$ combines features of the algorithms presented in the previous two sections. It iteratively grows a relation $\mathcal{R}$ representing the indices $(i,j)$ for which $X_{ij}$ is not identically zero for all $X \in \mathcal{S}_{0/1}$. It also maintains a partition $\mathcal{P}$ of $\mathcal{R}$ whose characteristic matrices (at termination) are the 0/1 basis for $\mathcal{S}_{0/1}$. The procedure and statement of its correctness follow, where $\operatorname{part}_{\mathcal{R}}(T)$ denotes the partition of $\mathcal{R} \subseteq [n] \times [n]$ induced by the unique entries of a matrix $T$ with support contained in $\mathcal{R}$.

**Theorem 7.2.3.** *Let $\mathcal{R}$ and $\mathcal{P}$ be the relation and partition of $\mathcal{R}$ returned by*

*Initialize* $\mathcal{R}$ *to* $\text{supp}(C_{\mathcal{L}}) \bigcup \text{supp}(X_{0,\mathcal{L}^\perp})$
*Initialize* $\mathcal{P}$ *to* $\text{part}_{\mathcal{R}}(C_{\mathcal{L}}) \bigwedge \text{part}_{\mathcal{R}}(X_{0,\mathcal{L}^\perp})$
**repeat**
  **for** $f \in \{f_{\mathcal{L}}, f_{X^2}\}$ **do**
    *Replace* $\mathcal{R}$ *with* $\mathcal{R} \cup \text{supp}\,(f(\mathcal{B}_{\mathcal{P}}))$
    *Add class* $\mathcal{R} \setminus (\cup_{P\in\mathcal{P}} P)$ *to* $\mathcal{P}$
    *Replace* $\mathcal{P}$ *with refinement* $\mathcal{P} \bigwedge \text{part}_{\mathcal{R}}\,(f(\mathcal{B}_{\mathcal{P}}))$
  **end**
**until** converged.

*The minimal* $0/1$ *subspace* $\mathcal{S}_{0/1}$ *equals the span of the characteristic matrices* $\mathcal{B}_{\mathcal{P}}$*.*

As indicated, this algorithm alternates between growing the relation $\mathcal{R}$, adding a single class $\mathcal{R} \setminus (\cup_{P\in\mathcal{P}} P)$ to $\mathcal{P}$ (such that it partitions $\mathcal{R}$), and refining $\mathcal{P}$.

## ■ 7.2.4 Randomization via sampling

Note that each algorithm need not explicitly construct symbolic representations of $f_{\mathcal{L}}(\mathcal{B})$ and $f_{X^2}(\mathcal{B})$; instead, one can evaluate these matrices at a generic point by evaluating the maps $X \mapsto X^2$ and $P_{\mathcal{L}} : \mathbb{S}^n \to \mathbb{S}^n$ at a random combination of basis elements in $\mathcal{B}$. Consider, for instance, a point evaluation of $f_{X^2}(\mathcal{B}_{\mathcal{P}})$ at $t^* \in \mathbb{R}^{|\mathcal{B}_{\mathcal{P}}|}$, i.e., consider

$$f_{X^2}(\mathcal{B}_{\mathcal{P}})_{|t=t^*} := \left( \sum_{B\in\mathcal{B}_{\mathcal{P}}} t^*_B B \right)^2 .$$

Clearly the support of $f_{X^2}(\mathcal{B}_{\mathcal{P}})$ and $f_{X^2}(\mathcal{B}_{\mathcal{P}})_{|t=t^*}$ are the same except for $t^*$ in a measure-zero subset of $\mathbb{R}^{|\mathcal{B}_{\mathcal{P}}|}$. Similarly, $f_{X^2}(\mathcal{B}_{\mathcal{P}})$ and $f_{X^2}(\mathcal{B}_{\mathcal{P}})_{|t=t^*}$ induce the same partition of $[n] \times [n]$ , i.e.,

$$\text{part}\,(f_{X^2}(\mathcal{B}_{\mathcal{P}})) = \text{part}\left( f_{X^2}(\mathcal{B}_{\mathcal{P}})_{|t=t^*} \right),$$

except for $t^*$ also in a measure-zero subset. The following illustrates this latter fact.

**Example 7.2.1** (continued)**.** *For* $\mathcal{B}$ *defined previously, the point evaluation* $f_{X^2}(\mathcal{B})_{|t=t^*_{\mathcal{B}}}$ *at* $t^*_{\mathcal{B}} = (2,3,4)$ *is*

$$f_{X^2}(\mathcal{B})_{|t=t^*_{\mathcal{B}}} = (2U + 3V + 4W)^2 = \begin{pmatrix} 20 & 0 & 8 & 6 \\ 0 & 20 & 6 & 8 \\ 8 & 6 & 13 & 0 \\ 6 & 8 & 0 & 13 \end{pmatrix}.$$

*We see this point evaluation induces the same partition as the polynomial matrix* $f_{X^2}(\mathcal{B})$ *given by* (7.1)*. In other words, the partitions* $\text{part}\,(f_{X^2}(\mathcal{B}_{\mathcal{P}}))$ *and* $\text{part}\left(f_{X^2}(\mathcal{B})_{|t=t^*_{\mathcal{B}}}\right)$ *are the same.*

## 7.3 Combinatorial subalgebras

We study the subalgebras of $\mathbb{S}^n$ that are partition or coordinate subspaces in more detail. We first characterize these subalgebras. We then discuss existence of sparse isomorphisms onto these algebras.

### 7.3.1 Characterizations

#### Coordinate subalgebras and transitive relations

Coordinate subalgebras—i.e., coordinate subspaces that are subalgebras—are in one-to-one correspondence with symmetric relations that are transitive. Consider the following.

**Lemma 7.3.1.** *Let $\mathcal{R} \subseteq [n] \times [n]$ be a symmetric relation. The following are equivalent.*

- *The relation $\mathcal{R}$ is transitive.*
- *The coordinate subspace $\mathcal{S}_{\mathcal{R}}$ spanned by $\{E_{ij} + E_{ji} : (i,j) \in \mathcal{R}\}$ is a subalgebra of $\mathbb{S}^n$, i.e., $\mathcal{S} \supseteq \{X^2 : X \in \mathcal{S}\}$.*

*Proof.* Transitivity easily follows from the Peirce Multiplication Rules (Lemma 6.1.1). For the converse, note that $\mathcal{R}$ partitions $\{i : (i,i) \in \mathcal{R}\}$ into equivalence classes, where $(i,j) \in \mathcal{R}$ iff $i \equiv j$. This implies $\mathcal{S}$ is block-diagonal up to simultaneous permutation of rows and columns; specifically, up to permutation, it equals

$$\left(\oplus_{i=1}^{d} \mathbb{S}^{d_i}\right) \oplus 0_{n-r},$$

where $r$ is the cardinality of $\{i : (i,i) \in \mathcal{R}\}$ and $d_i$ is the cardinality of the $i^{th}$ equivalence class. From this decomposition, invariance under $X \mapsto X^2$ is obvious. □

This allows one to replace a step of the Theorem 7.2.2 algorithm with computation of a transitive closure:

$\mathcal{R} \leftarrow \operatorname{supp}(C_{\mathcal{L}}) \cup \operatorname{supp}\left(X_{0,\mathcal{L}^\perp}\right)$
**repeat**
  $\mathcal{R} \leftarrow \mathcal{R} \cup \operatorname{supp}(f_{\mathcal{L}}(\mathcal{B}_{\mathcal{R}}))$
  $\mathcal{R} \leftarrow$ The transitive closure of $\mathcal{R}$
**until** converged.

One can find the transitive closure from scratch in $n^3$ time (e.g., [55]). Algorithms also exist for *maintaining* a transitive closure (e.g., [78]) as elements are added to $\mathcal{R}$.

#### Partition subalgebras and coherent configurations

Partition subalgebras—i.e., partition subspaces that are subalgebras— do not have a clean characterization (at least that we are aware of). They do, however, relate to well studied partitions called *coherent configurations*.

**Definition 7.3.1** (e.g., [29])**.** *A partition $\mathcal{P}$ of $[n] \times [n]$ is a* coherent configuration *if its characteristic matrices $\mathcal{B}_\mathcal{P}$ satisfy*

- $X^T \in \mathcal{B}_\mathcal{P}$ *for all* $X \in \mathcal{B}_\mathcal{P}$*;*
- $XY \in \operatorname{span} \mathcal{B}_\mathcal{P}$ *for all* $X, Y \in \mathcal{B}_\mathcal{P}$*;*
- $I \in \operatorname{span} \mathcal{B}_\mathcal{P}$*.*

Note that the characteristic matrices of a coherent configuration form a basis for a *-subalgebra of $\mathbb{R}^{n \times n}$ containing the identity matrix $I$. Based on this, we define a *Jordan configuration* as a partition of $[n] \times [n]$ whose characteristic matrices form a basis for a Jordan subalgebra of $\mathbb{S}^n$ containing $I$:

**Definition 7.3.2.** *A partition $\mathcal{P}$ of $[n]\times[n]$ is a* Jordan configuration *if its characteristic matrices $\mathcal{B}_\mathcal{P}$ satisfy*

- $X = X^T$ *for all* $X \in \mathcal{B}_\mathcal{P}$*;*
- $XY + YX \in \operatorname{span} \mathcal{B}_\mathcal{P}$ *for all* $X, Y \in \mathcal{B}_\mathcal{P}$*;*
- $I \in \operatorname{span} \mathcal{B}_\mathcal{P}$*.*

The following 'characterization' is then obtained as a restatement of definitions.

**Proposition 7.3.1.** *Let $\mathcal{S}_\mathcal{P}$ be a partition subspace that contains the identity matrix $I$. The following statements are equivalent.*

- *The subspace $\mathcal{S}_\mathcal{P}$ is a subalgebra, i.e., it is invariant under $X \mapsto X^2$.*
- *The partition $\mathcal{P}$ is a Jordan configuration.*

Note that if $\mathcal{B}_\mathcal{C}$ is the set of characteristic matrices of a coherent configuration $\mathcal{C}$, then

$$\left\{B + B^T : B \in \mathcal{B}_\mathcal{C}\right\}$$

is the set of characteristic matrices of a Jordan configuration. It is an open question if all Jordan configurations arise this way. Cameron posed this question for *Jordan schemes* [29]—the Jordan configurations whose characteristic matrices include $I$.

Given a partition $\mathcal{P}$ of $[n] \times [n]$, an algorithm of Weisfeiler [142] finds the coarsest coherent configuration refining $\mathcal{P}$; see also [8]. An algorithm for finding the analogous Jordan configuration (assuming $\mathcal{P}$ has symmetric characteristic matrices) is easily stated using the notation of Section 7.2.1:

$\mathcal{P} \leftarrow \mathcal{P} \bigwedge \operatorname{part} I$
**repeat**
| $\mathcal{P} \leftarrow \mathcal{P} \bigwedge \operatorname{part} f_{X^2}(\mathcal{B}_\mathcal{P})$
**until** converged.

Recall that $f_{X^2}(\mathcal{B}_\mathcal{P})$ is the polynomial matrix $(\sum_{B\in\mathcal{B}_\mathcal{P}} t_B B)^2$ in the set $\{t_B\}_{B\in\mathcal{B}_\mathcal{P}}$ of *commuting* indeterminates whose unique entries induce the partition part $(f_{X^2}(\mathcal{B}_\mathcal{P}))$. By replacing $\{t_B\}_{B\in\mathcal{B}_\mathcal{P}}$ with *non-commuting* indeterminates, i.e., by treating $t_A t_B \neq t_B t_A$ in construction of part $\left((\sum_{B\in\mathcal{B}_\mathcal{P}} t_B B)^2\right)$, this algorithm reduces to that of [142].

## ■ 7.3.2 Sparse isomorphisms

### Coordinate subalgebras

As just shown, the coordinate subspace $\mathcal{S}_\mathcal{R}$ is a subalgebra if and only if the symmetric relation $\mathcal{R}$ is transitive. As a consequence, decomposing a coordinate subalgebra into minimal ideals is trivial: one simply finds the disjoint subsets of $[n] := \{1, \ldots, n\}$ induced by $\mathcal{R}$. These subsets are the *equivalence classes* of $\mathcal{R}$ when it is also reflexive (and hence an equivalence relation).

**Example 7.3.1.** *Examples of coordinate subalgebras $\mathcal{S}_\mathcal{R}$ and the corresponding subsets of $[n]$ are*

$$\begin{pmatrix} * & 0 & * & 0 \\ 0 & * & 0 & * \\ * & 0 & * & 0 \\ 0 & * & 0 & * \end{pmatrix}, \qquad \begin{pmatrix} * & * & 0 & 0 \\ * & * & 0 & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} * & 0 & 0 & * \\ 0 & * & * & 0 \\ 0 & * & * & 0 \\ * & 0 & 0 & * \end{pmatrix}.$$

$$\{1,3\},\{2,4\} \qquad \{1,2\},\{3\} \qquad \{1,4\},\{2,3\}$$

*Here, $*$ marks the $(i,j)^{th}$ entry when $(i,j) \in \mathcal{R}$. In the first and third examples, $\mathcal{R}$ is an equivalence relation and hence induces a partition of $[n]$.*

If $d_1, \ldots, d_r$ are the cardinalities of these subsets, then $\mathcal{S}_\mathcal{R}$ is isomorphic to the direct-sum $\oplus_{i=1}^r \mathbb{S}^{d_i}$. Further, we can express the isomorphism using a permutation matrix. Specifically, letting $q = \sum_{i=1}^r d_i$ and $0_{n-q}$ equal the $(n-q)\times(n-q)$ zero matrix, we have that

$$\mathcal{S}_\mathcal{R} = \Phi \cdot \left(\left(\oplus_{i=1}^r \mathbb{S}^{d_i}\right) \oplus 0_{n-q}\right)$$

where the isomorphism $\Phi$ satisfies

$$\Phi(X) = PXP^T \qquad \forall X \in \left(\oplus_{i=1}^r \mathbb{S}^{d_i}\right) \oplus 0_{n-q}$$

for some permutation matrix $P \in \mathbb{R}^{n\times n}$. Note that constructing a projected reformulation with this isomorphism *never* destroys sparsity. For instance, a projected reformulation of

$$\begin{array}{ll} \text{minimize} & \operatorname{Tr} CX \\ \text{subject to} & \operatorname{Tr} A_i X = b_i \quad \forall i \in [m] \\ & X \in \mathbb{S}^n_+ \end{array}$$

takes the following form

$$\begin{array}{ll} \text{minimize} & \langle \Phi^*(C), Z\rangle \\ \text{subject to} & \langle \Phi^*(A_i), Z\rangle = b_i \quad \forall i \in [m], \\ & Z \in \mathbb{S}^{d_1}_+ \times \cdots \times \mathbb{S}^{d_r}_+ \times 0_{n-r}, \end{array}$$

where $\Phi^*(W)$ simply permutes the rows and columns of $W \in \mathbb{S}^n$ and sets certain entries to zero.

#### Partition subalgebras

As discussed in Section 6.6, isomorphic matrix algebras with diagonal idempotents admit sparse isomorphisms. The fundamental reason is the *Peirce components* $E_iXE_j + E_jXE_i$ (Section 6.1.1) of any $X$ in these algebras are sparse when the idempotents $E_i$ and $E_j$ are diagonal. Hence if such idempotents exist, sparse Peirce components map to sparse Peirce components under some (easily constructed) isomorphism.

A partition subalgebra almost always has diagonal idempotents. (The only exception is the algebra spanned by the all-ones-matrix.) Further, these idempotents sum to $I$ when the algebra contains $I$. The following example illustrates this:

$$\mathcal{S} = \left\{ \begin{bmatrix} a & b & c \\ b & a & c \\ c & c & d \end{bmatrix} : (a,b,c,d) \in \mathbb{R}^4 \right\}, \qquad E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad E_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Note that any $X$ in this algebra has sparse Peirce components

$$X_{11} = \begin{bmatrix} a & b & 0 \\ b & a & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad X_{12} = \begin{bmatrix} 0 & 0 & c \\ 0 & 0 & c \\ c & c & 0 \end{bmatrix}, \quad X_{22} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & d \end{bmatrix}.$$

Specifically, the Peirce component $X_{ij}$ is nonzero only on the submatrices induced by the support of $E_i$ and $E_j$. Given this block sparsity, some authors call partition subalgebras *cellular* [142].

## 7.4 Combinatorial invariant subspaces

### 7.4.1 Partition subspaces and equitable partitions

The partition subspaces invariant under an orthogonal projection $P_{\mathcal{L}} : \mathbb{S}^n \to \mathbb{S}^n$ correspond to the *equitable partitions* of $P_{\mathcal{L}}$. Before defining these objects, we first consider the equitable partitions of a symmetric matrix $A \in \mathbb{S}^m$, which are the partitions of $[m]$ that induce sub-matrices with constant row and column sums [60]. For instance, the following partitions are equitable:

$$\left(\begin{array}{cc|cc|c} a & b & c & c & b \\ b & a & c & c & b \\ \hline c & c & d & d & d \\ c & c & d & d & d \\ \hline b & b & d & d & e \end{array}\right) \qquad \left(\begin{array}{cc|c|c|c} a & b & c & c & b \\ b & a & c & c & b \\ \hline c & c & d & d & d \\ \hline c & c & d & d & d \\ \hline b & b & d & d & e \end{array}\right) \qquad \left(\begin{array}{c|c|cc|c} a & b & c & c & b \\ \hline b & a & c & c & b \\ \hline c & c & d & d & d \\ c & c & d & d & d \\ \hline b & b & d & d & e \end{array}\right)$$

$$\{1,2\},\{3,4\},\{5\} \qquad\qquad \{1,2\},\{3\},\{4\},\{5\} \qquad\qquad \{1\},\{2\},\{3,4\},\{5\}.$$

Observe if a partition is equitable for $A$, then the subspace spanned by its characteristic vectors is invariant under $A$. (Lemma 9.3.2 of [60] proves this for adjacency matrices.) Indeed, for the partition $\{1,2\},\{3,4\},\{5\}$, the characteristic vector $(1,1,0,0,0)^T$ of $\{1,2\}$ satisfies

$$\left(\begin{array}{cc|cc|c} a & b & c & c & b \\ b & a & c & c & b \\ \hline c & c & d & d & d \\ c & c & d & d & d \\ \hline b & b & d & d & e \end{array}\right)\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = (a+b)\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + 2c\begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} + 2b\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Hence, it is in the span of the characteristic vectors of $\{1,2\}$ and $\{3,4\}$ and $\{5\}$.

To generalize equitable partitions to linear maps on $\mathbb{S}^n$, we make the following observation: if a partition of $[m]$ is equitable for $A \in \mathbb{S}^m$, then all pairs of characteristic vectors $x, y \in \{0,1\}^m$ satisfy

$$y \star (Ax) \in \operatorname{span} y,$$

where $\star$ denotes the element-wise product. This suggests the following definition.

**Definition 7.4.1.** *Let $\mathcal{P}$ be a partition of $[n] \times [n]$ with symmetric characteristic matrices $\mathcal{B}_{\mathcal{P}} \subset \mathbb{S}^n$. Let $\Psi : \mathbb{S}^n \to \mathbb{S}^n$ be a self-adjoint linear map. Then, $\mathcal{P}$ is a* matrix equitable partition *of $\Psi$ if for all $X, Y \in \mathcal{B}_{\mathcal{P}}$,*

$$\Psi(X) \star Y \in \operatorname{span} Y,$$

*where* $\Psi(X) \star Y$ *denotes the Schur (i.e., entrywise) product.*

The following is then essentially immediate:

**Lemma 7.4.1.** *Let* $\mathcal{P}$ *be a partition of* $[n] \times [n]$ *with symmetric characteristic matrices* $\mathcal{B}_{\mathcal{P}} \subset \mathbb{S}^n$*. Let* $\Psi : \mathbb{S}^n \to \mathbb{S}^n$ *be a self-adjoint linear map. The following are equivalent:*

- *The span of* $\mathcal{B}_{\mathcal{P}}$ *is an invariant subspace of* $\Psi$*.*
- $\mathcal{P}$ *is a matrix equitable partition of* $\Psi$*.*

*Proof.* Suppose the first statement holds. Then, for all $X \in \mathcal{B}_{\mathcal{P}}$, it holds that $\Psi(X) \in \operatorname{span} \mathcal{B}_{\mathcal{P}}$. Hence, for all $Y \in \mathcal{B}_{\mathcal{P}}$, it holds that $Y \star \Psi(X) \in \{0, Y\}$ since matrices in $\mathcal{B}_{\mathcal{P}}$ have disjoint support. For the other direction, assume $\Psi(X) \star Y \in \operatorname{span} Y$ for all $X, Y \in \mathcal{B}_{\mathcal{P}}$. Then, clearly,

$$\sum_{Y \in \mathcal{B}_{\mathcal{P}}} \Psi(X) \star Y \in \operatorname{span} \mathcal{B}_{\mathcal{P}}.$$

Since $\mathcal{P}$ partitions $[n] \times [n]$, the sum $\sum_{Y \in \mathcal{B}_{\mathcal{P}}} Y$ equals the all-ones-matrix. Hence, $\Psi(X) \in \operatorname{span} \mathcal{B}_{\mathcal{P}}$. The claim then follows by linearity of $\Psi$.

□

Note equitable partitions and matrix equitable partitions form a lattice, reflecting the lattice structure of invariant subspaces. As a subset of the partition lattice, equitable partitions are also closed under join (Lemma 5.3 of [92]), reflecting the fact invariant subspaces are closed under intersection.

## 7.4.2 Coordinate subspaces and connected components

The invariant coordinate subspaces of $P_{\mathcal{L}}$—or any linear map— depend only on the sparsity of a particular matrix representation. For self-adjoint maps, this dependence is naturally expressed using the connected components of a graph. To establish intuition, suppose that the span of $\{e_1, \dots, e_m\} \subseteq \mathbb{R}^n$ is an invariant subspace of a matrix $T \in \mathbb{R}^{n \times n}$. Then, necessarily

$$T = \begin{pmatrix} T_{11} & T_{21} \\ 0 & T_{22} \end{pmatrix},$$

where $T_{11} \in \mathbb{R}^{m \times m}$, $T_{21} \in \mathbb{R}^{m \times n-m}$ and $T_{22} \in \mathbb{R}^{(n-m)\times(n-m)}$. Further, if $T$ is symmetric, then $T_{21} = 0$, which shows the span of $\{e_{m+1}, \dots, e_n\}$ is also an invariant subspace. Treating $T$ as an adjacency matrix of a graph, it follows the connected components—and unions of connected components (since the sum of invariant subspaces is invariant)—define the invariant coordinate subspaces. The following formalizes this statement for linear maps on $\mathbb{S}^n$.

**Lemma 7.4.2.** *Consider an orthogonal basis $\mathcal{B} \subset \mathbb{S}^n$ for $\mathbb{S}^n$ and a self-adjoint linear map $\Psi : \mathbb{S}^n \to \mathbb{S}^n$. Let $G = (\mathcal{B}, \mathcal{E})$ denote the undirected graph for which $\{A, B\} \in \mathcal{E}$ if and only if*

$$\langle A, \Psi(B)\rangle = \langle \Psi(A), B\rangle \neq 0,$$

*and let $\{(\mathcal{B}_k, \mathcal{E}_k)\}_{i=1}^p$ denote its connected components. Finally, let $\mathcal{T} \subseteq \mathcal{B}$. Then the subspace spanned by $\{B\}_{B\in\mathcal{T}}$ is invariant under $\Psi$ if and only if $\mathcal{T} = \bigcup_{k\in S} \mathcal{B}_k$ for some $S \subseteq \{1, 2, \ldots, p\}$.*

*Proof.* If $\mathcal{T} = \bigcup_{k\in S} \mathcal{B}_k$, then $\{A, B\} \notin \mathcal{E}$ for all $A \in \mathcal{T}$ and $B \notin \mathcal{T}$; hence, $\langle \Psi(A), B\rangle = 0$, which shows $\Psi(A)$ is in the orthogonal complement of $\text{span}\{B\}_{B\in\mathcal{B}\setminus\mathcal{T}}$; in other words, $\Psi(A)$ is in $\text{span}\{B\}_{B\in\mathcal{T}}$. Hence, $\text{span}\{B\}_{B\in\mathcal{T}}$ is an invariant subspace.

Conversely, if $\text{span}\{B\}_{B\in\mathcal{T}}$ is invariant subspace, then $\langle \Psi(A), B\rangle = 0$ for all $A \in \mathcal{T}$ and $B \notin \mathcal{T}$. We conclude no edge connects $\mathcal{T}$ with its complement. Hence, for each connected component, $\mathcal{B}_k \subseteq \mathcal{T}$ or $\mathcal{B}_k \subseteq \mathcal{B} \setminus \mathcal{T}$. Taking $S = \{k \in [p] : \mathcal{B}_k \subseteq \mathcal{T}\}$, it follows $\mathcal{T} = \cup_{k\in S}\mathcal{B}_k$. □

To describe the invariant coordinate subspaces of the projection $P_{\mathcal{L}}$, we can apply Lemma 7.4.2 to the orthogonal basis $\{E_{ij} + E_{ji} : (i, j) \in [n] \times [n]\}$. Constructing the graph $G$ then identifies subsets of this basis (and corresponding subsets of $[n] \times [n]$) that span invariant coordinate subspaces.

## ■ 7.5 Examples

We now find admissible subspaces for several example SDPs, exploring trade-offs in dimension, complexity of the cone constraint, and sparsity. To simplify presentation, we will use a common format for original instances and reduced instances. We now overview these formats and their complexity parameters.

**Format of original SDPs** Each primal-dual pair is originally expressed in either SeDuMi [129] or SDPA [58] format and may have a mix of free and conic variables, where the cones are either orthants or cones of psd matrices.[3] From these formats, we eliminate free variables, reformatting the primal problem as

$$\begin{array}{ll} \text{minimize} & \langle C, X\rangle \\ \text{subject to} & \langle A_i, X\rangle = b_i \quad \forall i \in [m] \\ & X \in \mathbb{S}_+^{n_1} \times \cdots \times \mathbb{S}_+^{n_r}, \end{array} \tag{7.2}$$

[3] These formats also allow for Lorentz cones. None of the examples presented, however, use this type of cone.

where $C, A_i \in \mathbb{S}^{n_1} \times \cdots \times \mathbb{S}^{n_r}$ are fixed, and $\langle \cdot, \cdot \rangle$ denotes the inner product obtained by equipping each product term $\mathbb{S}^{n_i}$ with the trace inner product. Note once free variables are eliminated, reformatting amounts to relabeling linear inequalities as semidefinite constraints of order one.

We will in some cases report the number of nonzero (nnz) entries in a description of (7.2); this equals the number of nonzero floating-point numbers needed to store $C$ and $\{A_i\}_{i=1}^m$. We also report a tuple of ranks for (7.2), which is simply the tuple $(n_1, \ldots, n_r)$.

**Formats of projected reformulations** For each SDP, we construct a projected reformulation (Chapter 1.2.5) over $\mathcal{K} \cap \mathcal{S}$ by finding a Jordan isomorphism $\Phi : \mathbf{J} \to \mathcal{S}$ satisfying

$$\mathcal{S} \cap \mathcal{K} = \Phi(\mathcal{K}_1 \times \cdots \times \mathcal{K}_q)$$

for irreducible symmetric cones $\mathcal{K}_i$ and a Jordan algebra $\mathbf{J}$. The projected reformulation takes the following form

$$\begin{array}{ll} \text{minimize} & \langle \Phi^*(C), \hat{X} \rangle \\ \text{subject to} & \langle \Phi^*(A_i), \hat{X} \rangle = b_i \quad \forall i \in T \subseteq [m] \\ & \hat{X} \in \mathcal{K}_1 \times \cdots \times \mathcal{K}_q, \end{array} \tag{7.3}$$

where $T$ indexes a maximal linearly-independent subset of equations. The isomorphism is found using techniques from Chapter 6.

We will in some cases report the number of nonzero (nnz) entries in a description of (7.3); this equals the number of nonzero floating-point numbers needed to store $\Phi^*(C)$ and $\{\Phi^*(A_i)\}_{i \in T}$. We also report a tuple $(r_1, \ldots, r_q)$ of ranks for (7.3). Here, each cone $\mathcal{K}_i$ is the cone-of-squares of a simple algebra; the reported tuple consists of the ranks of these simple algebras.

**Remark 7.5.1.** *For most examples, $\mathcal{K}_1 \times \cdots \times \mathcal{K}_q$ is a product of psd cones $\mathbb{S}^{r_1}_+ \times \cdots \times \mathbb{S}^{r_q}_+$ and the tuple $(r_1, \ldots, r_q)$ indicates their orders—in other words, all minimal ideals of $\mathcal{S}$ are isomorphic to algebras of real symmetric matrices. Given this, one can also find the isomorphism $\Phi$ using techniques from [89]. The only exception is discussed in Section 7.5.1. We also note $\mathbb{S}^2$ is isomorphic to a spin-factor algebra—hence, $\mathbb{S}^2_+$ is isomorphic to a Lorentz cone.*

**Reference subspaces and inclusions** For convenience, we will let $\mathcal{S}_{full} := \mathbb{S}^{n_1} \times \cdots \times \mathbb{S}^{n_r}$ denote the full ambient space of the original instance (7.2). As discussed in [37], an SDP can be restricted to the *-algebra generated by its data matrices. To compare with this restriction, we let

$$\mathcal{S}_{data} := \mathcal{C}_{data} \cap (\mathbb{S}^{n_1} \times \cdots \times \mathbb{S}^{n_r}),$$

where $\mathcal{C}_{data}$ is the *-algebra generated by the problem data $C$ and $\{A_i\}_{i=1}^m$ (using matrix multiplication as a product and transposition as the *-involution). Recall that

$$\mathcal{S}_{opt} \subseteq \mathcal{S}_{0/1} \subseteq \mathcal{S}_{coord} \subseteq \mathcal{S}_{full}.$$

As examples indicate, different inclusions can hold strictly for different examples. We also have

$$\mathcal{S}_{opt} \subseteq \mathcal{S}_{data} \subseteq \mathcal{S}_{full}.$$

As examples show, it often holds that $\mathcal{S}_{data} = \mathcal{S}_{full}$ even when $\mathcal{S}_{opt}$ is (much) smaller than $\mathcal{S}_{full}$.

## ■ 7.5.1 Libraries of problem instances

The first set of SDPs are selected from three publicly-available sources: the parser SOSTOOLS [100], the DIMACS library [104] and a library of structured SDP instances from [38]. Table 7.1 reports the dimensions of the subspaces $\mathcal{S}_{opt}$, $\mathcal{S}_{0/1}$, $\mathcal{S}_{coord}$, $\mathcal{S}_{data}$ and $\mathcal{S}_{full}$. Note the inclusions $\mathcal{S}_{opt} \subseteq \mathcal{S}_{0/1} \subseteq \mathcal{S}_{coord} \subseteq \mathcal{S}_{full}$ hold as expected, and, as Table 7.1 indicates, different ones hold strictly for different instances. For a large fraction of instances, $\mathcal{S}_{full}$ equals $\mathcal{S}_{data}$, implying generating a *-subalgebra from the problem data [37] does not provide reductions for these instances.

**Remark 7.5.2.** *We note the libraries [104, 38] have additional instances on which our method was not effective ($\mathcal{S}_{opt} = \mathcal{S}_{full}$); we do not report results for these instances. The library [38] also has other instances with group symmetry that were too large or too poorly conditioned for a simple MATLAB implementation of our algorithm. We omit these instances.*

#### The Lovasz number of Hamming graphs

We give special attention to the Table 7.2 instances denoted

`hamming_q_x`,     `hamming_q_x_y`,

that were taken from [104]. The optimal values of these SDPs equal the *Lovász number* of a particular graph. For a graph $G$ with vertices $\{1, \ldots, n\}$ and edge set $E$, the Lovász number is the optimal value of

$$\begin{array}{ll} \text{maximize} & \operatorname{Tr} \mathbf{1}\mathbf{1}^{\mathbf{T}} X \\ \text{subject to} & \operatorname{Tr} X = 1 \\ & \operatorname{Tr}(E_{ij} + E_{ji}) X = 0 \qquad \forall (i,j) \in E, \end{array} \tag{7.4}$$

| instance | $\mathcal{S}_{opt}$ | $\mathcal{S}_{0/1}$ | $\mathcal{S}_{coord}$ | $\mathcal{S}_{data}$ | $\mathcal{S}_{full}$ | References |
|---|---|---|---|---|---|---|
| sosdemo2 | 25 | 25 | 28 | 103 | 103 | Instances from [100] |
| sosdemo4 | 11 | 11 | 85 | 630 | 630 | |
| sosdemo5 | 226 | 816 | 816 | 816 | 816 | |
| sosdemo6 | 49 | 49 | 327 | 462 | 462 | |
| sosdemo7 | 40 | 40 | 68 | 68 | 68 | |
| sosdemo9 | 26 | 26 | 26 | 78 | 78 | |
| sosdemo10 | 78 | 78 | 78 | 254 | 254 | |
| hamming_7_5_6 | 5 | 5 | 8256 | 8256 | 8256 | Instances from [104] |
| hamming_8_3_4 | 5 | 5 | 32896 | 32896 | 32896 | |
| hamming_9_5_6 | 6 | 6 | 131328 | 131328 | 131328 | |
| hamming_9_8 | 6 | 6 | 131328 | 131328 | 131328 | |
| hamming_10_2 | 7 | 7 | 524800 | 524800 | 524800 | |
| copo14 | 73 | 73 | 1834 | 1834 | 1834 | |
| copo23 | 188 | 188 | 8119 | 8119 | 8119 | |
| copos68 | 1576 | 1576 | 209644 | 209644 | 209644 | |
| ThetaPrimeER23_red | 86 | 762 | 777 | 101 | 1712 | Instances from [38] |
| ThetaPrimeER29_red | 104 | 1125 | 1143 | 122 | 2486 | |
| ThetaPrimeER31_red | 110 | 1262 | 1281 | 129 | 2776 | |
| crossing_K_7n | 113 | 577 | 3138 | 113 | 3138 | |
| crossing_K_8n | 479 | 18577 | 72630 | 479 | 72630 | |
| kissing_3_5_5 | 811 | 811 | 3796 | 3796 | 3796 | |
| kissing_4_7_7 | 3723 | 3723 | 19760 | 19760 | 19760 | |

**Table 7.1:** Dimensions of admissible subspaces $\mathcal{S}_{opt}$, $\mathcal{S}_{0/1}$ and $\mathcal{S}_{coord}$ compared with dimensions of the ambient space $\mathcal{S}_{full}$ and $\mathcal{S}_{data}$—the (symmetric part) of the *-algebra generated by $C$ and $\{A_i\}_{i=1}^m$.

| instance | $\mathcal{S}_{opt}$ | $\mathcal{S}_{0/1}$ | $\mathcal{S}_{coord}$ | Orig. |
|---|---|---|---|---|
| ThetaPrimeER23_red | $(3, 2_{12\times}, 1_{44\times})$ | $(27, 25, 5, 1_{44\times})$ | $(27, 25, 5, 1_{59\times})$ | $(57, 1_{59\times})$ |
| ThetaPrimeER29_red | $(3, 2_{15\times}, 1_{53\times})$ | $(33, 31, 5, 1_{53\times})$ | $(33, 31, 5, 1_{71\times})$ | $(69, 1_{71\times})$ |
| ThetaPrimeER31_red | $(3, 2_{16\times}, 1_{56\times})$ | $(35, 33, 5, 1_{56\times})$ | $(35, 33, 5, 1_{75\times})$ | $(73, 1_{75\times})$ |

**Table 7.2:** Tuple of ranks for select examples satisfying the strict inclusions $\mathcal{S}_{opt} \subset \mathcal{S}_{0/1} \subset \mathcal{S}_{coord}$. Here, $s_{t\times}$ means $s$ repeated $t$ times, i.e., $3_{2\times} := (3, 3)$.

where $\mathbf{1}\mathbf{1}^{\mathbf{T}} \in \mathbb{S}^n$ is the all-ones-matrix and $E_{ij}$ is a standard basis matrix of $\mathbb{R}^{n\times n}$. The graphs for these instances are closely related to Hamming graphs, which, for parameters $q$ and $d$, have $2^q$ vertices labeled uniquely by $q$-bit Boolean vectors. For these graphs, vertices are adjacent iff their labels have *Hamming distance* at least $d$. The graphs of `hamming_q_x` and `hamming_q_x_y` are Hamming graphs with certain edges removed; precisely, vertices are adjacent iff the Hamming distance of their labels equals `x` or equals `x` or `y`.

When $G$ is a Hamming graph, it is well known the SDP (7.4) can be converted into a linear program using the theory of association schemes [123]. Here, we find similar simplifications for the instances of [104]; precisely, $\mathbb{S}^n_+ \cap \mathcal{S}_{opt}$ is isomorphic to a nonnegative orthant of order equal to the dimension of $\mathcal{S}_{opt}$, i.e.,

$$\mathbb{S}^n_+ \cap \mathcal{S}_{opt} = \Phi(\mathbb{R}_+^{\dim \mathcal{S}_{opt}})$$

for a Jordan isomorphism $\Phi$.

Note the other automated approach—generating a $*$-algebra from the cost matrix $\mathbf{1}\mathbf{1}^{\mathbf{T}}$ and constraint matrices $I$, $\{E_{ij} + E_{ji}\}_{(i,j)\in E}$ fails completely for these instances; that is, $\mathcal{S}_{data}$ (the symmetric part of this $*$-algebra) equals $\mathbb{S}^n$.

#### Decompositions and majorization

In Table 7.2 we report the tuple of ranks for the subspaces $\mathcal{S}_{opt}$, $\mathcal{S}_{0/1}$ and $\mathcal{S}_{coord}$ for select examples. Specifically, we select examples satisfying the strict inclusions:

$$\mathcal{S}_{opt} \subset \mathcal{S}_{0/1} \subset \mathcal{S}_{coord}.$$

Given these strict inclusions, Theorem 5.4.1 implies the ranks of $\mathcal{S}_{0/1}$ and $\mathcal{S}_{coord}$ weakly majorize those of $\mathcal{S}_{opt}$ in the sense of Definition 5.4.1. Similarly, it implies the ranks of $\mathcal{S}_{coord}$ weakly majorize those of $\mathcal{S}_{0/1}$. This is confirmed in Table 7.2. The first row, for instance, reports the following tuples $r_1 \in \mathbb{Z}^{l_1}$ and $r_2 \in \mathbb{Z}^{l_2}$ for $\mathcal{S}_{opt}$ and $\mathcal{S}_{0/1}$, respectively:

$$r_1 := (3, \underbrace{2, 2, \ldots, 2}_{12\times}, \underbrace{1, 1, \ldots, 1}_{44\times}) \qquad r_2 := (27, 25, 5, \underbrace{1, 1, \ldots, 1}_{44\times}).$$

It easily follows $r_2$ weakly majorizes $r_1$, i.e., for all positive integers $q \in \mathbb{Z}$,

$$\sum_{i=1}^{\min\{q,l_2\}} [r_2]_i \geq \sum_{i=1}^{\min\{q,l_1\}} [r_1]_i.$$

This illustrates the major result of Chapter 5: $\mathcal{S}_{opt}$ is not only optimal with respect to dimension but also its decomposition.

#### An algebra with a complex direct-summand

The example `sosdemo5` is an SDP that bounds a quantity from robust control theory—the structured singular value $\mu(M, \mathbf{\Delta})$ [99]:

$$\mu(M, \mathbf{\Delta}) := \frac{1}{\inf\{\|\Delta\| : \Delta \in \mathbf{\Delta}, \det(I - M\Delta) = 0\}}. \tag{7.5}$$

Here, $M$ is a complex matrix and $\mathbf{\Delta}$ is a set of complex matrices. Though the parameters of $\mu(M, \mathbf{\Delta})$ are complex, one can formulate an SDP with real data matrices to bound $\mu(M, \mathbf{\Delta})$. This is done in `sosdemo5` for particular $M$ and $\mathbf{\Delta}$. Nevertheless, upon decomposing $\mathcal{S}_{opt}$ into a direct-sum of minimal ideals, we find one of the direct-summands is isomorphic to an algebra of complex Hermitian matrices. Precisely, $\mathcal{S}_{opt} = \oplus_{i=1}^{11} \mathcal{S}_i$ for minimal ideals $\mathcal{S}_i$. Letting $r := (\operatorname{rank} S_1, \ldots, \operatorname{rank} S_{11})$ and $d := (\dim S_1, \ldots, \dim S_{11})$, we have

$$\begin{aligned} r &= (1, 1, 1, 1, 4, 4^*, 4, 6, 10, 10, 10) \\ d &= (1, 1, 1, 1, 10, 16^*, 10, 21, 55, 55, 55). \end{aligned}$$

Note with the exception of the entries marked $*$, the relation $d_i = \binom{r_i+1}{2}$ holds, showing $\mathcal{S}_i$ is isomorphic to the algebra of real symmetric matrices of order $r_i$. The exception satisfies $d_i = r_i^2$, showing the corresponding ideal $\mathcal{S}_i$ is isomorphic to the algebra of complex Hermitian matrices of order $r_i$. We remark this is the only example considered where the direct-summands are not all isomorphic to $\mathbb{S}^n$ for some $n$.

### 7.5.2 Comparison with LP method of Grohe, Kersting, Mladenov, and Selman

In [65], Grohe et al. describe a reduction method for *linear* programming (LP) and show it outperforms a symmetry reduction method of [17] on a collection of LPs; indeed, they show their method theoretically subsumes [17]. The linear programs used for comparison are relaxations of integer programs studied in [90]. By treating each linear inequality as a semidefinite constraint of order one, we applied our method to the same LP relaxations. Of the 57 relaxations, we find the same reductions on 56. For the remaining instance (`cov1054sb`), we significantly outperform [65]. For space reasons, Table 7.3 reports results for just a small subset of these LP relaxations. To match [65], we give the number of *dual* variables and inequality constraints. In terms of SDP (7.2) and the SDP (7.3), the number of dual variables and constraints equals the number of linear equations and the sum of the ranks, respectively.

That the method of [65] and ours exhibits similar performance is not surprising. The method of [65] is based on equitable partitions, which, as we discussed in Section 7.4.1, define invariant subspaces of linear maps. This and the empirical evidence of Table 7.3

| | Constraints | | | Variables | | |
|---|---|---|---|---|---|---|
| | Orig. | CR | $\mathcal{S}_{opt}$ | Orig. | CR | $\mathcal{S}_{opt}$ |
| `cov1053` | 252 | 1 | 1 | 679 | 5 | 5 |
| `cov1054` | 252 | 1 | 1 | 889 | 6 | 6 |
| `cov1054sb` | 252 | 252 | 1 | 898 | 898 | 6 |
| `cov1075` | 120 | 1 | 1 | 877 | 7 | 7 |
| `cov1076` | 120 | 1 | 1 | 835 | 7 | 7 |
| `cov1174` | 330 | 1 | 1 | 1221 | 6 | 6 |
| `cov954` | 126 | 1 | 1 | 507 | 6 | 6 |

**Table 7.3:** Dual variables and constraints of original LP, the LP formulated via the color refinement (CR) method of [65], and the LP formulated via restriction to $\mathcal{S}_{opt}$. Columns labeled (CR) use numbers reported in [65].

suggest a projection satisfying the conditions of Constraint Set Invariance and Unitality (Definition 1.4.1) is implicitly constructed in the method of [65]; the instance `cov1054sb` shows this projection isn't always minimum rank.

### 7.5.3 Completely-positive rank, the subspace $\mathcal{S}_{0/1}$, and decomposition trade-offs

Our next example illustrates restrictions to $\mathcal{S}_{0/1}$, the optimal subspace with an orthogonal basis of 0/1 matrices. The considered SDP family yields lower-bounds of *completely-positive rank*, or cp-rank for short. The cp-rank of $W \in \mathbb{S}^n_+$ measures the size of the smallest nonnegative factorization of $W$. Precisely, it is the smallest $r$ for which $V \in \mathbb{R}^{n\times r}_+$ exists satisfying $W = VV^T$. Note cp-rank need not be finite—that is, a nonnegative factorization of $W$ need not exist for any $r$. As shown in [53], the cp-rank of $W \in \mathbb{S}^n$ is lower bounded by the optimal value of the following SDP:

$$
\begin{aligned}
&\text{minimize } t\\
&\text{subject to}\\
&\qquad \begin{pmatrix} t & \operatorname{vect} W^T \\ \operatorname{vect} W & X \end{pmatrix} \succeq 0\\
&\qquad X_{ij,ij} \le W_{ij}^2 \quad \forall i,j \in \{1,\ldots,n\}\\
&\qquad X \preceq W \otimes W\\
&\qquad X_{ij,kl} = X_{il,jk} \quad \forall (1,1) \le (i,j) < (k,l) \le (n,n).
\end{aligned}
$$

Here, $W \otimes W$ denotes the *Kronecker product* and $\operatorname{vect} W$ denotes the $n^2 \times 1$ vector obtained by stacking the columns of $W$. The double subscript $ij$ indexes the $n^2$ rows (or columns) of $X$ and the inequalities on $(i, j)$ hold iff they hold element-wise (see [53] for further clarification on this notation).

In this example, we solve three instances of this SDP taking $W$ equal to the matrices

$Z$, $Z \otimes Z$, and $Z \otimes Z \otimes Z$, where

$$Z = \begin{pmatrix} 4 & 0 & 1 \\ 0 & 4 & 1 \\ 1 & 1 & 3 \end{pmatrix}.$$

Table 7.4 reports computational savings obtained by restricting to $\mathcal{S}_{0/1}$. (In Chapter 2, we applied facial reduction techniques to the same set of SDPs.)

**Alternative reformulation** For these examples, we compare (7.3) against the following alternative reformulation:

$$\begin{array}{ll} \text{minimize} & \langle P_{\mathcal{S}_{0/1}}(C), X\rangle \\ \text{subject to} & \langle P_{\mathcal{S}_{0/1}}(A_i), X\rangle = b_i \quad \forall i \in T \subseteq [m] \\ & X \in \mathbb{S}_+^{n_1} \times \cdots \times \mathbb{S}_+^{n_r}, \end{array} \tag{7.6}$$

where $T \subseteq [m]$ indexes a maximal subset of linearly-independent equations. The dual of (7.6) is

$$\begin{array}{ll} \text{maximize} & \sum_{i \in T} y_i b_i \\ \text{subject to} & P_{\mathcal{S}_{0/1}}(C) - \sum_{i \in T} P_{\mathcal{S}_{0/1}}(A_i) \in \mathbb{S}_+^{n_1} \times \cdots \times \mathbb{S}_+^{n_r}. \end{array}$$

We can interpret this dual SDP as the dual of (7.2) restricted to the subspace $\mathcal{S}_{0/1}$, recalling by Proposition 5.1.1 that $\mathcal{S}_{0/1}$ contains both primal and dual solutions.

Table 7.4 shows solving (7.6) achieves computational savings and, indeed, can be preferred to solving (7.3). As indicated, for the largest instance, we cannot even find the Jordan isomorphism needed to construct (7.3) due to memory constraints. The formulation (7.6) also preserves sparsity.

## ■ 7.5.4 Sparse isomorphisms

Finally, we illustrate how coordinate and 0/1 subspaces can lead to sparse reformulations (as discussed in Section 7.3.2).

### Coordinate subspaces

We find the minimal coordinate subspace $\mathcal{S}_{coord}$ for instances taken from

`http : //www.aem.umn.edu/~AerospaceControl/`,

which build upon the parser SOSOPT [124]. Table 7.5 shows that sparsity is always preserved. Though these examples are of small size, they illustrate $\mathcal{S}_{coord}$ is a proper subspace of $\mathcal{S}_{full}$ for surprisingly many examples.

**Remark 7.5.3.** *Note many of these scripts construct several SDPs; reported results*

| SDP | ranks | num eq | nnz | $t_{pre}$ | $t_{solve}$ |
|---|---|---|---|---|---|
| Orig. (7.2) | $(10, 9, 1_{9\times})$ | 37 | 172 | – | 1.05 |
| Reform. (7.3) | $(5, 4, 2_{4\times}, 1_{6\times})$ | 14 | 798 | 0.30 | 0.37 |
| Reform. (7.6) | $(10, 9, 1_{9\times})$ | 14 | 172 | 0.018 | 0.16 |

**(a)** Instance: $Z$

| SDP | ranks | num eq | nnz | $t_{pre}$ | $t_{solve}$ |
|---|---|---|---|---|---|
| Orig. (7.2) | $(82, 81, 1_{81\times})$ | 2026 | 13204 | – | 39.17 |
| Reform. (7.3) | $(12, 11, 10_{4\times}, 6_{4\times}, 4_{8\times}, 2_{2\times}, 1_{11\times})$ | 167 | 157303 | 8.9 | 2.1 |
| Reform. (7.6) | $(82, 81, 1_{81\times})$ | 167 | 13204 | .11 | 4.35 |

**(b)** Instance: $Z \otimes Z$

| SDP | ranks | num eq | nnz | $t_{pre}$ | $t_{solve}$ |
|---|---|---|---|---|---|
| Orig. (7.2) | $(730, 729, 1_{729\times})$ | 142885 | 1063612 | Out of memory | |
| Reform. (7.3) | Out of memory | | | Out of memory | |
| Reform. (7.6) | $(730, 729, 1_{729\times})$ | 1883 | 1063612 | 7.1 | 2008 |

**(c)** Instance: $Z \otimes Z \otimes Z$

**Table 7.4:** The first row corresponds to the original SDP (7.2) and the others reformulations over $\mathcal{S}_{0/1}$. Here, $t_{pre}$ is time spent (in seconds) finding $\mathcal{S}_{0/1}$ and constructing the reformulation. Solve time $t_{solve}$ is also in seconds.

*are for the* first *SDP constructed.*

#### Zero-one subspaces

Finally, we compare sparsity of $\Phi(A_i)$ for two isomorphisms $\Phi_{rand}$ and $\Phi_{sparse}$ induced by two types of Jordan frames for $\mathcal{S}_{0/1}$. The original SDP instances are from earlier examples of this section and Chapter 2. The map $\Phi_{rand}$ arises from the Jordan frame induced by the spectral decomposition of a randomly sampled element. The map $\Phi_{sparse}$ arises from a sparse Jordan frame obtained by refining diagonal idempotents (Chapter 6.6). Results appear in Table 7.6, along with the section that describes the original SDP instance. (See also Example 6.6.1 for the same comparison on an illustrative example.) As indicated using diagonal idempotents dramatically increases sparsity.

## ■ 7.6 Conclusion

We proposed combinatorial variations of the Jordan reduction methodology introduced in Chapter 5. These variations restrict to subspaces whose bases have low storage complexity and exact combinatorial descriptions (immune to floating-point round-off error)—allowing one to preserve sparsity, accurately store a basis, and trade-off preprocessing effort with the size of the obtained reductions. We also illustrated how these

| | Orig. | | $\mathcal{S}_{coord}$ | |
|---|---|---|---|---|
| | ranks | nnz | ranks | nnz |
| `Chesi(1|4)_IterationWithVlin` | $(9,5)$ | 181 | $(6,3_{2\times},2)$ | 97 |
| `Chesi3_GlobalStability` | $(14,5)$ | 341 | $(8,6,3,2)$ | 193 |
| `Chesi(5|6) _Bootstrap` | $(19,9)$ | 928 | $(13,6_{2\times},3)$ | 520 |
| `Chesi(5|6) _IterationWithVlin` | $(19,9)$ | 928 | $(13,6_{2\times},3)$ | 520 |
| `Coutinho3_IterationWithVlin` | $(9,5)$ | 181 | $(6,3_{2\times},2)$ | 97 |
| `HachichoTibken_Bootstrap` | $(19,9)$ | 685 | $(12,7,6,3)$ | 373 |
| `HachichoTibken_IterationWithVlin` | $(19,9)$ | 685 | $(12,7,6,3)$ | 373 |
| `Hahn_IterationWithVlin` | $(9,5)$ | 156 | $(6,3_{2\times},2)$ | 84 |
| `KuChen_IterationWithVlin` | $(19,9)$ | 928 | $(13,6_{2\times},3)$ | 520 |
| `Parrilo1_GlobalStabilityWithVec` | $(3,2)$ | 20 | $(2,1_{3\times})$ | 14 |
| `Parrilo2_GlobalStabilityWithMat` | $(3,2)$ | 16 | $(2,1_{3\times})$ | 10 |
| `Pendubot_IterationWithVlin` | $(14,4)$ | 372 | $(10,4_{2\times})$ | 292 |
| `VDP_IterationWithVball` | $(5,4)$ | 82 | $(3_{2\times},2,1)$ | 55 |
| `VDP_IterationWithVlin` | $(9,5)$ | 181 | $(6,3_{2\times},2)$ | 97 |
| `VDP_LinearizedLyap` | $(9,5)$ | 156 | $(6,3_{2\times},2)$ | 84 |
| `VDP_MultiplierExample` | $(5,2)$ | 37 | $(3,2,1_{2\times})$ | 23 |
| `VannelliVidyasagar2_Bootstrap` | $(19,9)$ | 928 | $(13,6_{2\times},3)$ | 520 |
| `VannelliVidyasagar2_IterationWithVlin` | $(19,9)$ | 928 | $(13,6_{2\times},3)$ | 520 |
| `VincentGrantham_IterationWithVlin` | $(9,5)$ | 181 | $(6,3_{2\times},2)$ | 97 |
| `WTBenchmark_IterationWithVlin` | $(19,9)$ | 685 | $(13,6_{2\times},3)$ | 385 |

**Table 7.5:** Ranks and number of nonzero (nnz) entries in problem description of original instance and its restriction (7.3) to $\mathcal{S}_{coord}$. The notation $r_{s\times}$ indicates $r$ repeated $s$ times. The table illustrates $\mathcal{S}_{coord}$ has a sparse decomposition—that is the restriction (7.3) is also sparse.

| Instance | nnz A | nnz $A\Phi_{rand}$ | nnz $A\Phi_{sparse}$ | Section |
|---|---|---|---|---|
| `vamos_5_34` | 2704 | 199300 | 4235 | 2.8.5 |
| `copos_1` | 1225 | 3144 | 123 | 2.8.2 |
| `copos_2` | 14400 | 66363 | 492 | 2.8.2 |
| `copos_3` | 81796 | OOM | 1436 | 2.8.2 |
| `copos_4` | 313600 | OOM | 2842 | 2.8.2 |
| `cprank_2` | 13204 | 1914038 | 77879 | 7.5.3 |

**Table 7.6:** Number of nonzero entries (nnz) for indicated maps. OOM indicates construction of map failed due to an out-of-memory error. Descriptions of original SDP instances are given in the indicated section.

bases lead to sparse isomorphisms.

# Part III

# Applications to polynomial optimization

Chapter 8

# Reduction of sum-of-squares programs

A multivariate polynomial $f$ is a sum-of-squares (sos) if

$$f = \sum_{f \in S} f^2$$

for some set of polynomials $S$. The sum-of-squares polynomials in $n$ variables of degree at most $2d$ form a finite-dimensional convex cone, denoted $\Sigma_{n,2d}$. Further, $\Sigma_{n,2d}$ is a linear transformation of the psd cone of order $\binom{n+d}{d}$. Hence, one can solve cone programs formulated over $\Sigma_{n,2d}$—so-called sum-of-squares programs—using semidefinite programming (SDP). For this reason sum-of-squares programs are powerful tools for problems involving polynomial nonnegativity [15, Chapter 3].

This chapter studies partial facial reduction (Chapter 2) and combinatorial Jordan reduction (Chapter 7) of sum-of-squares programs by exploiting their connection with SDP. For partial facial reduction, we show that diagonal approximations of the psd cone induce approximations of $\Sigma^*_{n,2d}$ based on polynomial sparsity (i.e., nonzero coefficients). For Jordan reduction, we characterize admissible coordinate subspaces in terms of polynomial sparsity. These characterizations show existing techniques [85, 35] for simplifying sum-of-squares programs implicitly find such subspaces. These subspaces are not necessarily minimal, and, using these techniques, take exponential time to identify. In constrast, our algorithm from Chapter 7.2.2 finds the minimal coordinate subspace in polynomial time; it also simplifies for the sum-of-squares programs considered.

We organize this chapter as follows. Section 8.1 reviews the basics of polynomial vector spaces. Section 8.2 overviews sum-of-squares polynomials and their connections to semidefinite programming. Sections 8.3 and 8.4 study partial facial reduction and Jordan reduction of sum-of-squares programs, respectively.

## 8.1 Preliminaries

**Vector spaces of polynomials** Let $\mathbb{R}[x]$ denote the ring of polynomials with real coefficients and indeterminates $x_1, \ldots, x_n$. Any finite subset $M$ of $\mathbb{N}^n$ induces a finite-dimensional vector space $\mathbb{R}[x]_M$ of polynomials contained in $\mathbb{R}[x]$. Letting $x^\alpha$ denote the monomial $x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$, this vector space takes the form

$$\mathbb{R}[x]_M := \left\{ \sum_{\alpha \in M} c_\alpha x^\alpha : c_\alpha \in \mathbb{R} \right\}. \tag{8.1}$$

A natural inner product for $\mathbb{R}[x]_M$ is just the dot product between coefficient vectors:

$$\langle f, g \rangle := \sum_{\alpha \in M} c_\alpha d_\alpha,$$

where $f = \sum_{\alpha \in M} c_\alpha x^\alpha$ and $g = \sum_{\alpha \in M} d_\alpha x^\alpha$. We equip $\mathbb{R}[x]_M$ with this inner product and also identify the dual space $\mathbb{R}[x]_M^*$ with $\mathbb{R}[x]_M$.

**Newton polytopes, support, and nonnegativity** The support of $f = \sum_{\alpha \in M} c_\alpha x^\alpha$, denoted $\operatorname{supp}(f)$, is the set of exponents with nonzero coefficients, i.e.,

$$\operatorname{supp}(f) = \{\alpha \in M : c_\alpha \neq 0\}.$$

The *Newton polytope* of $f$, denoted $\operatorname{new}(f) \subseteq \mathbb{R}^n$, is the convex hull of its support (Figure 8.1(a)).

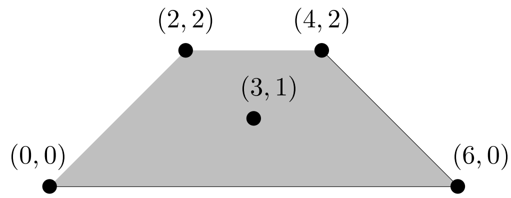
The Newton polytope induces necessary conditions for polynomial nonnegativity. Consider the following theorem of Reznick.

**Proposition 8.1.1** ([119, Section 3])**.** *Suppose $f \in \mathbb{R}[x]_M$ is nonnegative and that $f = \sum_{\alpha \in M} c_\alpha x^\alpha$. If $\alpha$ is a vertex of the Newton polytope* $\operatorname{new}(f)$*, then $c_\alpha \geq 0$.*

The inequalities associated with vertices are illustrated in Figure 8.1(b) for a specific polynomial. If $f$ is univariate, i.e., $f(x) = c_0 + c_1 x + \cdots + c_d x^d$, these inequalities simply state the leading term $c_d$ and constant term $c_0$ of $f$ are nonnegative. Also note that [120, Theorem 3.6] generalizes this proposition from vertices and nonnegative coefficients to arbitrary faces $\mathcal{F}$ and nonnegative *$\mathcal{F}$-restrictions* of $f$—polynomials obtained by setting each coefficient $c_\alpha$ to zero if $\alpha \notin \mathcal{F}$.

## 8.2 Sums-of-squares polynomials

We are interested in nonnegative polynomials that are sums-of-squares. We focus on a subset $\Sigma_M$ of these polynomials induced by a given $M \subseteq \mathbb{N}^n$ which, as we will shortly see, is a linear transformation of the psd cone of order $|M|$. This subset $\Sigma_M$ and corresponding linear transformation $A_M$ are defined as follows.




**(a)** Newton polytope of $f(x)$

$$c_{00} \geq 0, \quad c_{22} \geq 0,$$
$$c_{42} \geq 0, \quad c_{60} \geq 0.$$

**(b)** Valid ineqs. if $f(x) \geq 0$ for all $x$

**Figure 8.1:** $f(x) = c_{00} + c_{31}x_1^3x_2 + c_{22}x_1^2x_2^2 + c_{42}x_1^4x_2^2 + c_{60}x_1^6$

**Definition 8.2.1.** *For a finite set $M \subseteq \mathbb{N}^n$, let $\Sigma_M \subseteq \mathbb{R}[x]_{M+M}$ denote the polynomials that are sums-of-squares of finitely many $f \in \mathbb{R}[x]_M$, i.e., let*

$$\Sigma_M := \left\{ \sum_{f \in S} f^2 : S \subset \mathbb{R}[x]_M, \ \ |S| \text{ is finite} \right\}.$$

*Further, let $A_M : \mathbb{S}^{|M|} \to \mathbb{R}[x]_{M+M}$ denote the unique linear map whose adjoint $A_M^* : \mathbb{R}[x]_{M+M} \to \mathbb{S}^{|M|}$ satisfies*

$$[A_M^* x^\gamma]_{\alpha,\beta} = \begin{cases} 1 & \text{if } \alpha + \beta = \gamma, \\ 0 & \text{otherwise,} \end{cases} \qquad \forall \gamma \in M + M,$$

*where the rows and columns of $A_M^* x^\gamma$ are indexed by $M$.*

That $\Sigma_M$ equals the image of a psd cone under $A_M$ is a special case of a result of Nesterov. This result also yields a description of the dual cone $\Sigma_M^*$. Formally:

**Proposition 8.2.1** (Special case of Theorem 17.1 of [98])**.** *For a finite subset $M$ of $\mathbb{N}^n$, the cone $\Sigma_M$ and linear map $A_M : \mathbb{S}^{|M|} \to \mathbb{R}[x]_{M+M}$ (Definition 8.2.1) satisfy*

- $\Sigma_M = A_M \cdot \mathbb{S}_+^{|M|}$, *where* $A_M \cdot \mathbb{S}_+^{|M|} := \left\{ A_M(X) : X \in \mathbb{S}_+^{|M|} \right\}$;
- $\Sigma_M^* = \left\{ y \in \mathbb{R}[x]_{M+M} : A_M^* y \in \mathbb{S}_+^{|M|} \right\}$.

*Further, $\Sigma_M$ is closed and convex.*

As a practical consequence, one can check membership in $\Sigma_M$ by solving a semidefinite program. Specifically, $\Sigma_M$ contains the polynomial $f \in \mathbb{R}[x]_{M+M}$ if and only if there exists a symmetric matrix $Q \in \mathbb{S}^{|M|}$ that solves

$$\begin{array}{ll} \text{Find} & Q \in \mathbb{S}_+^{|M|} \\ \text{subject to} & A_M(Q) = f. \end{array} \qquad \text{(SOS-SDP)}$$

To see that existence of a solution implies that $f \in \Sigma_M$, note that $Q \in \mathbb{S}^{|M|}$ satisfies $A_M(Q) = f$ if and only if

$$f = \sum_{\alpha \in M} \sum_{\beta \in M} Q_{\alpha\beta} x^\alpha x^\beta$$

by definition of the map $A_M : \mathbb{S}^{|M|} \to \mathbb{R}[x]_{M+M}$. In addition, any $Q \in \mathbb{S}^{|M|}_+$ has a factorization $Q = L^T L$. Hence, if $Q$ solves (SOS-SDP), then

$$f = \sum_{\alpha \in M} \sum_{\beta \in M} [L^T L]_{\alpha\beta} x^\alpha x^\beta = (Lx_M)^T L x_M,$$

where $x_M$ is a vector of monomials indexed by $M$; specifically, $[x_M]_\alpha = x^\alpha$. This shows that $f$ equals the sum of the squared entries of $Lx_M$, a polynomial vector. Conversely, if $f = \sum_{i=1}^p f_i^2$ for $f_i = \sum_{\alpha \in M} c^i_\alpha x^\alpha$ and $c^i \in \mathbb{R}^{|M|}$, then the matrix

$$Q = \sum_{i=1}^p c^i (c^i)^T$$

solves (SOS-SDP).

**Remark 8.2.1.** *The set $M \subseteq \mathbb{N}^n$ in* (SOS-SDP) *is usually picked* from *the polynomial $f$ with the following guarantee: if $f$ is a sum-of-squares (of polynomials in any set), then $f \in \Sigma_M$; see, e.g., [119, Theorem 1], [15, Chapter 3] and [80]. See also [16, Section 6] for further study of $\Sigma_{\mathrm{conv}(M) \cap \mathbb{N}^n}$ and its relationship with the nonnegative polynomials with support contained in* conv$(2M)$.

## ■ 8.3 Partial facial reduction

Given a cone program with feasible set $\mathcal{A} \cap \mathcal{K}$, the facial reduction algorithm (Algorithm 1.1) finds a hyperplane containing the affine set $\mathcal{A}$ that exposes a face of the cone $\mathcal{K}$. To find a hyperplane, the algorithm solves an auxiliary cone program over $\mathcal{K}^*$, which can be expensive to solve. To reduce the cost of solving this auxiliary problem, we proposed replacing $\mathcal{K}^*$ with a computationally efficient inner approximation—a methodology we called partial facial reduction (Chapter 2).

In this section, we apply partial facial reduction to (SOS-SDP). For this SDP, the affine set $\mathcal{A}$ is the solution set of $A_M(Q) = f$ for some given set of monomial exponents $M$ and polynomial $f \in \mathbb{R}[x]_{M+M}$. Further, $\mathcal{K}$ is the psd cone $\mathbb{S}^{|M|}_+$, which is self dual, i.e., $\mathbb{S}^{|M|}_+ = (\mathbb{S}^{|M|}_+)^*$. Specifically, we study when the auxiliary problem of (SOS-SDP), given by

$$\begin{array}{ll} \text{Find} & (S, y) \in \mathbb{S}^{|M|}_+ \times \mathbb{R}[x]_{M+M} \\ \text{subject to} & S = A_M^* y, \langle f, y \rangle = 0, \end{array} \qquad \text{(SOS-SDP-AUX)}$$

has a nonzero solution $(S, y)$ in $\mathcal{D}^{|M|} \times \mathbb{R}[x]_{M+M}$, where $\mathcal{D}^{|M|}$ denotes the nonnegative diagonal matrices of order $|M|$. (Note that $\mathcal{D}^{|M|}$ inner approximates $\mathbb{S}_+^{|M|}$, i.e., $\mathcal{D}^{|M|} \subset \mathbb{S}_+^{|M|}$.) Results of this type first appeared in [139].

## ■ 8.3.1 Nonnegativity of coefficients

Suppose $\sum_{\alpha \in M+M} c_\alpha x^\alpha$ is nonnegative. Then, $c_\alpha \geq 0$ for all vertices $\alpha$ of $\operatorname{conv}(M)$ by Proposition 8.1.1. Hence, the vertices of $\operatorname{conv}(M)$ induce a polyhedral approximation of $\Sigma_M^*$. It turns out, and its not hard to show, that this polyhedral approximation is contained in the polyhedral approximation induced by diagonal matrices $\mathcal{D}^{|M|}$. To be precise, let $\mathcal{P}^*_{M,\text{new}}$ and $\Sigma^*_{M,\mathcal{D}}$ denote these two approximations, i.e.,

$$\mathcal{P}^*_{M,\text{new}} := \left\{ \sum_{\gamma \in \operatorname{ext}(M+M)} \lambda_\gamma x^\gamma : \lambda_\gamma \geq 0 \right\}, \qquad \Sigma^*_{M,\mathcal{D}} := \left\{ y \in \Sigma^*_M : A^*_M y \in \mathcal{D}^{|M|} \right\}.$$

The following establishes the mentioned containment $\mathcal{P}^*_{M,\text{new}} \subseteq \Sigma^*_{M,\mathcal{D}}$.

**Lemma 8.3.1.** *Let $\gamma$ be an extreme point of* $\operatorname{conv}(M+M)$*. Then $A^*_M x^\gamma \in \mathcal{D}^{|M|}$. Hence, $\mathcal{P}^*_{M,\text{new}} \subseteq \Sigma^*_{M,\mathcal{D}}$.*

*Proof.* We only need to show that $A^*_M x^\gamma$ is diagonal if $\gamma$ is an extreme point of $\operatorname{conv}(M+M)$. By definition, $\operatorname{supp}(A^*_M x^\gamma) = \{(\alpha, \beta) \in M \times M : \alpha + \beta = \gamma\}$. Suppose $\gamma$ is an extreme point and $A^*_M x^\gamma$ is not diagonal. Then by definition of $A_M$, there exists $\alpha \neq \beta$ for which $\alpha + \beta = \gamma$, implying that $\frac{1}{2}\gamma = \frac{1}{2}(\alpha + \beta)$. Since $\frac{1}{2}\operatorname{conv}(M+M)$ contains $\alpha$, $\beta$ and $\frac{1}{2}\gamma$, this shows that $\frac{1}{2}\gamma$ is *not* an extreme point of $\frac{1}{2}\operatorname{conv}(M+M)$, a contradiction. □

It turns out that the inclusion $\mathcal{P}^*_{M,\text{new}} \subseteq \Sigma^*_{M,\mathcal{D}}$ can be strict. To explain, we define the set $M^+ \subseteq M$ as follows.

**Definition 8.3.1.** *For a subset $M$ of $\mathbb{N}^n$, let $M^+$ be the subset of points that cannot be written as the midpoint of distinct points in $M$, i.e.,*

$$M^+ := M \setminus \left\{ \frac{\alpha + \beta}{2} : \alpha, \beta \in M, \alpha \neq \beta \right\}.$$

As shown next, $2M^+$ is precisely the set of exponents $\gamma$ for which $A^*_M x^\gamma$ is contained in $\mathcal{D}^{|M|}$. Hence, $2M^+$ contains the extreme points of $\operatorname{conv}(M+M)$.

**Lemma 8.3.2.** $A^*_M x^\gamma \in \mathcal{D}^{|M|}$ *if and only if $\gamma = 2\zeta$ for $\zeta \in M^+$.*

*Proof.* Consider $\zeta \in M^+$ and suppose that $(\alpha, \beta) \in \operatorname{supp} A^*_M x^{2\zeta}$. Then $\zeta = \frac{1}{2}(\alpha + \beta)$; hence, $\alpha = \beta$ by definition of $M^+$. Conversely, if $A^*_M x^\gamma \in \mathcal{D}^{|M|}$, then $\alpha + \beta = \gamma$ implies that $\alpha = \beta$. Hence, $\gamma = 2\alpha$. Suppose that $\alpha \notin M^+$. Then, $\alpha = \frac{\mu+\lambda}{2}$ for $\mu \neq \lambda$. But

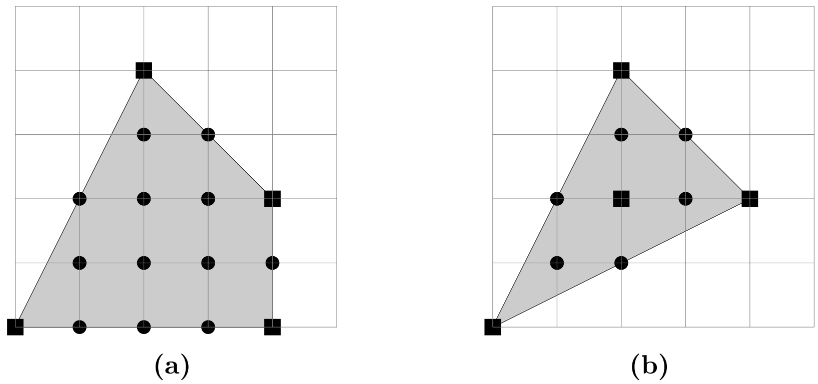


**Figure 8.2:** Convex hulls of $(M + M)$ for two choices of $M$. The set $2M^+$ is marked with rectangles. The inclusion $2M^+ \subseteq 2M$ is not strict in Figure 8.2(a), where $M = \{(0,0),(1,0),(2,0),(1,1),(1,2),(2,1)\}$. It is strict in Figure 8.2(b), where $M = \{(0,0),(1,1),(1,2),(2,1)\}$.

$\mu + \lambda = 2\alpha = \gamma$. Hence, $(\mu,\lambda) \in \operatorname{supp} A_M^* x^\gamma$, contradicting the assumption $A_M^* x^\gamma$ is diagonal. □

The inclusion $\mathcal{P}^*_{M,\text{new}} \subseteq \Sigma^*_{M,\mathcal{D}}$ is strict whenever $2M^+$ is a strict superset of the extreme points of $\operatorname{conv}(M+M)$. Further, when $2M^+$ is a strict superset, $\Sigma^*_{M,\mathcal{D}}$ has more extreme rays than $\mathcal{P}^*_{M,\text{new}}$ and hence strictly contains it. As an example, this containment is strict when

$$M = \{(0,0),(1,1),(1,2),(2,1)\}.$$

For this choice of $M$, we have that $M = M^+$. Hence, $2M^+$ contains $(2,2)$, which is not an extreme point of $\operatorname{conv}(M+M)$. This is illustrated by Figure 8.2(b).

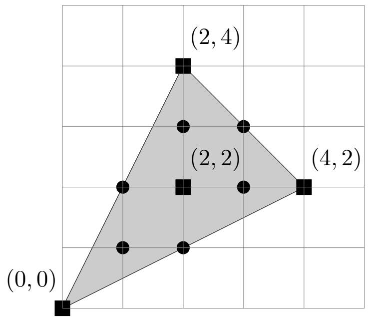
## ■ 8.3.2 Inequalities violated by nonnegative polynomials

For $y \in \mathcal{P}^*_{M,\text{new}}$, it holds that $\langle f, y\rangle \geq 0$ for all nonnegative polynomials $f \in \mathbb{R}[x]_{M+M}$ (Proposition 8.1.1). The same is not necessarily true for $y \in \Sigma^*_{M,\mathcal{D}}$: an inequality $\langle f, y\rangle \geq 0$ induced by some $y \in \Sigma^*_{M,\mathcal{D}}$ may be violated by some nonnegative polynomial $f \in \mathbb{R}[x]_{M+M}$, as illustrated by Figure 8.3.

To see this, note for $M = \{(0,0),(1,1),(1,2),(2,1)\}$ that $(1,1) \in M^+$; hence, $x_1^2x_2^2 \in \Sigma^*_{M,\mathcal{D}}$ by Lemma 8.3.2. However, the nonnegative polynomial $g \in \mathbb{R}[x]_{M+M}$ given by

$$g = x_1^4x_2^2 + x_1^2x_2^4 - 3x_1^2x_2^2 + 1$$

(the so-called Motzkin polynomial) satisfies $\langle g, x_1^2x_2^2\rangle = -3$. Note that this implies $g \notin \Sigma_M$—i.e., that $g$ is not a sum of squares of polynomials with support contained by




**(a)** conv$(M+M)$

$$c_{00} \geq 0, c_{24} \geq 0,$$
$$c_{42} \geq 0,$$

**(b)** Valid ineqs. if $f(x) \geq 0$

$$c_{00} \geq 0, c_{24} \geq 0,$$
$$c_{42} \geq 0, c_{22} \geq 0,$$

**(c)** Valid ineqs. if $f(x) \in \Sigma_M$

**Figure 8.3:** $f(x) = \sum_{\alpha \in M+M} c_\alpha x^\alpha$, where $2M^+ \subseteq 2M$ is marked with rectangles. Here, $M$ is the same as in Figure 8.2(b).

this particular $M$. In fact, it is well known that $g \notin \Sigma_M$ for any choice of $M$.

## ■ 8.4 Jordan reduction

Recall the main idea in Jordan reduction: finding a subspace $\mathcal{S}$ that is *admissible.* For (SOS-SDP), this means the orthogonal projection $P_\mathcal{S} : \mathbb{S}^{|M|} \to \mathbb{S}^{|M|}$ onto $\mathcal{S}$ satisfies the conditions

$$P_\mathcal{S} \cdot \mathbb{S}^{|M|}_+ \subseteq \mathbb{S}^{|M|}_+, \qquad P_\mathcal{S} \cdot \mathcal{A} \subseteq \mathcal{A},$$

which together imply that $\mathcal{S}$ contains solutions to (SOS-SDP) when they exist. In this section, we characterize admissible *coordinate* subspaces (Chapter 7) for (SOS-SDP) in terms of the exponents $M$ and the polynomial $f$. For this, we let $\mathcal{S}_\mathcal{R} \subseteq \mathbb{S}^{|M|}$ denote the coordinate subspace of matrices with supports contained in $\mathcal{R} \subseteq M \times M$, i.e.,

$$\mathcal{S}_\mathcal{R} := \operatorname{span}\left\{E_{\alpha\beta} + E_{\beta\alpha} : (\beta, \alpha) \in \mathcal{R}\right\},$$

where $E_{\alpha\beta} \in \mathbb{R}^{|M|\times|M|}$ is the 0/1 matrix with support equal to $(\alpha, \beta)$. We will show that $\mathcal{S}_\mathcal{R}$ is admissible only if $\mathcal{R}$ is a relation of the following type.

**Definition 8.4.1.** *Let $M$ be a finite subset of $\mathbb{N}^n$. For $\mathcal{T} \subseteq M+M$, define the relation $R_\mathcal{T} \subseteq M \times M$ as follows*

$$R_\mathcal{T} = \left\{(\alpha, \beta) \in M \times M : \alpha + \beta \in \mathcal{T}\right\}.$$

*Equivalently, let $R_\mathcal{T} := \{\text{supp}(A^*_M x^\gamma) : \gamma \in \mathcal{T}\}$.*

Our main result is the following characterization.

**Theorem 8.4.1** (Main result)**.** *The subspace* $\mathcal{S}_{\mathcal{R}}$ *is admissible for* (SOS-SDP) *if and only if there is a subset* $\mathcal{T}$ *of* $M+M$ *such that* $\mathcal{R} = \mathcal{R}_{\mathcal{T}}$*, where* $\mathcal{T}$ *satisfies the following properties:*

- $\mathcal{R}_{\mathcal{T}}$ *is transitive*
- $\mathcal{T}$ *contains the support of* $f$*.*

The next example illustrates this theorem for a fixed $M$ and different polynomials $f$.

**Example 8.4.1.** *Consider* $M = \{(0,0),(1,1),(1,2),(2,1)\}$ *and let*

$$f = \sum_{(\alpha,\beta)\in(M+M)} c_{\alpha\beta} x_1^{\alpha} x_2^{\beta}.$$

*The matrix* $A_M^*(f)$ *and two polynomials in* $\mathbb{R}[x]_{M+M}$ *are given by:*

$$A_M^*(f) = \begin{bmatrix} c_{00} & c_{11} & c_{12} & c_{21} \\ c_{11} & c_{22} & c_{23} & c_{32} \\ c_{12} & c_{23} & c_{24} & c_{33} \\ c_{21} & c_{32} & c_{33} & c_{42} \end{bmatrix}, \qquad \begin{aligned} g &= 1 + x_1^2x_2^2 + x_1^2x_2^4 + x_1^4x_2^2 \\ h &= 1 + x_1^3x_2^2 + x_1^4x_2^2 + x_1^3x_2^3 + x_1^4x_2^2 \end{aligned}$$

*A subset* $\mathcal{T} \subseteq M+M$ *and the coordinate subspace* $\mathcal{S}_{R_{\mathcal{T}}}$ *it induces are*

$$\mathcal{T} = \left\{ \begin{array}{c} (0,0),(2,2),(2,3), \\ (3,2),(2,4),(3,3) \\ (4,2) \end{array} \right\}, \qquad \mathcal{S}_{R_{\mathcal{T}}} = \begin{bmatrix} * & 0 & 0 & 0 \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{bmatrix}.$$

*The subspace is admissible if* $f = g$ *or* $f = h$*. It is minimal if* $f = h$*. Another subset* $\mathcal{T} \subseteq M+M$ *and induced coordinate subspace* $\mathcal{S}_{R_{\mathcal{T}}}$ *is*

$$\mathcal{T} = \left\{ \begin{array}{c} (0,0),(2,2), \\ (2,4),(4,2) \end{array} \right\}, \qquad \mathcal{S}_{R_{\mathcal{T}}} = \begin{bmatrix} * & 0 & 0 & 0 \\ 0 & * & 0 & 0 \\ 0 & 0 & * & * \\ 0 & 0 & * & * \end{bmatrix}.$$

*This subspace is admissible and minimal if* $f = g$*. It is not admissible for* $f = h$ *since* $\mathcal{T}$ *doesn't contain the support of* $h$*.*

Theorem 8.4.1 yields a procedure (Algorithm 8.1) for finding the minimal coordinate subspace $\mathcal{S}_{\mathcal{R}_{\mathcal{T}}}$. Note this algorithm is expressed solely in terms of the polynomial $f$. It also performs at most $|M+M|$ iterations, each with complexity polynomial in the cardinality of $M$. In the remainder of this section, we prove Theorem 8.4.1. We then interpret other techniques for simplifying (SOS-SDP) as less powerful and less efficient versions of Algorithm 8.1.

**Algorithm 8.1:** Finds minimal admissible coord. subspace $\mathcal{S}_{\mathcal{R}_\mathcal{T}}$ for (SOS-SDP)

**Inputs:** a polynomial $f \in \mathbb{R}[x]_{M+M}$.
**Output:** relation $R_\mathcal{T}$ inducing the minimal coord. subspace $\mathcal{S}_{R_\mathcal{T}}$ of (SOS-SDP)
Initialize $\mathcal{T}$ to $\text{supp}(f)$
**repeat**
| $\mathcal{T} \leftarrow \{\alpha + \beta : (\alpha, \beta) \in$ the transitive closure of $R_\mathcal{T}\}$
**until** converged

## 8.4.1 Proof of Theorem 8.4.1

From Chapter 5, the subspace $\mathcal{S}_\mathcal{R}$ is admissible if and only if

- $\mathcal{S}_\mathcal{R}$ is invariant under $X \mapsto X^2$.
- $\mathcal{S}_\mathcal{R}$ contains $A_M^*(A_M A_M^*)^{-1} f$
- $\mathcal{S}_\mathcal{R}$ is an invariant subspace of $A_M^*(A_M A_M^*)^{-1} A_M$.

Note in Chapter 7, we saw that the first condition holds if and only if $\mathcal{R}$ is transitive. In this section we prove the following theorem by additionally characterizing the second and third conditions using special sparsity properties of the linear map $A_M$.

The proof uses sparsity properties of $A_M^*$ which, by definition, satisfies

$$\text{supp}(A_M^* x^\gamma) = \{(\alpha, \beta) \in M \times M : \beta + \alpha = \gamma\}.$$

This shows that $A_M^*$ maps polynomials with disjoint (resp., equal) support to matrices with disjoint (resp., equal) support. We first use these properties to show the following.

**Lemma 8.4.1.** *For all $g, f \in \mathbb{R}_{M+M}[x]$, the following statements hold.*

1. *The polynomials $f$ and $g$ have disjoint support if and only if the matrices $A_M^* g$ and $A_M^* f$ have disjoint support.*
2. *The polynomials $f$ and $g$ have equal support if and only if the matrices $A_M^* g$ and $A_M^* f$ have equal support.*
3. *If $A_M X \neq 0$ and $\text{supp}(X) \subseteq \text{supp}(A_M^* x^\alpha)$, then $\text{supp}(A_M X) = \text{supp}(x^\alpha)$.*
4. $\text{supp}(A_M A_M^* f) = \text{supp}(f)$ *for all $f$.*

*Proof.* The first two statements are immediate from the definition of $A_M^*$. For the third, we need to show $A_M X$ and $x^\alpha$ have equal support. For this, we note that for all $\beta \neq \alpha$,

$$\langle x^\beta, A_M X \rangle = \langle A_M^* x^\beta, X \rangle = 0$$

given that $A_M^* x^\beta$ and $X$ have disjoint support if $\mathrm{supp}(X) \subseteq \mathrm{supp}(A_M^* x^\alpha)$. This shows that $A_M X = \lambda x^\alpha$ for some $\lambda \in \mathbb{R}$. Further, $\lambda \neq 0$ since $A_M X \neq 0$ by assumption.

We now show the last statement. By the first statement,

$$\mathrm{supp}(A_M^* f) = \cup_{\alpha \in \mathrm{supp}(f)} \mathrm{supp}(A_M^* x^\alpha).$$

Further,

$$\mathrm{supp}(A_M A_M^* f) = \cup_{\alpha \in \mathrm{supp}(f)} \mathrm{supp}(A_M A_M^* x^\alpha).$$

Since $A_M A_M^*$ is invertible, $A_M A_M^* x^\alpha \neq 0$. Hence, by the third statement, $\mathrm{supp}(A_M A_M^* x^\alpha) = \mathrm{supp}(x^\alpha)$, showing that

$$\mathrm{supp}(A_M A_M^* f) = \cup_{\alpha \in \mathrm{supp}(f)} \mathrm{supp}(x^\alpha) = \mathrm{supp}(f).$$

□

We use this to prove the following which, combined with Lemma 7.3.1 (which established a correspondence between transitive relations and coordinate projections that leave the psd cone invariant), proves Theorem 8.4.1.

**Lemma 8.4.2.** *The following statements hold*

- *The coordinate subspace $\mathcal{S}_\mathcal{R}$ contains $A_M^*(A_M A_M^*)^{-1} f$ if and only if the relation $\mathcal{R}$ contains the support of $A_M^* f$.*
- *The coordinate subspace $\mathcal{S}_\mathcal{R}$ is an invariant subspace of $A_M^*(A_M A_M^*)^{-1} A_M$ if and only if there exists a subset $\mathcal{T} \subseteq M + M$ for which*

$$\mathcal{R} = \bigcup_{\gamma \in \mathcal{T}} \mathrm{supp}(A_M^* x^\gamma)$$

*Proof.* By definition, the subspace $\mathcal{S}_\mathcal{R}$ contains a point $X$ if and only if the relation $\mathcal{R}$ contains the support of $X$. By Lemma 8.4.1-(4), the polynomials $f$ and $(A_M A_M^*)^{-1} f$ have equal support. Hence, by Lemma 8.4.1-(2), the matrices $A_M^*(A_M A_M^*)^{-1} f$ and $A_M^* f$ have equal support. Hence, $\mathcal{S}_\mathcal{R}$ contains $A_M^*(A_M A_M^*)^{-1} f$ if and only if $\mathcal{R}$ contains the support of $A_M^* f$.

Let $\Psi$ denote $A_M^*(A_M A_M^*)^{-1} A_M$ and let

$$\mathcal{B} = \{E_{\alpha\beta} + E_{\beta\alpha} : \alpha, \beta \in M\}.$$

Finally, let $G$ be the graph with node set $\mathcal{B}$ for which $X, Y \in \mathcal{B}$ are adjacent if and only if $\langle X, \Psi(Y) \rangle \neq 0$. Finally, let $B_{\alpha\beta} = E_{\alpha\beta} + E_{\beta\alpha}$. Let $S_\gamma := \{B_{\alpha\beta} : (\alpha, \beta) \in \mathrm{supp}(A_M^* x^\gamma)\}$. By Lemma 7.4.2, statement 2 follows if the collection of sets $\{S_\gamma : \gamma \in M + M\}$

partitions $\mathcal{B}$ into connected components of $G$. For this, it suffices to show that $\Psi(B_{\alpha\beta})$ and $A_M^* x^\gamma$ have equal support if $\alpha+\beta=\gamma$. To see they have equal support, note that $\operatorname{supp}(B_{\alpha\beta}) \subseteq \operatorname{supp}(A_M^* x^\gamma)$ by definition of $A_M$. Further,

$$\operatorname{supp}\left((A_M A_M^*)^{-1} A_M(B_{\alpha\beta})\right) = \operatorname{supp}\left(A_M(B_{\alpha\beta})\right) = \operatorname{supp}(x^\gamma)$$

where the first equality follows by Lemma 8.4.1-(4) and the next by Lemma 8.4.1-(3). Hence, $\Psi(B_{\alpha\beta})$ and $A_M^* x^\gamma$ have equal support by Lemma 8.4.1-(2).

□

## ■ 8.4.2 Characterization of transitivity and comparisons

By Theorem 8.4.1, the coordinate subspace $\mathcal{S}_{\mathcal{R}_\mathcal{T}}$ is admissible for (SOS-SDP) if and only if $\mathcal{T} \subseteq (M+M)$ contains the support of the polynomial $f$ and $R_\mathcal{T} \subseteq M \times M$ is transitive. We now study transitivity of $R_\mathcal{T}$ in more detail. First we describe the partition of $M$ induced by $R_\mathcal{T}$ when it is transitive. We then show how methods from the literature for simplying (SOS-SDP) implicitly construct a transitive relation $R_\mathcal{T}$ and hence can be viewed as weaker versions of Algorithm 8.1; incidentally, these algorithms run in exponential time (in $n$) whereas Algorithm 8.1 runs in polynomial time.

**A characterization of transitivity** When the symmetric relation $R_\mathcal{T} \subseteq M \times M$ is transitive, it defines a collection of disjoint subsets of $M$ (and a partition of $M$ when it is also reflexive). It turns out one can decompose $\mathcal{T}$ into unions of Minkowski sums using this collection. Moreover, existence of this decomposition implies transitivity. Formally:

**Lemma 8.4.3.** *Let $M$ be a finite subset of $\mathbb{N}^n$. For $\mathcal{T} \subseteq M+M$, the following statements are equivalent.*

1. *The relation $R_\mathcal{T}$ is transitive.*
2. *$\mathcal{T} = \bigcup_{i=1}^p (S_i + S_i)$, where $S_0, S_1, S_2, \ldots, S_p$ form a partition of $M$ and satisfy*

$$\begin{aligned} \mathcal{T} \cap (S_i + S_j) &= \emptyset \qquad \forall i \neq j, \\ \mathcal{T} \cap (S_0 + S_0) &= \emptyset. \end{aligned} \tag{8.2}$$

*Proof.* $(2 \Rightarrow 1)$: Suppose $\gamma \in S_i$ and $\beta \in S_j$ for $i, j \in [0, p]$ and that $(\beta, \gamma) \in R_\mathcal{T}$, i.e, $\beta + \gamma \in \mathcal{T}$. By (8.2), we conclude $i = j$ and $i \neq 0$. By the exact same argument, if $(\gamma, \mu) \in R_\mathcal{T}$ for $\mu \in S_k$, then $S_j = S_k$. Hence, $\beta + \mu \in S_i + S_i \subseteq \mathcal{T}$, showing that $(\beta, \mu) \in R_\mathcal{T}$.

$(1 \Rightarrow 2)$. Suppose $R_\mathcal{T}$ is transitive. Then, there exists disjoint subsets $S_0, S_1, \ldots, S_p$ of $M$ for which

$$M = \cup_{i=0}^p S_i,$$

where $S_1, \dots, S_p$ form a partition of $\bar{M} := \{\beta : (\beta, \beta) \in R_{\mathcal{T}}\}$, $S_0 = M \setminus \bar{M}$ and $\beta, \gamma \in S_i$ if and only if $(\beta, \gamma) \in R_{\mathcal{T}}$ for all $i \in [1, p]$. By definition of $R_{\mathcal{T}}$, it follows that $\mathcal{T} = \cup_{i=1}^p S_i + S_i$.

If $\mathcal{T} \cap (S_0 + S_0) \neq \emptyset$, there exists $\beta, \gamma \in S_0$ such that $(\beta, \gamma) \in R_{\mathcal{T}}$, which, using the fact $R_{\mathcal{T}}$ is symmetric and transitive, implies $(\beta, \beta) \in R_{\mathcal{T}}$, a contradiction of the definition of $S_0$. Similarly, $\mathcal{T} \cap (S_i + S_j) \neq \emptyset$ cannot hold unless $i = j$ by definition of $S_i$.

□

Transitivity of $R_{\mathcal{T}}$ also implies existence of structured sums-of-squares decompositions. Specifically, if a polynomial $f$ with $\operatorname{supp}(f) \subseteq \mathcal{T}$ is a sum-of-squares of polynomials supported by $M$, then $f$ is also a sum-of-squares of polynomials supported by the subsets $S_i$ of Lemma 8.4.3. Formally:

**Corollary 8.4.1.** *Let $M$ be a finite subset of $\mathbb{N}^n$, $\mathcal{T}$ a finite subset of $M + M$ and suppose $R_{\mathcal{T}}$ is transitive. Finally, for $f \in \mathbb{R}[x]_{M+M}$ suppose $\operatorname{supp}(f) \subseteq \mathcal{T}$. The following statements are equivalent.*

1. *$f$ is a sum-of-squares of polynomials supported in $M$, i.e.*

$$f = \sum_i f_i^2, \qquad \operatorname{supp}(f_i) \subseteq M.$$

2. *$f$ is a sum-of-squares of polynomials supported in $S_j$, i.e.*

$$f = \sum_{j=1}^p \sum_i f_{i,j}^2 \qquad \operatorname{supp}(f_{i,j}) \subseteq S_j,$$

*where $S_0, \dots, S_p$ are disjoint subsets of $M$, satisfying (8.2), for which $\mathcal{T} = \cup_{j=1}^p S_j + S_j$.*

#### Comparison with other methods

Since $R_{\mathcal{T}}$ is transitive, the associated coordinate subspace is block diagonal up to permutation, where the blocks correspond to the partition of $|M|$ induced by $R_{\mathcal{T}}$. We next show how other block-diagonalization strategies in the literature implicitly construct transitive relations $R_{\mathcal{T}}$ and hence admissible coordinate subspaces (Theorem 8.4.1). Methods of [35] are based on Newton-polytope arguments and a generalization, whereas a method of [85] is based on polynomials with sign-symmetries, which we discuss first.

**Transitive relations from sign-symmetries** Löfberg [85] shows (SOS-SDP) can be block-diagonalized by identifying its *sign-symmetries*, where $f \in \mathbb{R}[x]_{M+M}$ has a sign-symmetry if jointly flipping the sign of a subset of indeterminates leaves $f$ invariant. For instance, if $f$ is in three variables, $f$ has a sign-symmetry if, e.g., $f(x_1, x_2, x_3) = f(x_1, -x_2, -x_3)$, $f(x_1, x_2, x_3) = f(-x_1, x_2, x_3)$, or $f(x_1, x_2, x_3) = f(-x_1, -x_2, -x_3)$. For a polynomial in $n$ variables, there are $2^n - 1$ possible sign-symmetries, each corresponding to a nonempty subset of $\{1, \ldots, n\}$. In three variables, the other symmetries are:

$$\begin{aligned} f(x_1, x_2, x_3) &= f(x_1, x_2, -x_3), & f(x_1, x_2, x_3) &= f(-x_1, x_2, -x_3), \\ f(x_1, x_2, x_3) &= f(x_1, -x_2, x_3), & f(x_1, x_2, x_3) &= f(-x_1, -x_2, x_3). \end{aligned}$$

To identify every sign-symmetry, Löfberg [85] uses a set of binary vectors that label sign flips, e.g., $f(x_1, x_2, x_3) = f(-x_1, x_2, x_3)$ is labeled by the vector $r = (1, 0, 0)^T$. He then exploits the fact $f$ has a sign-symmetry associated with $r \in \{0, 1\}^n$ if $r^T \gamma$ is an even number for all $\gamma \in \text{supp}(f)$. The next lemma shows the binary vectors identifying sign-symmetries also define a transitive relation:

**Lemma 8.4.4.** *Let $M$ be a finite subset of $\mathbb{N}^n$ and let $r_1, \ldots, r_p \in \{0, 1\}^n$ be a set of non-zero binary vectors. If $\mathcal{T} = \{\alpha \in M + M : r_i^T \alpha \in 2\mathbb{N}$ for all $i \in [p]\}$, the relation $R_{\mathcal{T}} := \{(\alpha, \beta) \in M \times M : \alpha + \beta \in \mathcal{T}\}$ is transitive.*

*Proof.* If $\beta + \gamma$ and $\gamma + \mu$ are in $\mathcal{T}$, then $r_i^T(\beta + \gamma)$ and $r_i^T(\gamma + \mu)$ are even integers. We conclude if $r_i^T \gamma$ is odd (resp., even), then $r_i^T \beta$ and $r_i^T \mu$ are both odd (resp., even). Hence, $r_i^T(\beta + \mu)$ is even, showing that $\beta + \mu \in \mathcal{T}$. □

**Transitive relations from Newton polytopes** Fix $f \in \mathbb{R}[x]_{M+M}$ and recall that the Newton polytope new$(f)$ of $f$ is the convex hull of supp$(f)$ is called the Newton polytope of $f$, which we denote new$(f)$. In [35], the authors show that if $f$ is supported on disjoint faces of new$(f)$, then $f$ is a sum-of-squares if and only if its *restriction* to each face is a sum-of-squares. In other words, if

$$f = \sum_{i=1}^{p} \sum_{\alpha \in \mathcal{F}_i} b_\alpha x^\alpha$$

for pairwise-disjoint faces $\mathcal{F}_i$ of new$(f)$, then $f$ is a sum-of-squares if and only if for each $i$ the polynomial

$$\sum_{\alpha \in \mathcal{F}_i} b_\alpha x^\alpha$$

is a sum-of-squares. (An analogous result holds for nonnegative polynomials [120, Theorem 3.6].)

The polynomial $f(x_1, x_2) = 1 + b_1 x_1^4 + b_2 x_2^4$, for instance, is supported only on vertices of new$(f)$, and therefore has this property. If $M$ is a set of monomial exponents for which $(M + M) = \text{new}(f) \cap \mathbb{N}^n$, we can prove this fact by constructing a transitive relation $R_{\mathcal{T}}$ as follows:

**Lemma 8.4.5.** *Let $M$ be a finite subset of $\mathbb{N}^n$ and let $\mathcal{F}_1, \ldots, \mathcal{F}_p \subseteq \mathbb{R}^n$ denote pairwise disjoint faces of a convex polytope that contains $M+M$. If $\mathcal{T} \subseteq M+M$ satisfies*

$$\mathcal{T} = \cup_{i=1}^{p}(M+M) \cap \mathcal{F}_i,$$

*then the relation $R_{\mathcal{T}} = \{(\alpha, \beta) \in M \times M : \alpha + \beta \in \mathcal{T}\}$ is transitive.*

*Proof.* Fix $(\beta, \gamma) \in R_{\mathcal{T}}$ and $(\gamma, \mu) \in R_{\mathcal{T}}$. By assumption, there exists faces $\mathcal{F}_i$ and $\mathcal{F}_j$ for which

$$\beta + \gamma \in \mathcal{F}_i, \qquad \gamma + \mu \in \mathcal{F}_j,$$

or, equivalently,

$$\frac{1}{2}(2\beta + 2\gamma) \in \mathcal{F}_i, \qquad \frac{1}{2}(2\gamma + 2\mu) \in \mathcal{F}_j.$$

Since $\mathcal{F}_i$ and $\mathcal{F}_j$ are faces of a polytope containing $M+M$, we conclude that $2\gamma \in \mathcal{F}_i \cap \mathcal{F}_j$, which, using the pairwise-disjointness assumption, implies that $\mathcal{F}_i = \mathcal{F}_j$. We also have that $2\beta, 2\mu \in \mathcal{F}_i$, which, since $\mathcal{F}_i$ is convex, shows that

$$\frac{1}{2}(2\beta + 2\mu) = \beta + \mu \in \mathcal{F}_i,$$

implying that $\beta + \mu \in \mathcal{T}$, i.e., $(\beta, \mu) \in R_{\mathcal{T}}$.

□

**Transitive relations from functions** Another block-diagonalization technique from [35] constructs a function $\psi : M + M \to 2^M$ with the following properties:

$$\beta + \gamma = \alpha \quad \Rightarrow \psi(2\beta) \cup \psi(2\gamma) \subseteq \psi(\alpha). \tag{8.3}$$

To gain intuition behind the function $\psi$, observe that the *minimal face* operation $\text{face}(\cdot, P)$ of a polytope $P$ satisfies a variant of property (8.3), i.e., if $2\beta$, $2\gamma$, and $\alpha$ are contained in some polytope $P$, then

$$\frac{2\beta + 2\gamma}{2} = \alpha \Rightarrow \text{face}(2\gamma, P) \cup \text{face}(2\beta, P) \subseteq \text{face}(\alpha, P).$$

This parallel helps the authors of [35] generalize the Newton-polytope-based block-diagonalization technique described in Section 8.4.2. While we will not state the specific definition of $\psi$ used in [35], we nevertheless show it defines a transitive relation when paired with pairwise-disjoint subsets $P_i$ of $M$. Consider the following.

**Lemma 8.4.6.** *Let $M$ be a finite subset of $\mathbb{N}^n$. Let $\psi : M + M \to 2^M$ be a function with the property* (8.3)*, and in addition assume that $\psi(\alpha) \neq \emptyset$ for all $\alpha \in M + M$.*

*Next, suppose $P_1, \ldots, P_p$ are pairwise-disjoint subsets of $M$ for which*

$$\psi(\beta+\gamma) \subseteq P_i \text{ when } \psi(2\beta), \psi(2\gamma) \subseteq P_i.$$

*If $\mathcal{T} = \{\alpha \in M+M : \psi(\alpha) \subseteq P_i \text{ for some } i \in [p]\}$, then $R_{\mathcal{T}} = \{(\alpha, \beta) \in M \times M : \alpha+\beta \in \mathcal{T}\}$ is transitive.*

*Proof.* For $\gamma, \beta, \mu \in M$, suppose that $(\beta, \gamma) \in R_{\mathcal{T}}$ and $(\gamma, \mu) \in R_{\mathcal{T}}$. Then, $\gamma+\beta$ and $\gamma+\mu$ are in $\mathcal{T}$, which implies

$$\psi(2\gamma) \cup \psi(2\beta) \subseteq \psi(\gamma+\beta) \subseteq P_i$$

and

$$\psi(2\gamma) \cup \psi(2\mu) \subseteq \psi(\gamma+\mu) \subseteq P_j$$

for some $P_i$ and $P_j$. Since $\psi(2\gamma) \subseteq P_i \cap P_j$, the pairwise-disjointness assumption implies that $P_i = P_j$. Since $\psi(2\beta), \psi(2\mu) \subseteq P_i$, we have, by our assumption on $P_i$, that $\psi(\beta+\mu) \subseteq P_i$, showing that $\beta+\mu \in \mathcal{T}$, i.e., $(\beta, \mu) \in R_{\mathcal{T}}$.

□

# Notation

### Sets

| | |
|---|---|
| $\mathcal{V}, \mathcal{W}$ | finite-dimensional inner product spaces |
| $\mathcal{A}$ | an affine subset of an inner product space $\mathcal{V}$ |
| $[n]$ | the finite set $\{1, 2, \ldots, n\}$ |

### Linear maps

| | |
|---|---|
| $\Phi : \mathcal{W} \to \mathcal{V}$ | a linear map between inner product spaces $\mathcal{W}$ and $\mathcal{V}$ |
| $\Phi^* : \mathcal{V} \to \mathcal{W}$ | the adjoint of $\Phi$ |
| $\Phi \cdot \mathcal{X}$ | the image $\{\Phi x : x \in \mathcal{X}\}$ of $\mathcal{X} \subseteq \mathcal{W}$ under the map $\Phi : \mathcal{W} \to \mathcal{V}$ |
| $\operatorname{range} \Phi$ | the range of $\Phi : \mathcal{W} \to \mathcal{V}$ |
| $\operatorname{null} \Phi$ | the null space (kernel) of $\Phi : \mathcal{W} \to \mathcal{V}$ |

### Convex cones

| | |
|---|---|
| $\mathcal{K}$ | a convex cone |
| $\mathcal{K}^*$ | the dual cone of $\mathcal{K}$ |
| $\mathcal{F}$ | a face of $\mathcal{K}$ |
| $\operatorname{relint} \mathcal{K}$ | the relative interior of $\mathcal{K}$ |

### Subspaces

| | |
|---|---|
| $\mathcal{L}$ | a linear subspace of an inner product space $\mathcal{V}$ |
| $\mathcal{L}^\perp$ | the orthogonal complement of $\mathcal{L}$ |
| $\operatorname{span} \mathcal{X}$ | the linear subspace spanned by $\mathcal{X} \subseteq \mathcal{V}$ |
| $s^\perp$ | the orthogonal complement of $\operatorname{span}\{s\}$ if $s \in \mathcal{V}$ (a hyperplane) |

### Matrices

| | |
|---|---|
| $\operatorname{Tr} X$ | the trace of a matrix $X$ |
| $\mathbb{S}^n$ | the vector space of $n \times n$ symmetric matrices equipped with trace inner product |
| $\mathbb{S}^n_+$ | the cone of $n \times n$ symmetric positive semidefinite matrices |

### Jordan algebras

| | |
|---|---|
| $\mathbf{J}$ | a Euclidean Jordan algebra |
| $x \circ y$ | the Jordan product between $x, y \in \mathbf{J}$ |
| $x^2$ | the square $x \circ x$ of $x \in \mathbf{J}$ |

# Bibliography


[1] A. A. Ahmadi and A. Majumdar. *DSOS* and *SDSOS* optimization: LP and SOCP-based alternatives to sum of squares optimization. In *Proceedings of the 48th Annual Conference on Information Sciences and Systems*, pages 1–5, 2014.

[2] B. Alipanahi, N. Krislock, A. Ghodsi, H. Wolkowicz, L. Donaldson, and M. Li. Protein structure by semidefinite facial reduction. In *Research in Computational Molecular Biology*, pages 1–11. Springer, 2012.

[3] F. Alizadeh. An introduction to formally real Jordan algebras and their applications in optimization. In M. F. Anjos and J. B. Lasserre, editors, *Handbook on Semidefinite, Conic and Polynomial Optimization*, pages 297–337. Springer, 2012.

[4] E. Andersen, C. Roos, and T. Terlaky. Notes on Duality in Second Order and p-Order Cone Optimization. *Optimization*, 51(4):627–643, 2002.

[5] E. D. Andersen and K. D. Andersen. Presolving in linear programming. *Mathematical Programming*, 71(2):221–245, 1995.

[6] M. F. Anjos and H. Wolkowicz. Strengthened semidefinite relaxations via a second lifting for the max-cut problem. *Discrete Applied Mathematics*, 119(1):79–106, 2002.

[7] W. Arveson. *An invitation to C*-algebras*, volume 39. Springer Science & Business Media, 2012.

[8] L. Babel, I. V. Chuvaeva, M. Klin, and D. V. Pasechnik. Algebraic combinatorics in mathematical chemistry. methods and algorithms. ii. program implementation of the weisfeiler-leman algorithm. *arXiv preprint arXiv:1002.1921*, 2010.

[9] V. Balakrishnan and L. Vandenberghe. Semidefinite programming duality and linear time-invariant systems. *Automatic Control, IEEE Transactions on*, 48(1): 30–41, 2003.

[10] G. P. Barker and D. Carlson. Cones of diagonally dominant matrices. *Pacific Journal of Mathematics*, 57(1):15–32, 1975.

[11] V. Baston. Extreme copositive quadratic forms. *Acta Arithmetica*, 15(3):319–327, 1969.

[12] A. Berman and N. Shaked-Monderer. *Completely positive matrices*. World Scientific, River Edge (NJ), London, Singapore, 2003. ISBN 981-238-368-9. URL `http://opac.inria.fr/record=b1130077`.

[13] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar. *Convex analysis and optimization*. Athena Scientific Belmont, 2003.

[14] R. Bhatia. *Positive definite matrices*. Princeton university press, 2009.

[15] G. Blekherman, P. A. Parrilo, and R. R. Thomas. *Semidefinite optimization and convex algebraic geometry*. SIAM, 2013.

[16] G. Blekherman, G. Smith, and M. Velasco. Sums of squares and varieties of minimal degree. *Journal of the American Mathematical Society*, 29(3):893–913, 2016.

[17] R. Bödi, T. Grundhöfer, and K. Herr. Symmetries of linear programs. *Note di Matematica*, 30(1):129–132, 2011.

[18] R. Bödi, K. Herr, and M. Joswig. Algorithms for highly symmetric linear and integer programs. *Mathematical Programming*, pages 1–26, 2013.

[19] E. G. Boman, D. Chen, O. Parekh, and S. Toledo. On factor width and symmetric $H$-matrices. *Linear algebra and its applications*, 405:239–248, 2005.

[20] J. Borwein and H. Wolkowicz. Regularizing the abstract convex program. *Journal of Mathematical Analysis and Applications*, 83(2):495–530, 1981.

[21] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, 2009.

[22] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear matrix inequalities in system and control theory*, volume 15. SIAM, 1994.

[23] S. Boyd, P. Diaconis, P. Parrilo, and L. Xiao. Fastest mixing markov chain on graphs with symmetries. *SIAM Journal on Optimization*, 20(2):792–819, 2009.

[24] S. Boyd, M. Mueller, B. O'Donoghue, and Y. Wang. Performance bounds and suboptimal policies for multi-period investment. *Foundations and Trends in Optimization*, 1(1):1–69, 2013.

[25] P. Brändén. Polynomials with the half-plane property and matroid theory. *Advances in Mathematics*, 216(1):302–320, 2007.

[26] D. Bremner, M. D. Sikirić, D. V. Pasechnik, T. Rehn, and A. Schürmann. Computing symmetry groups of polyhedra. *LMS Journal of computation and mathematics*, 17(1):565–581, 2014.

[27] F. Burkowski, Y.-L. Cheung, and H. Wolkowicz. Efficient use of semidefinite programming for selection of rotamers in protein conformations. Technical report, Technical Report CORR 2011, in progress, University of Waterloo, Waterloo, Ontario, 2011.

[28] S. Burton, Y. Youm, and C. Vinzant. A real stable extension of the Vámos matroid polynomial, In preparation.

[29] P. J. Cameron. Coherent configurations, association schemes and permutation groups. *Groups, combinatorics and geometry*, pages 55–72, 2003.

[30] D. Chen and S. Toledo. Combinatorial characterization of the null spaces of symmetric h-matrices. *Linear algebra and its applications*, 392:71–90, 2004.

[31] Y.-L. Cheung, S. Schurr, and H. Wolkowicz. Preprocessing and regularization for degenerate semidefinite programs. In *Computational and Analytical Mathematics*, pages 251–303. Springer, 2013.

[32] Y.-B. Choe, J. G. Oxley, A. D. Sokal, and D. G. Wagner. Homogeneous multivariate polynomials with the half-plane property. *Advances in Applied Mathematics*, 32(1):88–187, 2004.

[33] M.-D. Choi. Completely positive linear maps on complex matrices. *Quantum Computation and Quantum Information Theory: Reprint Volume with Introductory Notes for ISI TMR Network School, 12-23 July 1999, Villa Gualino, Torino, Italy*, 10:174, 2000.

[34] R. Cogill, S. Lall, and P. A. Parrilo. Structured semidefinite programs for the control of symmetric systems. *Automatica*, 44(5):1411–1417, 2008.

[35] L. Dai and B. Xia. Smaller SDP for SOS decomposition. *arXiv preprint arXiv:1407.2679*, 2014.

[36] C. Danielson and F. Borrelli. Symmetric linear model predictive control. *IEEE Transactions on Automatic Control*, 60(5):1244–1259, 2015.

[37] E. de Klerk. Exploiting special structure in semidefinite programming: A survey of theory and applications. *European Journal of Operational Research*, 201(1): 1–10, 2010.

[38] E. de Klerk and R. Sotirov. A new library of structured semidefinite programming instances. *Optimization Methods & Software*, 24(6):959–971, 2009.

[39] E. de Klerk and R. Sotirov. Improved semidefinite programming bounds for quadratic assignment problems with suitable symmetry. *Mathematical programming*, 133(1):75–91, 2012.

[40] E. de Klerk, C. Roos, and T. Terlaky. Initialization in semidefinite programming via a self-dual skew-symmetric embedding. *Operations Research Letters*, 20(5): 213–221, 1997.

[41] E. de Klerk, T. Terlaky, and K. Roos. *Self-Dual Embeddings*, pages 111–138. Springer US, Boston, MA, 2000. ISBN 978-1-4615-4381-7. doi: 10.1007/978-1-4615-4381-7_5. URL http://dx.doi.org/10.1007/978-1-4615-4381-7_5.

[42] E. de Klerk, C. Dobre, and D. V. Pasechnik. Numerical block diagonalization of matrix*-algebras with application to semidefinite programming. *Mathematical programming*, 129(1):91–111, 2011.

[43] R. Deits and R. Tedrake. Efficient mixed-integer planning for uavs in cluttered environments. In *Robotics and Automation (ICRA), 2015 IEEE International Conference on*, pages 42–49. IEEE, 2015.

[44] P. H. Diananda. On non-negative forms in real variables some or all of which are non-negative. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 58, pages 17–25. Cambridge Univ Press, 1962.

[45] C. Dobre and J. Vera. Exploiting symmetry in copositive programs via semidefinite hierarchies. *Mathematical Programming*, 151(2):659–680, 2015.

[46] D. Drusvyatskiy and H. Wolkowicz. The many faces of degeneracy in conic optimization. *arXiv preprint arXiv:1706.03705*, 2017.

[47] D. Drusvyatskiy, G. Pataki, and H. Wolkowicz. Coordinate shadows of semidefinite and euclidean distance matrices. *SIAM Journal on Optimization*, 25(2): 1160–1178, 2015.

[48] D. Drusvyatskiy, G. Li, and H. Wolkowicz. A note on alternating projections for ill-posed semidefinite feasibility problems. 2016.

[49] W. Eberly and M. Giesbrecht. Efficient decomposition of associative algebras. In *Proceedings of the 1996 international symposium on Symbolic and algebraic computation*, pages 170–178. ACM, 1996.

[50] W. Eberly and M. Giesbrecht. Efficient decomposition of separable algebras. *Journal of Symbolic Computation*, 37(1):35–81, 2004.

[51] J. Faraut and A. Korányi. *Analysis on symmetric cones*. Oxford university press, 1994.

[52] D. Farenick. *Algebras of Linear Transformations*. Universitext. Springer New York, 2012. ISBN 9781461300977.

[53] H. Fawzi and P. A. Parrilo. Self-scaled bounds for atomic cone ranks: applications to nonnegative rank and cp-rank. *arXiv preprint arXiv:1404.3240*, 2014.

[54] L. Faybusovich. Euclidean Jordan algebras and interior-point algorithms. *Positivity*, 1(4):331–357, 1997.

[55] M. J. Fischer and A. R. Meyer. Boolean matrix multiplication and transitive closure. In *Switching and Automata Theory, 1971., 12th Annual Symposium on*, pages 129–131. IEEE, 1971.

[56] R. M. Freund. On the behavior of the homogeneous self-dual model for conic convex optimization. *Mathematical programming*, 106(3):527–545, 2006.

[57] H. A. Friberg. Presolving and regularization in mixed-integer second-order cone optimization. 2016.

[58] K. Fujisawa, M. Kojima, K. Nakata, and M. Yamashita. Sdpa (semidefinite programming algorithm) user's manual—version 6.2. 0. *Department of Mathematical and Com-puting Sciences, Tokyo Institute of Technology. Research Reports on Mathematical and Computing Sciences Series B: Operations Research*, 2002.

[59] K. Gatermann and P. A. Parrilo. Symmetry groups, semidefinite programs, and sums of squares. *Journal of Pure and Applied Algebra*, 192(1–3):95–128, 2004. ISSN 0022-4049. doi: 10.1016/j.jpaa.2003.12.011. URL `http://www.sciencedirect.com/science/article/pii/S0022404904000131`.

[60] C. Godsil and G. F. Royle. *Algebraic graph theory*, volume 207. Springer Science & Business Media, 2013.

[61] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)*, 42(6):1115–1145, 1995.

[62] M. X. Goemans et al. Semidefinite programming in combinatorial optimization. *Mathematical Programming*, 79(1):143–162, 1997.

[63] A. J. Goldman and A. W. Tucker. Theory of linear programming. *Linear inequalities and related systems*, 38:53–97, 1956.

[64] M. Grant and S. Boyd. CVX: MATLAB software for disciplined convex programming (web page and software). Online at http://cvxr.com/.

[65] M. Grohe, K. Kersting, M. Mladenov, and E. Selman. Dimension reduction via colour refinement. In *Algorithms-ESA 2014*, pages 505–516. Springer, 2014.

[66] M. Halická, E. de Klerk, and C. Roos. On the convergence of the central path in semidefinite optimization. *SIAM Journal on Optimization*, 12(4):1090–1099, 2002.

[67] H. Hanche-Olsen and E. Størmer. *Jordan operator algebras*, volume 21. Pitman Advanced Publishing Program, 1984.

[68] J. W. Helton, I. Klep, and S. McCullough. The matricial relaxation of a linear matrix inequality. *Mathematical Programming*, 138(1-2):401–445, 2013.

[69] D. Higman. Coherent configurations. *Geometriae Dedicata*, 4(1):1–32, 1975.

[70] D. Higman. Coherent algebras. *Linear Algebra and its Applications*, 93:209–239, 1987.

[71] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex analysis and minimization algorithms I: Fundamentals*, volume 305. Springer science & business media, 2013.

[72] G. Ivanyos and L. Rónyai. Computations in associative and lie algebras. In *Some tapas of computer algebra*, pages 91–120. Springer, 1999.

[73] N. Jacobson. Structure of alternative and Jordan bimodules. *Osaka Mathematical Journal*, 6(1):1–71, 1954.

[74] N. Jacobson. Composition algebras and their automorphisms. *Rendiconti del Circolo Matematico di Palermo*, 7(1):55–80, 1958.

[75] N. Jacobson. A coordinatization theorem for Jordan algebras. *Proceedings of the National Academy of Sciences*, 48(7):1154–1160, 1962.

[76] N. Jacobson. *Structure and representations of Jordan algebras*, volume 39. American Mathematical Soc., 1968.

[77] P. Jordan, J. von Neumann, and E. P. Wigner. On an algebraic generalization of the quantum mechanical formalism. In *The Collected Works of Eugene Paul Wigner*, pages 298–333. Springer, 1993.

[78] V. King and G. Sagert. A fully dynamic algorithm for maintaining the transitive closure. In *Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pages 492–498. ACM, 1999.

[79] M. Koecher. *The Minnesota notes on Jordan algebras and their applications*, volume 1710. Springer Science & Business Media, 1999.

[80] M. Kojima, S. Kim, and H. Waki. Sparsity in sums of squares of polynomials. *Mathematical Programming*, 103(1):45–62, 2005.

[81] N. Krislock and H. Wolkowicz. Explicit sensor network localization using semidefinite representations and facial reductions. *SIAM Journal on Optimization*, 20 (5):2679–2708, 2010.

[82] M. Laurent. Strengthened semidefinite programming bounds for codes. *Mathematical Programming*, 109(2-3):239–261, 2007.

[83] M. Liu and G. Pataki. Exact duals and short certificates of infeasibility and weak infeasibility in conic linear programming. 2015.

[84] J. Löfberg. YALMIP : A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004. URL `http://users.isy.liu.se/johanl/yalmip`.

[85] J. Löfberg. Pre-and post-processing sum-of-squares programs in practice. *IEEE Transactions on Automatic Control*, 54(5):1007–1011, 2009.

[86] B. F. Lourenço, M. Muramatsu, and T. Tsuchiya. Facial reduction and partial polyhedrality. *arXiv preprint arXiv:1512.02549*, 2015.

[87] Z.-Q. Luo, J. F. Sturm, and S. Zhang. Duality and self-duality for conic convex programming. Technical Report 9620/A, Econometric Institute, Erasmus University Rotterdam, 1996.

[88] Z.-Q. Luo, J. F. Sturm, and S. Zhang. Duality Results for Conic Convex Programming. Technical report, Erasmus University Rotterdam, Erasmus School of Economics (ESE), Econometric Institute, 1997.

[89] T. Maehara and K. Murota. A numerical algorithm for block-diagonal decomposition of matrix*-algebras with general irreducible components. *Japan journal of industrial and applied mathematics*, 27(2):263–293, 2010.

[90] F. Margot. Exploiting orbits in symmetric ilp. *Mathematical Programming*, 98 (1-3):3–21, 2003.

[91] K. McCrimmon. *A taste of Jordan algebras*. Springer Science & Business Media, 2006.

[92] B. D. McKay. *Backtrack programming and the graph isomorphism problem*. University of Melbourne, 1977.

[93] H. D. Mittelmann. An independent benchmarking of sdp and socp solvers. *Mathematical Programming*, 95(2):407–430, 2003.

[94] Mosek APS. The MOSEK optimization software. Online at http://www.mosek.com.

[95] A. Németh and S. Németh. Lattice-like operations and isotone projection sets. *Linear Algebra and its Applications*, 439(10):2815–2828, 2013.

[96] A. Németh and S. Németh. Lattice-like subsets of Euclidean Jordan algebras. *arXiv preprint arXiv:1401.3581*, 2014.

[97] Y. Nesterov. Infeasible start interior-point primal-dual methods in nonlinear programming. CORE Discussion Papers 1995067, Université catholique de Louvain, Center for Operations Research and Econometrics (CORE), 1995.

[98] Y. Nesterov. Squared functional systems and optimization problems. In *High performance optimization*, pages 405–440. Springer, 2000.

[99] A. Packard and J. Doyle. The complex structured singular value. *Automatica*, 29 (1):71–109, 1993.

[100] A. Papachristodoulou, J. Anderson, G. Valmorbida, S. Prajna, P. Seiler, and P. Parrilo. SOSTOOLS version 3.00 sum of squares optimization toolbox for MATLAB. *arXiv preprint arXiv:1310.4716*, 2013.

[101] G. Pataki. The geometry of semidefinite programming. In H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors, *Handbook of Semidefinite Programming*, pages 29–65. Springer, 2000.

[102] G. Pataki. Strong duality in conic linear programming: facial reduction and extended duals. *Computational and Analytical Mathematics*, pages 613–634, 2013.

[103] G. Pataki. Bad semidefinite programs: they all look the same. *arXiv preprint arXiv:1112.1436*, 2016.

[104] G. Pataki and S. Schmieta. The DIMACS library of semidefinite-quadratic-linear programs. Available at http://dimacs.rutgers.edu/Challenges/Seventh/Instances, 1999.

[105] F. Permenter and P. A. Parrilo. Partial facial reduction: simplified, equivalent SDPs via inner approximations of the PSD cone. http://arxiv.org/abs/1408.4685, 2014.

[106] F. Permenter and P. A. Parrilo. Basis selection for SOS programs via facial reduction and polyhedral approximations. In *Proceedings of the IEEE Conference on Decision and Control*, 2014.

[107] F. Permenter and P. A. Parrilo. Finding sparse, equivalent SDPs via minimal-coordinate-projections. In *IEEE 54th Annual Conference on Decision and Control (CDC)*. IEEE, 2015.

[108] F. Permenter and P. A. Parrilo. Dimension reduction for SDPs via Jordan Algebras. 2016. http://arxiv.org/abs/1608.02090.

[109] F. Permenter, H. Friberg, and E. Andersen. Solving conic optimization problems via self-dual embedding and facial reduction: a unified approach. *SIAM Journal of Optimization*, 2017 (to appear).

[110] M. Posa, M. Tobenkin, and R. Tedrake. Lyapunov analysis of rigid body systems with impacts and friction via sums-of-squares. In *Proceedings of the 16th International Conference on Hybrid Systems: Computation and Control*, 2013.

[111] F. A. Potra and R. Sheng. On homogeneous interrior-point algorithms for semidefinite programming. *Optimization Methods and Software*, 9(1-3):161–184, 1998.

[112] S. Prajna, A. Papachristodoulou, P. Seiler, and P. A. Parrilo. *SOSTOOLS: Sum of squares optimization toolbox for MATLAB*, 2004.

[113] A. J. Quist, E. de Klerk, C. Roos, and T. Terlaky. Copositive relaxation for general quadratic programming. *Optimization methods and software*, 9(1-3):185–208, 1998.

[114] M. V. Ramana. An exact duality theory for semidefinite programming and its complexity implications. *Mathematical Programming*, 77(1):129–162, 1997.

[115] M. V. Ramana, L. Tunçel, and H. Wolkowicz. Strong duality for semidefinite programming. *SIAM Journal on Optimization*, 7(3):641–662, 1997.

[116] H. Ramírez and D. Sossa. On the central paths in symmetric cone programming. *Journal of Optimization Theory and Applications*, pages 1–20, 2016.

[117] A. Raymond, M. Singh, and R. R. Thomas. Symmetry in tur\'an sums of squares polynomials from flag algebras. *arXiv preprint arXiv:1507.03059*, 2015.

[118] J. Renegar. Incorporating Condition Measures into the Complexity Theory of Linear Programming. *SIAM Journal on Optimization*, 5(3):506–524, 1995. ISSN 1052-6234. doi: 10.1137/0805026.

[119] B. Reznick. Extremal psd forms with few terms. *Duke Mathematical Journal*, 45 (2):363–374, 1978.

[120] B. Reznick. Forms derived from the arithmetic-geometric inequality. *Mathematische Annalen*, 283(3):431–464, 1989.

[121] R. T. Rockafellar. *Convex analysis*, volume 28. Princeton University Press, 1997.

[122] S. Schmieta and F. Alizadeh. Associative and Jordan algebras, and polynomial time interior-point algorithms for symmetric cones. *Mathematics of Operations Research*, 26(3):543–564, 2001.

[123] A. Schrijver. A comparison of the Delsarte and Lovász bounds. *Information Theory, IEEE Transactions on*, 25(4):425–429, 1979.

[124] P. Seiler. SOSOPT: A toolbox for polynomial optimization. *arXiv preprint arXiv:1308.1889*, 2013.

[125] S. Shrikhande. The uniqueness of the l_2 association scheme. *The annals of mathematical statistics*, pages 781–798, 1959.

[126] E. Størmer. *Positive linear maps of operator algebras*. Springer Science & Business Media, 2013.

[127] E. Størmer and E. G. Effros. Positive projections and Jordan structure in operator algebras. *Mathematica Scandinavica*, 45:127–138, 1979.

[128] J. F. Sturm. *Primal-dual interior point approach to semidefinite programming*. PhD thesis, Thesis Publishers Amsterdam,, The Netherlands, 1997.

[129] J. F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization methods and software*, 11(1-4):625–653, 1999.

[130] J. F. Sturm. Error bounds for linear matrix inequalities. *SIAM Journal on Optimization*, 10(4):1228–1248, 2000.

[131] K.-C. Toh, M. J. Todd, and R. Tütüncü. SDPT3 version 4.0: a MATLAB software for semidefinite-quadratic-linear programming. *http://www.math.nus.edu.sg/ mattohkc/sdpt3.html*, 2009.

[132] L. Tunçel. *Polyhedral and semidefinite programming methods in combinatorial optimization*, volume 27. American Mathematical Soc., 2016.

[133] L. Tunçel and H. Wolkowicz. Strong duality and minimal representations for cone optimization. *Computational optimization and applications*, 53(2):619–648, 2012.

[134] F. Vallentin. Symmetry in semidefinite programs. *Linear Algebra and Its Applications*, 430(1):360–369, 2009.

[135] E. R. van Dam and R. Sotirov. Semidefinite programming and eigenvalue bounds for the graph partition problem. *Mathematical Programming*, 151(2):379–404, 2015.

[136] D. G. Wagner and Y. Wei. A criterion for the half-plane property. *Discrete Mathematics*, 309(6):1385–1390, 2009.

[137] H. Waki. How to generate weakly infeasible semidefinite programs via Lasserre's relaxations for polynomial optimization. *Optimization Letters*, 6(8):1883–1896, 2012.

[138] H. Waki and M. Muramatsu. Facial reduction algorithms for conic optimization problems. *Journal of Optimization Theory and Applications*, pages 188–215.

[139] H. Waki and M. Muramatsu. A facial reduction algorithm for finding sparse sos representations. *Operations Research Letters*, 38(5):361–365, 2010.

[140] H. Waki, M. Nakata, and M. Muramatsu. Strange behaviors of interior-point methods for solving semidefinite programming problems in polynomial optimization. *Computational Optimization and Applications*, 53(3):823–844, 2012.

[141] H. Waki, Y. Ebihara, and N. Sebe. Reduction of SDPs in $\mathcal{H}_\infty$ control of SISO systems and performance limitations analysis. In *Decision and Control (CDC), 2016 IEEE 55th Conference on*, pages 646–651. IEEE, 2016.

[142] B. Weisfeiler. *On construction and identification of graphs*. Springer, 1977.

[143] M. M. Wolf. Quantum channels & operations: Guided tour. *Lecture notes available at http://www-m5. ma. tum. de/foswiki/pub M*, 5, 2012.

[144] H. Wolkowicz and Q. Zhao. Semidefinite programming relaxations for the graph partitioning problem. *Discrete Applied Mathematics*, 96:461–479, 1999.

[145] Y. Ye, M. J. Todd, and S. Mizuno. An $\mathcal{O}(\sqrt{n}L)$-iteration homogeneous and self-dual linear programming algorithm. *Mathematics of Operations Research*, 19(1): 53–67, 1994.

[146] Q. Zhao, S. E. Karisch, F. Rendl, and H. Wolkowicz. Semidefinite programming relaxations for the quadratic assignment problem. *Journal of Combinatorial Optimization*, 2(1):71–109, 1998.