



Dimension reduction for semidefinite programs via Jordan algebras

Frank Permenter¹ · Pablo A. Parrilo¹

Received: 2 December 2016 / Accepted: 5 February 2019 / Published online: 8 March 2019
© Springer-Verlag GmbH Germany, part of Springer Nature and Mathematical Optimization Society 2019

Abstract

We propose a new method for simplifying semidefinite programs (SDP) inspired by symmetry reduction. Specifically, we show if an orthogonal projection map satisfies certain invariance conditions, restricting to its range yields an equivalent primal–dual pair over a lower-dimensional symmetric cone—namely, the cone-of-squares of a Jordan subalgebra of symmetric matrices. We present a simple algorithm for minimizing the rank of this projection and hence the dimension of this subalgebra. We also show that minimizing rank optimizes the direct-sum decomposition of the algebra into simple ideals, yielding an optimal “block-diagonalization” of the SDP. Finally, we give combinatorial versions of our algorithm that execute at reduced computational cost and illustrate effectiveness of an implementation on examples. Through the theory of Jordan algebras, the proposed method easily extends to linear and second-order-cone programming and, more generally, symmetric cone optimization.

Mathematics Subject Classification 90 Operations research, mathematical programming · 49 Calculus of variations and optimal control; optimization (including geometric integration theory)

1 Introduction

Many practically relevant optimization problems can be posed as semidefinite programs (SDPs)—convex optimization problems over the cone of positive semidefinite (psd) matrices. While SDPs are efficiently solved in theory, specific instances may be intractable in practice unless one exploits special structure. Existing techniques for structure exploitation include facial reduction [6,12,32] and symmetry reduction [3,8,19,41]. In this paper, we present a method that builds on this latter technique.

✉ Frank Permenter
fperment@mit.edu

¹ Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139, USA

To explain, we first recall a key step in symmetry reduction: finding an orthogonal projection map whose range intersects the solution set. This projection (called a *Reynolds* or *group-average* operator) maps feasible points to feasible points without changing the objective function, which implies that its range (called the *fixed-point subspace*) contains solutions. This leads to a simple statement of our method: minimize rank—or, equivalently, the dimension of the range—over a tractable subset of maps with this property. As we show, this minimization problem is efficiently solved for arbitrary SDP instances by a simple algorithm. Further, the subset of projections considered strictly contains those implicit in existing symmetry reduction procedures (Sect. 2.1.3); hence, our method is more general.

Symmetry reduction not only reduces the dimension of the feasible set, it also simplifies the semidefinite constraint. This simplification process is informally called block-diagonalization, and it amounts to finding a canonical direct-sum decomposition of the fixed-point subspace. The projection we identify enables similar simplifications. Precisely, the range is always a subalgebra of the *Jordan algebra* of real, symmetric matrices and hence also has a canonical direct-sum decomposition into *simple ideals*. Further, its intersection with the psd cone (the *cone-of-squares* of the subalgebra) has a corresponding decomposition into irreducible *symmetric cones* [14, Chapter 3]. As we review (Sect. 2.3.3), finding this decomposition generalizes current block-diagonalization techniques based on $*$ -algebras [10, 25]. As we show, minimizing the rank of the projection optimizes this decomposition in a precise sense.

Finally, our method easily extends to any symmetric cone optimization problem (including linear and second-order-cone programs). Indeed, via Jordan algebra theory, our algorithm for finding projections extends “word-by-word”, mirroring similar extensions of interior-point methods [1].

We organize this paper as follows. Section 2 contains preliminaries. Section 3 gives an algorithm for finding a minimum-rank projection. Section 4 shows that minimizing rank yields an algebra with an optimal direct-sum decomposition. Section 5 gives combinatorial (but less powerful) versions of our algorithm that can be less costly to execute. Computational results appear in Sect. 6.

2 Preliminaries

We consider a primal–dual pair of semidefinite programs (SDPs) expressed in conic form [29, Chapter 4]:

$$\begin{array}{ll} \text{minimize} & C \cdot X \\ \text{subject to} & X \in Y + \mathcal{L} \\ & X \in \mathbb{S}_+^n \end{array} \quad \begin{array}{ll} \text{maximize} & -Y \cdot S \\ \text{subject to} & S \in C + \mathcal{L}^\perp \\ & S \in \mathbb{S}_+^n. \end{array} \quad (1)$$

Here, $X \in \mathbb{S}^n$ and $S \in \mathbb{S}^n$ are decision variables in the vector space \mathbb{S}^n of real symmetric matrices equipped with trace inner-product $X \cdot Y := \text{Tr } XY$, $\mathbb{S}_+^n \subseteq \mathbb{S}^n$ denotes the (self-dual) cone of psd matrices, $\mathcal{L} \subseteq \mathbb{S}^n$ is a linear subspace with orthogonal complement $\mathcal{L}^\perp \subseteq \mathbb{S}^n$, and $Y + \mathcal{L}$ and $C + \mathcal{L}^\perp$ are affine sets defined by fixed $C \in \mathbb{S}^n$ and $Y \in \mathbb{S}^n$. We refer to X and S as the primal and dual decision variables, respectively,

noting that we have identified the dual space $(\mathbb{S}^n)^*$ with \mathbb{S}^n . (Note that in this form, the complementary slackness condition $X \cdot S = 0$ does not necessarily imply the primal and dual objective values are equal. Rather, they differ by a constant that depends on the particular choice of C and Y .)

Throughout this paper we also, for a subspace $\mathcal{S} \subseteq \mathbb{S}^n$, let $P_{\mathcal{S}} : \mathbb{S}^n \rightarrow \mathbb{S}^n$ denote the corresponding orthogonal projection map, i.e., the unique self-adjoint and idempotent map with range equal to \mathcal{S} .

2.1 Constraint set invariance

Our goal is to find a subspace $\mathcal{S} \subseteq \mathbb{S}^n$ that contains primal and dual solutions of (1) if they exist. To do this, we will find a projection that maps feasible points to feasible points without changing the cost function (which implies the range contains solutions), a key idea from symmetry reduction [3,8,19,41]. Precisely, we will search over the orthogonal projections that satisfy the following set of conditions, which we'll show are also implicit in existing symmetry reduction approaches (Sect. 2.1.3).

Definition 2.1 (*Constraint Set Invariance Conditions*) We say the orthogonal projection map $P_{\mathcal{S}} : \mathbb{S}^n \rightarrow \mathbb{S}^n$ satisfies the *Constraint Set Invariance Conditions* for the primal–dual pair (1) if

- (a) $P_{\mathcal{S}}(\mathbb{S}_+^n) \subseteq \mathbb{S}_+^n$, i.e., $P_{\mathcal{S}}$ is a *positive* map;
- (b) $P_{\mathcal{S}}(Y + \mathcal{L}) \subseteq Y + \mathcal{L}$;
- (c) $P_{\mathcal{S}}(C + \mathcal{L}^\perp) \subseteq C + \mathcal{L}^\perp$.

Under these conditions, $P_{\mathcal{S}} : \mathbb{S}^n \rightarrow \mathbb{S}^n$ maps primal/dual feasible points to primal/dual feasible points (by definition). For C and all $X \in Y + \mathcal{L}$, these conditions also imply that

$$X - P_{\mathcal{S}}(X) \in \mathcal{L} \quad C - P_{\mathcal{S}}(C) \in \mathcal{L}^\perp, \quad (2)$$

which in turn implies that $P_{\mathcal{S}}$ preserves the cost function on the primal feasible set:

$$C \cdot X = P_{\mathcal{S}}(C) \cdot P_{\mathcal{S}}(X) = C \cdot P_{\mathcal{S}}P_{\mathcal{S}}(X) = C \cdot P_{\mathcal{S}}(X).$$

(Here, the first equality holds given (2), and the second and third given that $P_{\mathcal{S}}$ is self-adjoint and idempotent.) A similar argument shows $Y \cdot S = Y \cdot P_{\mathcal{S}}(S)$ for all dual feasible S . In summary, we've proven the following.

Proposition 2.1 (Preservation of optimal values) *Suppose $P_{\mathcal{S}} : \mathbb{S}^n \rightarrow \mathbb{S}^n$ satisfies the Constraint Set Invariance Conditions for the primal–dual pair (1). Let $\theta_p := \inf \{C \cdot X : X \in \mathbb{S}_+^n \cap (Y + \mathcal{L})\}$ and $\theta_d := \sup \{-Y \cdot S : S \in \mathbb{S}_+^n \cap (C + \mathcal{L}^\perp)\}$. Then,*

$$\begin{aligned} \theta_p &= \inf \{C \cdot X : X \in \mathbb{S}_+^n \cap (Y + \mathcal{L}) \cap \mathcal{S}\}, \\ \theta_d &= \sup \{-Y \cdot S : S \in \mathbb{S}_+^n \cap (C + \mathcal{L}^\perp) \cap \mathcal{S}\}. \end{aligned}$$

Further, \mathcal{S} contains points that attain θ_p and θ_d when such points exist.

In other words, we've shown that restricting the primal/dual feasible set to \mathcal{S} doesn't change the primal/dual optimal value or its attainment.

2.1.1 Infeasibility certificates

A dual improving direction is $S \in \mathbb{S}_+^n \cap \mathcal{L}^\perp$ satisfying $Y \cdot S < 0$. Analogously, a primal improving direction is $X \in \mathbb{S}_+^n \cap \mathcal{L}$ satisfying $C \cdot X < 0$. The existence of primal (resp. dual) improving directions implies infeasibility of the dual (resp. primal). It turns out that if $P_{\mathcal{S}} : \mathbb{S}^n \rightarrow \mathbb{S}^n$ satisfies the Constraint Set Invariance Conditions, then the subspace \mathcal{S} contains improving directions whenever they exist for the original problem. To show this, we need the following lemma.

Lemma 2.1 (Invariance of linear subspaces) *Suppose $P_{\mathcal{S}} : \mathbb{S}^n \rightarrow \mathbb{S}^n$ satisfies the Constraint Set Invariance Conditions. Then \mathcal{L} and \mathcal{L}^\perp are both invariant subspaces of $P_{\mathcal{S}}$, i.e., $P_{\mathcal{S}}(\mathcal{L}) \subseteq \mathcal{L}$ and $P_{\mathcal{S}}(\mathcal{L}^\perp) \subseteq \mathcal{L}^\perp$.*

Proof For all $Z \in \mathcal{L}$, we have that $P_{\mathcal{S}}(Z) = P_{\mathcal{S}}(Y) - P_{\mathcal{S}}(Y - Z) \in \mathcal{L}$, where containment in \mathcal{L} follows given that $Y + \mathcal{L}$ contains both $P_{\mathcal{S}}(Y)$ and $P_{\mathcal{S}}(Y - Z)$ by the Constraint Set Invariance Conditions. This shows that \mathcal{L} is an invariant subspace; the proof for \mathcal{L}^\perp is identical. \square

We can now show the desired result.

Proposition 2.2 (Improving directions) *Suppose $P_{\mathcal{S}} : \mathbb{S}^n \rightarrow \mathbb{S}^n$ satisfies the Constraint Set Invariance Conditions. The following statements hold*

- If $S \in \mathbb{S}^n$ is a dual improving direction, then so is $P_{\mathcal{S}}(S)$.
- If $X \in \mathbb{S}^n$ is a primal improving direction, then so is $P_{\mathcal{S}}(X)$.

Proof Let S be a dual improving direction. Lemma 2.1 and the Constraint Set Invariance Conditions imply that $\mathbb{S}_+^n \cap \mathcal{L}^\perp$ contains $P_{\mathcal{S}}(S)$, that \mathcal{L}^\perp contains $S - P_{\mathcal{S}}(S)$ and that \mathcal{L} contains $Y - P_{\mathcal{S}}(Y)$. These latter two facts imply that $S \cdot Y = S \cdot P_{\mathcal{S}}(Y)$; hence, $P_{\mathcal{S}}(S)$ is a dual improving direction. Proof of the second statement is identical. \square

2.1.2 Restricted primal–dual pair

We've seen that intersecting the primal and dual feasible with \mathcal{S} does not change the primal and dual optimal value if $P_{\mathcal{S}}$ satisfies the Constraint Set Invariance Conditions (Proposition 2.1). Further, \mathcal{S} contains solutions or infeasibility certificates for (1) when such objects exist (Propositions 2.1–2.2). These facts allow us to solve (1) by first restricting the primal and dual to \mathcal{S} . The following shows that these restrictions are a primal–dual pair if we view \mathcal{S} as the ambient space.

Proposition 2.3 (Duality and restrictions) *Suppose that $P_{\mathcal{S}} : \mathbb{S}^n \rightarrow \mathbb{S}^n$ satisfies the Constraint Set Invariance Conditions (Definition 2.1). Then, treating the range \mathcal{S} as*

the ambient space, the pair of optimization problems

$$\begin{array}{ll} \text{minimize } P_S(C) \cdot X & \text{maximize } -P_S(Y) \cdot S \\ \text{subject to } X \in P_S(Y) + \mathcal{L} \cap S & \text{subject to } S \in P_S(C) + \mathcal{L}^\perp \cap S \\ & X \in \mathbb{S}_+^n \cap S \quad S \in \mathbb{S}_+^n \cap S \end{array} \quad (3)$$

is a primal–dual pair, i.e.,

$$(\mathbb{S}_+^n \cap S)^* \cap S = \mathbb{S}_+^n \cap S, \quad (\mathcal{L} \cap S)^\perp \cap S = \mathcal{L}^\perp \cap S. \quad (4)$$

Moreover,

$$(Y + \mathcal{L}) \cap S = P_S(Y) + \mathcal{L} \cap S, \quad (C + \mathcal{L}^\perp) \cap S = P_S(C) + \mathcal{L}^\perp \cap S. \quad (5)$$

Proof For any set $T \subseteq \mathbb{S}^n$, the condition $P_S(T) \subseteq T$ implies $P_S(T) = S \cap T$ given that P_S is the orthogonal projection onto S . Using this fact, we have that

$$(Y + \mathcal{L}) \cap S = P_S(Y + \mathcal{L}) = P_S(Y) + P_S(\mathcal{L}) = P_S(Y) + \mathcal{L} \cap S,$$

where the last equality uses the additional fact that $P_S(\mathcal{L}) \subseteq \mathcal{L}$ (Lemma 2.1). The other equality in (5) follows by identical argument. The inclusions \supseteq of (4) are obvious. To see the inclusions \subseteq , let T be any set satisfying $P_S(T) \subseteq T$. Then, for any $X \in (T \cap S)^* \cap S$,

$$\langle X, Y \rangle = \langle P_S(X), Y \rangle = \langle X, P_S(Y) \rangle \geq 0, \quad \forall Y \in T,$$

where the first equality holds since $X \in S$, the second equality since P_S is self-adjoint and the inequality since $P_S(Y) \in T \cap S$. Hence, $X \in T^*$. \square

We illustrate this proposition with the following example.

Example 2.1 Consider the following primal–dual pair of semidefinite programs:

$$\begin{array}{ll} \text{minimize } x_1 + x_2 & \text{maximize } -(s_5 + 2s_1) \\ \text{subject to} & \text{subject to} \\ \begin{pmatrix} x_1 & 1 & x_3 & x_4 \\ 1 & x_2 & x_4 & -x_3 \\ x_3 & x_4 & 1 & x_5 \\ x_4 & -x_3 & x_5 & 0 \end{pmatrix} \succeq 0 & \begin{pmatrix} 1 & s_1 & s_2 & s_3 \\ s_1 & 1 & -s_3 & s_2 \\ s_2 & -s_3 & s_5 & 0 \\ s_3 & s_2 & 0 & s_6 \end{pmatrix} \succeq 0. \end{array}$$

The projection $P_S : \mathbb{S}^4 \rightarrow \mathbb{S}^4$ satisfies the Constraint Set Invariance Conditions (Definition 2.1) if S equals the span of $\{E_{21} + E_{12}\} \cup \{E_{ii}\}_{i=1}^3$. Hence, one obtains primal

and dual optimal solutions by solving the following restrictions to \mathcal{S} :

$$\begin{array}{ll} \text{minimize} & x_1 + x_2 \\ \text{subject to} & \begin{pmatrix} x_1 & 1 & 0 & 0 \\ 1 & x_2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \succeq 0 \end{array} \qquad \begin{array}{ll} \text{maximize} & -(s_5 + 2s_1) \\ \text{subject to} & \begin{pmatrix} 1 & s_1 & 0 & 0 \\ s_1 & 1 & 0 & 0 \\ 0 & 0 & s_5 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \succeq 0. \end{array}$$

2.1.3 Relationship with prior work

A common symmetry reduction technique, described in [19], assumes existence of a subgroup $\mathcal{G} \subset \mathbb{R}^{n \times n}$ of the group of orthogonal matrices that, for all $U \in \mathcal{G}$, satisfies

$$UCU^T = C, \quad \left\{ UXU^T : X \in Y + \mathcal{L} \right\} \subseteq Y + \mathcal{L}. \quad (6)$$

Under this condition, one can restrict the primal problem to the *fixed-point subspace*

$$\mathcal{F}_{\mathcal{G}} := \{X \in \mathbb{S}^n : UXU^T = X \ \forall U \in \mathcal{G}\}, \quad (7)$$

without changing its optimal value [19, Theorem 3.3], in analogy with Proposition 2.1. (One can also derive analogues of Proposition 2.3 based on these conditions; see, e.g., [11, Proposition 2].) It turns out that the orthogonal projection onto $\mathcal{F}_{\mathcal{G}}$ (called the *Reynolds operator*) satisfies the Constraint Set Invariance Conditions.

To see this, first observe that $P_{\mathcal{F}_{\mathcal{G}}}$ is the map $X \mapsto \frac{1}{|\mathcal{G}|} \sum_{U \in \mathcal{G}} UXU^T$. As shown in [19],

$$P_{\mathcal{F}_{\mathcal{G}}}(\mathbb{S}_{+}^n) \subseteq \mathbb{S}_{+}^n, \quad P_{\mathcal{F}_{\mathcal{G}}}(Y + \mathcal{L}) \subseteq Y + \mathcal{L}, \quad P_{\mathcal{F}_{\mathcal{G}}}(C) = C, \quad (8)$$

given (6) and the fact that $P_{\mathcal{F}_{\mathcal{G}}}(X)$ is a convex combination of points in $\{UXU^T : U \in \mathcal{G}\}$. The proof of the next proposition shows that $P_{\mathcal{F}_{\mathcal{G}}}$ also satisfies $P_{\mathcal{F}_{\mathcal{G}}}(C + \mathcal{L}^{\perp}) \subseteq C + \mathcal{L}^{\perp}$ (and hence the full set of Constraint Set Invariance Conditions).

Proposition 2.4 (Constraint Set Invariance From Groups) *Let $\mathcal{G} \subset \mathbb{R}^{n \times n}$ be a finite group of orthogonal matrices that satisfies (6). Then, $P_{\mathcal{F}_{\mathcal{G}}} : \mathbb{S}^n \rightarrow \mathbb{S}^n$ satisfies the Constraint Set Invariance Conditions (Definition 2.1), and, in addition, the equality $P_{\mathcal{F}_{\mathcal{G}}}(C) = C$.*

Proof Given (8), we only need to show that $P_{\mathcal{F}_{\mathcal{G}}}(C + \mathcal{L}^{\perp}) \subseteq \mathcal{L}^{\perp}$. To begin, $P_{\mathcal{F}_{\mathcal{G}}}(Y + \mathcal{L}) \subseteq Y + \mathcal{L}$ implies that $P_{\mathcal{F}_{\mathcal{G}}}(\mathcal{L}) \subseteq \mathcal{L}$ (Lemma 2.1). Since $P_{\mathcal{F}_{\mathcal{G}}}$ is self-adjoint, $P_{\mathcal{F}_{\mathcal{G}}}(\mathcal{L}) \subseteq \mathcal{L}$ holds if and only if $P_{\mathcal{F}_{\mathcal{G}}}(\mathcal{L}^{\perp}) \subseteq \mathcal{L}^{\perp}$. Since $P_{\mathcal{F}_{\mathcal{G}}}(C) = C$, we conclude that $P_{\mathcal{F}_{\mathcal{G}}}(C + \mathcal{L}^{\perp}) \subseteq C + \mathcal{L}^{\perp}$, as desired. \square

Another technique, surveyed in [8], treats $\mathbb{R}^{n \times n}$ as a **-algebra* with matrix multiplication as a product and transposition as a **-involution*. It then finds any **-subalgebra*, i.e., any subspace closed under matrix multiplication and transposition, that contains

the primal affine set $Y + \mathcal{L}$. If $\mathcal{M} \subseteq \mathbb{R}^n$ is such a $*$ -subalgebra, then $\mathcal{S} := \mathcal{M} \cap \mathbb{S}^n$ contains primal and dual solutions [8, Theorem 2]. Further, the projection $P_{\mathcal{S}}$ satisfies

$$P_{\mathcal{S}}(\mathbb{S}_+^n) \subseteq \mathbb{S}_+^n, \quad P_{\mathcal{S}}(Y + \mathcal{L}) = Y + \mathcal{L}, \quad (9)$$

where the inclusion $P_{\mathcal{S}}(\mathbb{S}_+^n) \subseteq \mathbb{S}_+^n$ holds because \mathcal{M} is a $*$ -subalgebra. It turns out that $P_{\mathcal{S}}(C + \mathcal{L}^\perp) \subseteq C + \mathcal{L}^\perp$ (and hence the full set of Constraint Set Invariance Conditions) also holds.

Proposition 2.5 (Constraint Set Invariance From $*$ -algebras) *Let $\mathcal{M} \subseteq \mathbb{R}^{n \times n}$ be any $*$ -subalgebra containing $Y + \mathcal{L} \subseteq \mathbb{S}^n$. Let $\mathcal{S} = \mathcal{M} \cap \mathbb{S}^n$. Then $P_{\mathcal{S}} : \mathbb{S}^n \rightarrow \mathbb{S}^n$ satisfies the Constraint Set Invariance Conditions (Definition 2.1), and, in addition, the equality $P_{\mathcal{S}}(Y + \mathcal{L}) = Y + \mathcal{L}$.*

Proof Given (9), we only need to show that $P_{\mathcal{S}}(C + \mathcal{L}^\perp) \subseteq C + \mathcal{L}^\perp$.

To begin, since \mathcal{S} contains $Y + \mathcal{L}$, we have that $P_{\mathcal{S}}(Y + \mathcal{L}) = Y + \mathcal{L}$, which in turn implies that

$$\mathcal{L} = P_{\mathcal{S}}(\mathcal{L}). \quad (10)$$

From (10), we conclude that \mathcal{L} is an invariant subspace of $P_{\mathcal{S}}$ which in turn implies that \mathcal{L}^\perp is an invariant subspace of the adjoint of $P_{\mathcal{S}}$. But $P_{\mathcal{S}}$ is self-adjoint; hence,

$$P_{\mathcal{S}}(\mathcal{L}^\perp) \subseteq \mathcal{L}^\perp. \quad (11)$$

We conclude that

$$P_{\mathcal{S}}(C_{\mathcal{L}} + \mathcal{L}^\perp) = C_{\mathcal{L}} + P_{\mathcal{S}}(\mathcal{L}^\perp) \subseteq C_{\mathcal{L}} + \mathcal{L}^\perp,$$

where the equality follows from (10) and the inclusion from (11). Since $C_{\mathcal{L}} + \mathcal{L}^\perp = C + \mathcal{L}^\perp$, the conclusion follows. \square

Note that this proposition puts no condition on objective matrix C of the primal problem. Similarly, [8, Theorem 2] puts no condition on the dual objective function.

2.2 Reformulations over isomorphic, symmetric cones

The fixed-point subspace of symmetry reduction and $*$ -subalgebras have structured intersections with \mathbb{S}_+^n : each intersection is isomorphic to a direct product of psd cones of Hermitian matrices with real, complex, or quaternion entries. Such a product is an instance of a *symmetric cone*, a special type of self-dual cone that admits efficient optimization algorithms [1, 17]. To maintain this feature, Sect. 2.3.1 gives an additional condition on the projection $P_{\mathcal{S}} : \mathbb{S}^n \rightarrow \mathbb{S}^n$ that ensures $\mathbb{S}_+^n \cap \mathcal{S}$ is always isomorphic to a symmetric cone. This next proposition shows that the primal–dual pair (1) can be explicitly reformulated over such an isomorphic cone under the Constraint Set Invariance Conditions.

Proposition 2.6 (Reformulations over isomorphic cones) *Suppose $P_S : \mathbb{S}^n \rightarrow \mathbb{S}^n$ satisfies the Constraint Set Invariance Conditions (Definition 2.1). For an inner-product space \mathcal{V} , let $\Psi : \mathcal{V} \rightarrow \mathbb{S}^n$ be an injective linear map with range equal to \mathcal{S} and $\mathcal{C} \subseteq \mathcal{V}$ a self-dual cone that satisfies*

$$\Psi(\mathcal{C}) = \mathbb{S}_+^n \cap \mathcal{S}. \quad (12)$$

If $\hat{X} \in \mathcal{V}$ and $\hat{S} \in \mathcal{V}$ solve the primal–dual pair of conic optimization problems

$$\begin{array}{ll} \text{minimize } \langle \Psi^*(C), \hat{X} \rangle & \text{maximize } -\langle (\Psi^*\Psi)^{-1}\Psi^*(Y), \hat{S} \rangle \\ \text{subject to } \hat{X} \in (\Psi^*\Psi)^{-1}\Psi^*(Y + \mathcal{L}) & \text{subject to } \hat{S} \in \Psi^*(C + \mathcal{L}^\perp) \\ \hat{X} \in \mathcal{C} & \hat{S} \in \mathcal{C}, \end{array} \quad (13)$$

then $\Psi(\hat{X})$ and $\Psi(\Psi^\Psi)^{-1}(\hat{S})$ solve the primal–dual pair (3)—and hence the primal–dual pair (1).*

Proof We will show that Ψ and $(\Psi^*\Psi)^{-1}\Psi^*$ are mappings between primal feasible points of (13) and (3) that do not change the objective value. To see that Ψ has this property, let \hat{X} be a feasible point of (13). Then,

$$\Psi(\hat{X}) \in P_S(Y + \mathcal{L}), \quad \Psi(\hat{X}) \in \mathbb{S}_+^n \cap \mathcal{S}, \quad \langle C, \Psi(\hat{X}) \rangle = \langle \Psi^*(C), \hat{X} \rangle,$$

where the first containment follows given that $\Psi(\Psi^*\Psi)^{-1}\Psi^*$ equals P_S and the second by (12). Since $P_S(Y + \mathcal{L}) \subseteq Y + \mathcal{L}$, we conclude that $\Psi(\hat{X})$ is feasible for (3) with same objective as \hat{X} .

Now suppose X is feasible for (3). Then $X = \Psi(\hat{X})$ for a unique $\hat{X} \in \mathcal{C}$ since Ψ is injective. Indeed, we must have that $\hat{X} = (\Psi^*\Psi)^{-1}\Psi^*(X)$, since

$$\Psi(\Psi^*\Psi)^{-1}\Psi^*(X) = P_S(X) = X.$$

Hence, $\hat{X} = (\Psi^*\Psi)^{-1}\Psi^*(X)$ is a feasible point of (13) with objective

$$\langle \Psi^*(C), \hat{X} \rangle = \langle \Psi^*(C), (\Psi^*\Psi)^{-1}\Psi^*(X) \rangle = \langle C, \Psi(\Psi^*\Psi)^{-1}\Psi^*(X) \rangle = \langle C, P_S X \rangle = \langle C, X \rangle,$$

as desired.

For the dual, we similarly prove that Ψ^* and $\Psi(\Psi^*\Psi)^{-1}$ are mappings between the feasible sets that do not change the objective. The proof is almost the same, but exploits the additional fact that

$$\Psi^*\Psi(\mathcal{C}) = \mathcal{C}, \quad (14)$$

which we show in the “Appendix” (Lemma 7.4). To begin, if \hat{S} is dual feasible for (13), then $\Psi(\Psi^*\Psi)^{-1}(\hat{S})$ satisfies

$$\Psi(\Psi^*\Psi)^{-1}(\hat{S}) \in P_S(C + \mathcal{L}^\perp), \quad \Psi(\Psi^*\Psi)^{-1}(\hat{S}) \in \mathbb{S}_+^n \cap \mathcal{S},$$

$$\langle Y, \Psi(\Psi^* \Psi)^{-1}(\hat{S}) \rangle = \langle (\Psi^* \Psi)^{-1} \Psi^*(Y), \hat{S} \rangle,$$

Here, the first containment follows because $\hat{S} \in \Psi^*(C + \mathcal{L}^\perp)$ and $\Psi(\Psi^* \Psi)^{-1} \Psi^* = P_S$; the second by (12) and (14). Since $P_S(C + \mathcal{L}^\perp) \subseteq C + \mathcal{L}^\perp$, we conclude that $\Psi(\Psi^* \Psi)^{-1}(\hat{S})$ is feasible for (3) with same objective as \hat{S} .

On the other hand, if S is dual feasible for (3), then $\hat{S} := \Psi^* S$ must be the unique $\hat{S} \in \mathcal{V}$ satisfying $S = \Psi(\Psi^* \Psi)^{-1} \hat{S}$ since $\Psi(\Psi^* \Psi)^{-1} \Psi^* S = S$. Further, $\hat{S} \in \mathcal{C}$ since $\Psi^*(\mathbb{S}_+^n \cap \mathcal{S}) = \mathcal{C}$ by (12) and (14). Hence \hat{S} is dual feasible for (13). Further, its objective satisfies

$$\langle (\Psi^* \Psi)^{-1} \Psi^*(Y), \hat{S} \rangle = \langle (\Psi^* \Psi)^{-1} \Psi^*(Y), \Psi^* S \rangle = \langle Y, P_S S \rangle = \langle Y, S \rangle,$$

as desired. \square

2.3 Euclidean Jordan algebras

We now develop a condition that guarantees $\mathbb{S}_+^n \cap \mathcal{S}$ is isomorphic to a symmetric cone \mathcal{C} and discuss how to construct a linear map Ψ satisfying $\Psi(\mathcal{C}) = \mathbb{S}_+^n \cap \mathcal{S}$. For this, we first view \mathbb{S}^n as a *Euclidean Jordan algebra*, i.e., as an inner-product space \mathcal{J} equipped with bilinear product $(x, y) \mapsto x \circ y$ that is commutative, satisfies the *Jordan identity*

$$x^2 \circ (y \circ x) = (x^2 \circ y) \circ x \quad \forall x, y \in \mathcal{J}$$

(where $x^2 := x \circ x$), and, for all fixed x , is a self-adjoint, i.e., $\langle x \circ y, z \rangle = \langle y, x \circ z \rangle$ for all $y, z \in \mathcal{J}$. To satisfy these axioms, we equip \mathbb{S}^n with the trace inner-product $X \cdot S := \text{Tr } XY$ and product $X \circ Y := \frac{1}{2}(XY + YX)$. For any Euclidean Jordan algebra \mathcal{J} , the set of squares $\{x^2 : x \in \mathcal{J}\}$ is always a symmetric cone [14, Chapter 3], often called the *cone-of-squares* of \mathcal{J} . For the aforementioned product, the cone-of-squares of \mathbb{S}^n is just the psd cone \mathbb{S}_+^n .

It turns out that $\mathbb{S}_+^n \cap \mathcal{S}$ is isomorphic to a symmetric cone whenever \mathcal{S} is a *subalgebra* of \mathbb{S}^n , i.e., whenever \mathcal{S} contains $X \circ Y$ for all $X, Y \in \mathcal{S}$. This follows because \mathcal{S} satisfies the Euclidean-Jordan-algebra axioms (when viewed as the ambient space) and has cone-of-squares $\mathbb{S}_+^n \cap \mathcal{S}$. As a consequence, we can write $\mathbb{S}_+^n \cap \mathcal{S}$ as the linear image $\Psi(\mathcal{C})$ of the cone-of-squares \mathcal{C} of any isomorphic algebra \mathcal{J} using an injective *homomorphism* $\Psi : \mathcal{J} \rightarrow \mathbb{S}^n$, i.e., an injective linear map satisfying $\Psi(x \circ y) = \Psi(x) \circ \Psi(y)$. Formally:

Proposition 2.7 *Let \mathcal{S} be a subalgebra of \mathbb{S}^n . Let \mathcal{J} be any Euclidean Jordan algebra isomorphic to \mathcal{S} with cone-of-squares $\mathcal{C} \subseteq \mathcal{J}$. Let $\Psi : \mathcal{J} \rightarrow \mathbb{S}^n$ be an injective homomorphism with range equal to \mathcal{S} . Then,*

$$\Psi(\mathcal{C}) = \mathbb{S}_+^n \cap \mathcal{S}.$$

Example 2.2 Let \mathcal{S} denote the subalgebra of \mathbb{S}^n spanned by

$$E_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, E_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, T_1 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, T_2 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Let \mathcal{J} denote the spin-factor algebra $\mathbb{R} \times \mathbb{R}^3$ with cone-of-squares $\mathcal{Q} := \{(x_0, x) \in \mathcal{J} : \|x\|_2 \leq x_0\}$ and product

$$(x_0, x) \circ (y_0, y) := (x_0 y_0 + x^T y, x_0 y + y_0 x).$$

Finally, let $\Psi : \mathcal{J} \rightarrow \mathbb{S}^n$ denote the injective linear map satisfying

$$\Psi e_i = E_i, \quad \Psi t_i = T_i \quad i \in \{1, 2\},$$

where

$$e_1 = [\tfrac{1}{2} \ \tfrac{1}{2} \ 0 \ 0]^T, e_2 = [\tfrac{1}{2} \ -\tfrac{1}{2} \ 0 \ 0]^T, t_1 = [0 \ 0 \ 1 \ 0]^T, t_2 = [0 \ 0 \ 0 \ 1]^T.$$

Then, the image of \mathcal{Q} under Ψ is $\mathbb{S}_+^n \cap \mathcal{S}$. Further, Ψ is an injective homomorphism from \mathcal{J} into \mathbb{S}^n with range equal to \mathcal{S} .

Given \mathcal{S} , one can find a canonical isomorphic algebra \mathcal{J} and an injective homomorphism Ψ numerically [34, Chapter 6]; see Sect. 2.3.2 for more details.

2.3.1 Positive projections, unitality, and subalgebras

We can guarantee that \mathcal{S} is a subalgebra and hence that $\mathbb{S}^n \cap \mathcal{S}$ is isomorphic to a symmetric cone by revisiting the positivity constraint $P_{\mathcal{S}}(\mathbb{S}_+^n) \subseteq \mathbb{S}_+^n$ of the Constraint Set Invariance Conditions. Specifically, we obtain this guarantee by imposing positivity and, in addition, *unitality*.

Definition 2.2 (*Unitality Condition*) We say that $P_{\mathcal{S}} : \mathbb{S}^n \rightarrow \mathbb{S}^n$ is unital if the range \mathcal{S} contains a unit $E \in \mathcal{S}$ for the Jordan product $X \circ Y = \frac{1}{2}(XY + YX)$, i.e., if there exists $E \in \mathcal{S}$ for which $X \circ E = X$ for all $X \in \mathcal{S}$.

Theorem 2.1 (Characterization of positive, unital projections) *Let $\mathcal{S} \subseteq \mathbb{S}^n$ be subspace with orthogonal projection map $P_{\mathcal{S}} : \mathbb{S}^n \rightarrow \mathbb{S}^n$. The following statements are equivalent.*

1. The projection $P_{\mathcal{S}}$ is positive (Definition 2.1-(a)) and unital (Definition 2.2).
2. The subspace \mathcal{S} is subalgebra of \mathbb{S}^n .

Proof See “Appendix”. □

As we show in the “Appendix”, this theorem follows from basic linear algebra and arguments of [38] (who proves an analogous result for complex Jordan algebras). Note also that the unitality condition holds when \mathcal{S} arises from a group or a $*$ -subalgebra via Proposition 2.4 or 2.5.

2.3.2 Structure of subalgebras

We now discuss the structure of subalgebras in more detail. To begin, call an abstract Euclidean Jordan algebra \mathcal{J} *simple* if its only *ideals* are $\{0\}$ and \mathcal{J} , where an ideal $\mathcal{I} \subseteq \mathcal{J}$ is a subspace satisfying $x \circ y \in \mathcal{I}$ for all $y \in \mathcal{J}$ and $x \in \mathcal{I}$. Similarly, call an ideal *simple* if it is simple when viewed as an algebra. It is well known that any subalgebra \mathcal{S} equals an orthogonal direct-sum of its simple ideals (e.g., [14, Proposition III.4.4]). Further, the isomorphism classes of these ideals are fully understood [14, Chapter V]. As a consequence, \mathcal{S} is always, up-to linear transformation, a direct-sum of “canonical” algebras. Formally:

Proposition 2.8 (Structure theorem for subalgebras [14]) *Let $\oplus_{i=1}^r \mathcal{S}_i$ be the orthogonal direct-sum decomposition of a subalgebra $\mathcal{S} \subseteq \mathbb{S}^n$ into simple ideals. Then, there exists simple Jordan algebras $\mathcal{J}_1, \dots, \mathcal{J}_r$ and injective homomorphisms $\Psi_i : \mathcal{J}_i \rightarrow \mathbb{S}^n$ satisfying*

$$\mathcal{S}_i = \Psi_i(\mathcal{J}_i), \quad (15)$$

where each \mathcal{J}_i is one of the following¹

- A spin-factor algebra $\mathbb{R} \times \mathbb{R}^m$ with product $(x_0, x) \circ (y_0, y) := (x_0 y_0 + x^T y, x_0 y + y_0 x)$
- The set of Hermitian matrices $H_d(\mathcal{D})$ of order d with entries in \mathcal{D} and product $\frac{1}{2}(XY + YX)$, where \mathcal{D} denotes the real numbers, the complex numbers, or the quaternions.

Efficient algorithms exist for finding the ideals \mathcal{S}_i , the algebras \mathcal{J}_i and the homomorphisms Ψ_i given \mathcal{S} [34, Chapter 6].

2.3.3 Connections with *-subalgebras

In some cases, *-algebra techniques (currently used in the SDP literature) can find the decomposition of a Jordan subalgebra \mathcal{S} into its simple ideals, a crucial step in finding (15). To explain, view $\mathbb{R}^{n \times n}$ as a *-algebra with matrix multiplication as a product and transposition as a *-involution, and let $\mathcal{M} \subseteq \mathbb{R}^{n \times n}$ denote the *-subalgebra generated by \mathcal{S} . The *Wedderburn decomposition* [42] of \mathcal{M} is its direct-sum decomposition $\mathcal{M} = \oplus_{i=1}^q \mathcal{M}_i$ into simple ideals \mathcal{M}_i . If $\mathcal{S} = \mathcal{M} \cap \mathbb{S}^n$, then the ideals of \mathcal{M} identify the ideals of \mathcal{S} . Formally:

Proposition 2.9 (Ideals from the Wedderburn decomposition) *Let \mathcal{M} be the *-subalgebra of $\mathbb{R}^{n \times n}$ generated by a Jordan subalgebra \mathcal{S} of \mathbb{S}^n . Let \mathcal{M} have Wedderburn decomposition $\mathcal{M} = \oplus_{i=1}^q \mathcal{M}_i$. If $\mathcal{S} = \mathcal{M} \cap \mathbb{S}^n$, then $\oplus_{i=1}^q (\mathcal{M}_i \cap \mathbb{S}^n)$ is the decomposition of \mathcal{S} into simple ideals.*

¹ We omit the Albert algebra from this list since it is *exceptional*, i.e., it is an algebra that is not *special*. By definition, all subalgebras of \mathbb{S}^n are special [22, 2.3.1]; hence, no subalgebra of \mathbb{S}^n is isomorphic to the Albert algebra.

Proof We need to show that $\mathcal{S} = \bigoplus_{i=1}^q (\mathcal{M}_i \cap \mathbb{S}^n)$ and that $\mathcal{M}_i \cap \mathbb{S}^n$ is a simple ideal.

To begin, write $X \in \mathcal{S}$ as $X = \sum_{i=1}^q X_i$ for $X_i \in \mathcal{M}_i$. Then, $X = \sum_{i=1}^q \frac{1}{2}(X_i + X_i^T)$, where $X_i + X_i^T \in \mathcal{M}_i \cap \mathbb{S}^n$ since \mathcal{M}_i is closed under transposition. Hence, $\mathcal{S} \subseteq \bigoplus_{i=1}^q (\mathcal{M}_i \cap \mathbb{S}^n)$. The reverse containment follows because $\mathcal{S} = \mathcal{M} \cap \mathbb{S}^n = (\bigoplus_{i=1}^q \mathcal{M}_i) \cap \mathbb{S}^n \supseteq \bigoplus_{i=1}^q (\mathcal{M}_i \cap \mathbb{S}^n)$.

That $\mathcal{S}_i := \mathcal{M}_i \cap \mathbb{S}^n$ is an ideal of \mathcal{S} is obvious: if $X \in \mathcal{S}$ and $Y \in \mathcal{S}_i$ then $XY + YX \in \mathcal{M}_i$ since \mathcal{M}_i is an ideal of \mathcal{M} , hence $\frac{1}{2}(XY + YX) \in \mathcal{S}_i$. Further, by the Artin–Wedderburn theorem, each \mathcal{M}_i is isomorphic to the $*$ -algebra of real, complex, or quaternion matrices of some order; hence, \mathcal{S}_i is isomorphic to the Hermitian matrices of real, complex, or quaternion entries of some order and is therefore simple. \square

Algorithms for finding the Wedderburn decomposition of $*$ -subalgebras of $\mathbb{R}^{n \times n}$ include [13,25]; see also [10,20] for decompositions of complex $*$ -algebras.

Remark 1 A subalgebra \mathcal{S} that satisfies $\mathcal{S} = \mathcal{M} \cap \mathbb{S}^n$ is called *reversible*. If a subalgebra is not reversible, one of its simple ideals is isomorphic to a spin-factor algebra. Conversely, if a subalgebra is isomorphic to a spin-factor of dimension larger than 5, it is not reversible [22, Theorem 6.2.5].

3 Minimum rank projections and admissible subspaces

We now show how to find a projection $P_{\mathcal{S}} : \mathbb{S}^n \rightarrow \mathbb{S}^n$ satisfying the Constraint Set Invariance and Unitality Conditions (Definitions 2.1 and 2.2), which, as argued in the previous section, allows one to reformulate the primal–dual pair (1) over a symmetric cone isomorphic to $\mathbb{S}_+^n \cap \mathcal{S}$. As we’ll show, among projections that satisfy these conditions, there exists a unique one of minimum rank. Further, a simple algorithm finds this projection for any instance of the primal–dual pair (1). This will follow by characterizing subspaces whose orthogonal projections satisfy these conditions. We define such a subspace as *admissible*:

Definition 3.1 A subspace \mathcal{S} is *admissible* if its orthogonal projection $P_{\mathcal{S}} : \mathbb{S}^n \rightarrow \mathbb{S}^n$ satisfies the Constraint Set Invariance and Unitality Conditions (Definitions 2.1–2.2).

Theorem 2.1 provided a partial characterization of admissibility, showing that the ranges of positive, unital projections are the Jordan subalgebras of \mathbb{S}^n . To complete a characterization, we need the following result on invariance of the primal–dual affine sets.

Lemma 3.1 For affine sets $Y + \mathcal{L}$ and $C + \mathcal{L}^\perp$, let $Y_{\mathcal{L}^\perp} \in \mathbb{S}^n$ and $C_{\mathcal{L}} \in \mathbb{S}^n$ denote the projections of $Y \in \mathbb{S}^n$ and $C \in \mathbb{S}^n$ onto the subspaces \mathcal{L}^\perp and \mathcal{L} , respectively. The following are equivalent.

1. $C + \mathcal{L}$ and $Y + \mathcal{L}^\perp$ are invariant under the orthogonal projection $P_{\mathcal{S}} : \mathbb{S}^n \rightarrow \mathbb{S}^n$.
2. The subspace \mathcal{S} contains $C_{\mathcal{L}}$ and $Y_{\mathcal{L}^\perp}$ and is an invariant subspace of $P_{\mathcal{L}} : \mathbb{S}^n \rightarrow \mathbb{S}^n$, i.e., \mathcal{S} contains $Y_{\mathcal{L}^\perp}$, $C_{\mathcal{L}}$ and $P_{\mathcal{L}}(\mathcal{S})$.

Proof See “Appendix” (Sect. 7.2). \square

Remark 2 The invariant subspaces of $P_{\mathcal{L}}$ are precisely the invariant subspaces of $P_{\mathcal{L}^\perp}$ [15, Proposition 3.8], which explains the asymmetry of statement Lemma 3.1-(2) with respect to \mathcal{L} and \mathcal{L}^\perp .

We also use the following well-known characterization of subalgebras, which follows given that $XY + YX = (X + Y)^2 - X^2 - Y^2$.

Lemma 3.2 *A subspace $\mathcal{S} \subseteq \mathbb{S}^n$ is a Jordan subalgebra of \mathbb{S}^n (with product $\frac{1}{2}(XY + YX)$) if and only if $\mathcal{S} \supseteq \{X^2 : X \in \mathcal{S}\}$.*

Combining Theorem 2.1 with Lemmas 3.1–3.2 yields our characterization of admissibility.

Theorem 3.1 *A subspace \mathcal{S} is admissible (Definition 3.1) if and only if it satisfies the following conditions:*

$$\begin{aligned}\mathcal{S} &\ni Y_{\mathcal{L}^\perp}, C_{\mathcal{L}}, \\ \mathcal{S} &\supseteq P_{\mathcal{L}}(\mathcal{S}), \\ \mathcal{S} &\supseteq \{X^2 : X \in \mathcal{S}\},\end{aligned}$$

where $Y_{\mathcal{L}^\perp}$ and $C_{\mathcal{L}}$ are as in Lemma 3.1 and \mathcal{L} is the linear subspace of the primal–dual pair (1).

3.1 Optimal subspaces and minimum rank projections

Theorem 3.1 shows that the (arbitrary) intersection of admissible subspaces is admissible. This motivates the following definition.

Definition 3.2 The *optimal admissible subspace* \mathcal{S}_{opt} is the intersection of all admissible subspaces:

$$\mathcal{S}_{opt} = \bigcap \{\mathcal{S} : \mathcal{S} \text{ is admissible}\}.$$

Admissibility of \mathcal{S}_{opt} yields the following corollary of Theorem 3.1.

Corollary 3.1 *The map $P_{\mathcal{S}_{opt}} : \mathbb{S}^n \rightarrow \mathbb{S}^n$ is the minimum-rank orthogonal projection satisfying the Constraint Set Invariance and Unitality Conditions (Definitions 2.1–2.2).*

3.2 Solution algorithm

Theorem 3.1 also suggests a procedure for finding \mathcal{S}_{opt} . First, initialize \mathcal{S} to the subspace spanned by $C_{\mathcal{L}}$ and $Y_{\mathcal{L}^\perp}$. Then, add $P_{\mathcal{L}}(\mathcal{S})$ and the span of $\{X^2 : X \in \mathcal{S}\}$ to \mathcal{S} in an alternating fashion, terminating when the resulting ascending chain of subspaces stabilizes. Formally:

Theorem 3.2 *The optimal admissible subspace \mathcal{S}_{opt} (Definition 3.2) is the output of the following algorithm: where $C_{\mathcal{L}}, Y_{\mathcal{L}^\perp}$ are as in Lemma 3.1 and \mathcal{L} is the linear subspace of the primal–dual pair (1).*

```

 $\mathcal{S} \leftarrow \text{span}\{C_{\mathcal{L}}, Y_{\mathcal{L}^\perp}\}$ 
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 ascending chain stabilizes

```

Proof The algorithm computes an ascending chain of subspaces in finite dimensions which must stabilize to a subspace $\hat{\mathcal{S}}$. Stabilization implies that

$$\hat{\mathcal{S}} = \hat{\mathcal{S}} + P_{\mathcal{L}}(\hat{\mathcal{S}}), \quad \hat{\mathcal{S}} = \hat{\mathcal{S}} + \text{span}\{X^2 : X \in \hat{\mathcal{S}}\}$$

which, since $\hat{\mathcal{S}} \ni C_{\mathcal{L}}, Y_{\mathcal{L}^\perp}$, shows that $\hat{\mathcal{S}}$ is admissible (Theorem 3.1). At every iteration, \mathcal{S} is a subset of \mathcal{S}_{opt} (by induction); hence, $\hat{\mathcal{S}}$ is a subset. But since $\hat{\mathcal{S}}$ is admissible, it contains \mathcal{S}_{opt} (Definition 3.2). We conclude that $\hat{\mathcal{S}} = \mathcal{S}_{opt}$. \square

Note executing this algorithm may be impractical if storing a basis for \mathcal{S} is impractical. To deal with this, we introduce variations in Sect. 5 that restrict to subspaces spanned by bases with efficient combinatorial representations.

4 Optimal decompositions

Admissible subspaces are necessarily subalgebras of \mathbb{S}^n (Theorem 3.1). As a consequence, each admissible subspace has an orthogonal direct-sum decomposition into simple ideals (Sect. 2.3.2). We now prove the direct-sum decomposition of \mathcal{S}_{opt} is optimal in a precise sense.

Our notion of optimality is in terms of the *rank vector* of an algebra $\mathcal{W} = \oplus_{i=1}^w \mathcal{W}_i$

$$r_{\mathcal{W}} := (\text{rank } \mathcal{W}_1, \text{rank } \mathcal{W}_2, \dots, \text{rank } \mathcal{W}_w),$$

where each \mathcal{W}_i is a simple ideal and $\text{rank } \mathcal{W}_i$ is the maximum number of distinct eigenvalues of an $X \in \mathcal{W}_i$; as an example, the rank vectors of $\mathbb{S}^{n_1} \oplus \mathbb{S}^{n_2}$ and \mathbb{R}_+^4 are (n_1, n_2) and $(1, 1, 1, 1)$, respectively. Specifically, we show that the rank vector of \mathcal{S}_{opt} is *weakly majorized* by that of any other admissible subspace. Among other things, this means that \mathcal{S}_{opt} minimizes the rank vector's largest element and the sum of its elements.

Definition 4.1 The vector $x \in \mathbb{Z}^m$ *weakly majorizes* $y \in \mathbb{Z}^n$ if

$$\sum_{i=1}^{\min\{\ell, m\}} [x^\downarrow]_i \geq \sum_{i=1}^{\min\{\ell, n\}} [y^\downarrow]_i \quad \forall \ell \in \{1, \dots, \max\{m, n\}\},$$

where x^\downarrow and y^\downarrow denote x and y with entries sorted in descending order.

Example 4.1 For the following subalgebras \mathcal{U}_i (each parametrized by $t \in \mathbb{R}^m$), the rank vector $r_{\mathcal{U}_i}$ weakly majorizes $r_{\mathcal{U}_{i+1}}$:

$$\begin{aligned} \mathcal{U}_1 &:= \begin{pmatrix} t_1 & t_2 & 0 & 0 & 0 \\ t_2 & t_3 & 0 & 0 & 0 \\ 0 & 0 & t_4 & t_5 & t_6 \\ 0 & 0 & t_5 & t_7 & t_8 \\ 0 & 0 & t_6 & t_8 & t_9 \end{pmatrix} & \mathcal{U}_2 &:= \begin{pmatrix} t_1 & t_2 & 0 & 0 & 0 \\ t_2 & t_3 & 0 & 0 & 0 \\ 0 & 0 & t_4 & t_5 & 0 \\ 0 & 0 & t_5 & t_7 & 0 \\ 0 & 0 & 0 & 0 & t_9 \end{pmatrix} \\ r_{\mathcal{U}_1} &= (2, 3) & r_{\mathcal{U}_2} &= (2, 2, 1) \\ \\ \mathcal{U}_3 &:= \begin{pmatrix} t_1 & t_2 & 0 & 0 & 0 \\ t_2 & t_3 & 0 & 0 & 0 \\ 0 & 0 & t_1 & t_2 & 0 \\ 0 & 0 & t_2 & t_3 & 0 \\ 0 & 0 & 0 & 0 & t_4 \end{pmatrix} & \mathcal{U}_4 &:= \begin{pmatrix} t_1 & 0 & 0 & 0 & 0 \\ 0 & t_1 & 0 & 0 & 0 \\ 0 & 0 & t_1 & 0 & 0 \\ 0 & 0 & 0 & t_1 & 0 \\ 0 & 0 & 0 & 0 & t_2 \end{pmatrix} \\ r_{\mathcal{U}_3} &= (2, 1) & r_{\mathcal{U}_4} &= (1, 1) \end{aligned}$$

We now state our result.

Theorem 4.1 *Let $\mathcal{W} \subseteq \mathbb{S}^n$ be any admissible subspace (Definition 3.1). Let the optimal admissible subspace \mathcal{S}_{opt} (Definition 3.2) and \mathcal{W} have the following decompositions into simple ideals:*

$$\mathcal{S}_{opt} = \oplus_{i=1}^s \mathcal{S}_i, \quad \mathcal{W} = \oplus_{k=1}^w \mathcal{W}_k.$$

Then, $r_{\mathcal{W}} := (\text{rank } \mathcal{W}_1, \dots, \text{rank } \mathcal{W}_w)$ weakly majorizes $r_{\mathcal{S}_{opt}} := (\text{rank } \mathcal{S}_1, \dots, \text{rank } \mathcal{S}_s)$.

To prove this theorem, we will only use the fact that \mathcal{S}_{opt} is a subalgebra of all other admissible subspaces, which is immediate from its definition and Theorem 3.1.

4.1 Proof of Theorem 4.1

We prove the theorem by showing a more general result (Theorem 4.2) about the rank vectors of Euclidean Jordan algebras and their subalgebras. To our knowledge, these results are new. Towards this, we let $x \circ y$ denote the Jordan product of an abstract Euclidean Jordan algebra \mathcal{J} and recall some standard definitions. An *idempotent* is an $x \in \mathcal{J}$ satisfying $x \circ x = x$. An idempotent is *primitive* if it is nonzero and doesn't equal the sum of two different nonzero idempotents. Finally, as mentioned in Sect. 2.3.2, an algebra \mathcal{J} is *simple* if its only ideals are \mathcal{J} and $\{0\}$. We start with a needed technical lemma.

Lemma 4.1 *Let \mathcal{J} be a Euclidean Jordan algebra and let $\mathcal{V} \subseteq \mathcal{J}$ be a subalgebra that is simple (viewed as an algebra). Let $\mathcal{J} = \oplus_{k=1}^w \mathcal{J}_k$ denote the orthogonal direct-sum decomposition of \mathcal{J} into simple ideals. Finally, let $\Phi_k : \mathcal{J} \rightarrow \mathcal{J}$ denote the*

orthogonal projection onto \mathcal{J}_k . The following statements hold for all $k \in [w]$, where $[w] := \{1, \dots, w\}$:

1. If $e \in \mathcal{J}$ is an idempotent, then $\Phi_k e$ is an idempotent.
2. If $e, f \in \mathcal{J}$ are idempotents and $\langle e, f \rangle = 0$, then $\langle \Phi_k e, \Phi_k f \rangle = 0$.
3. Suppose $e, f \in \mathcal{V}$ are nonzero idempotents. If $\Phi_k e \neq 0$, then $\Phi_k f \neq 0$.

Proof Since \mathcal{J}_k is a simple ideal, the projection map Φ_k from \mathcal{J} onto \mathcal{J}_k is a Jordan homomorphism by [22, 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 $\mathcal{J} = \bigoplus_{k=1}^w \mathcal{J}_k$ is an orthogonal direct-sum decomposition of \mathcal{J} . We conclude

$$e = \sum_{k=1}^w \Phi_k e, \quad f = \sum_{k=1}^w \Phi_k f.$$

Since $\langle \Phi_i e, \Phi_j f \rangle \geq 0$ (since the cone-of-squares is self-dual) 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 $e_i = 2t \circ (t \circ f_j) - t^2 \circ f_j$ [14, Corollary IV.2.4]. Since Φ_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 idempotents and hence in the cone-of-squares, it follows $\Phi_k f \neq 0$ if $\Phi_k e \neq 0$. \square

The mentioned results on rank vectors and subalgebras follow.

Theorem 4.2 (Subalgebras and rank vectors) *Let $\mathcal{S} = \bigoplus_{i=1}^s \mathcal{S}_i$ and $\mathcal{W} = \bigoplus_{k=1}^w \mathcal{W}_k$ be Jordan subalgebras of \mathcal{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 $k \in [w]$,

$$\text{rank } \mathcal{W}_k \geq \sum_{i \in I_k} \text{rank } \mathcal{S}_i.$$

2. The vector $r_{\mathcal{W}}$ weakly majorizes $r_{\mathcal{S}}$, where

$$r_{\mathcal{W}} := (\text{rank } \mathcal{W}_1, \dots, \text{rank } \mathcal{W}_w), \quad r_{\mathcal{S}} := (\text{rank } \mathcal{S}_1, \dots, \text{rank } \mathcal{S}_s).$$

Proof First note \mathcal{S}_i contains a set $\mathcal{E}_i := \{e_j^i\}_{j=1}^{\text{rank } \mathcal{S}_i}$ of pairwise-orthogonal idempotents. 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 $\{\Phi_k f : f \in \cup_{i \in I_k} \mathcal{E}_i\}$ are nonzero (Lemma 4.1-3); moreover, they are idempotent (Lemma 4.1-1) and pairwise orthogonal (Lemma 4.1-2). It follows \mathcal{W}_k contains at least $\sum_{i \in I_k} \text{rank } \mathcal{S}_i$ nonzero idempotents that are pairwise orthogonal. Hence, $\text{rank } \mathcal{W}_k \geq \sum_{i \in I_k} \text{rank } \mathcal{S}_i$.

For the second statement, we note the first implies the following: for each $\ell \in \max\{s, w\}$, there is a subset $T \subseteq [w]$ for which

$$\sum_{k \in T} \text{rank } \mathcal{W}_k \geq \sum_{k \in T} \sum_{i \in I_k} \text{rank } \mathcal{S}_i \geq \sum_{i=1}^{\min\{\ell, s\}} [r_{\mathcal{S}}^\downarrow]_i.$$

Specifically, letting π be a permutation of $[s]$ satisfying $[r_{\mathcal{S}}^\downarrow]_i = [r_{\mathcal{S}}]_{\pi(i)}$, we can choose T to be subset of $[w]$ that satisfies

$$\cup_{k \in T} I_k \supseteq \{\pi(i)\}_{i=1}^{\min\{\ell, s\}}.$$

Further, we can choose T to have $|T| \leq \min\{\ell, w\}$, which implies that

$$\sum_{i=1}^{\min\{\ell, w\}} [r_{\mathcal{W}}^\downarrow]_i \geq \sum_{k \in T} \text{rank } \mathcal{W}_k.$$

Hence, the majorization inequality $\sum_{i=1}^{\min\{\ell, w\}} [r_{\mathcal{W}}^\downarrow]_i \geq \sum_{i=1}^{\min\{\ell, s\}} [r_{\mathcal{S}}^\downarrow]_i$ holds. \square

We see that Theorem 4.1 follows immediately from the second statement of Theorem 4.2 since, as mentioned, \mathcal{S}_{opt} is a subalgebra of all other admissible subspaces (Definition 4.1).

5 Combinatorial variations

This section introduces combinatorial restrictions on admissible subspaces (Definition 3.1), aiming to reduce the cost of storing a basis. We consider three types of subspaces (Fig. 1). The first two types have bases encoded by *relations* and *partitions*,

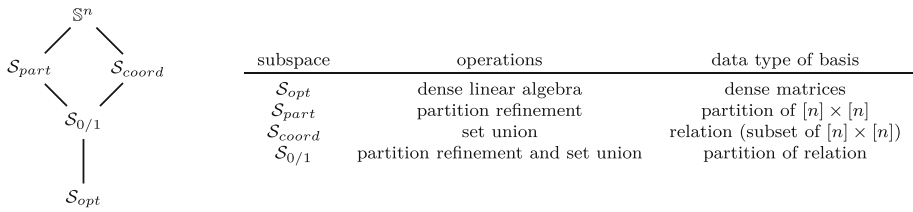


Fig. 1 Hasse diagram of subspace inclusions, key algorithmic operation needed to find subspace, and data type (mathematical object) used to represent a basis

respectively. The third type is a common generalization, whose discussion we defer to the end of this section.

To begin, let $[n] = \{1, \dots, n\}$. For a relation $\mathcal{R} \subseteq [n] \times [n]$, let $\mathcal{B}_{\mathcal{R}} := \{E_{ij} + E_{ji} : (i, j) \in \mathcal{R}\}$, where E_{ij} is the standard basis matrix of $\mathbb{R}^{n \times n}$ nonzero (and equal to 1) only at its (i, j) -th entry. For a partition \mathcal{P} of $[n] \times [n]$, let $\mathcal{B}_{\mathcal{P}}$ denote the corresponding set of characteristic matrices—i.e., let $\mathcal{B}_{\mathcal{P}} := \{\sum_{(i,j) \in C} E_{ij} : C \in \mathcal{P}\}$ where $C \subseteq [n] \times [n]$ denotes a subset in \mathcal{P} .

Definition 5.1 A *coordinate subspace* is the span of $\mathcal{B}_{\mathcal{R}}$ for some relation $\mathcal{R} \subseteq [n] \times [n]$. A *partition subspace* is the span of $\mathcal{B}_{\mathcal{P}}$ for some partition $\mathcal{P} \subseteq [n] \times [n]$.

Example 5.1 The following subspaces \mathcal{S}_1 and \mathcal{S}_2 are coordinate and partition subspaces, respectively.

$$\mathcal{S}_1 = \left\{ \begin{bmatrix} a & b & 0 \\ b & c & d \\ 0 & d & e \end{bmatrix} : (a, b, c, d, e) \in \mathbb{R}^5 \right\} \quad \mathcal{S}_2 = \left\{ \begin{bmatrix} a & a & b \\ a & a & b \\ b & b & c \end{bmatrix} : (a, b, c) \in \mathbb{R}^3 \right\}.$$

Specifically, \mathcal{S}_1 equals the span of $\mathcal{B}_{\mathcal{R}}$ for the relation

$$\mathcal{R} = \{(1, 1), (1, 2), (2, 1), (2, 2), (2, 3), (3, 2), (3, 3)\},$$

and \mathcal{S}_2 equals the span of $\mathcal{B}_{\mathcal{P}}$ for the partition

$$\mathcal{P} = \left\{ \{(1, 1), (1, 2), (2, 1), (2, 2)\}, \{(1, 3), (2, 3), (3, 1), (3, 2)\}, \{(3, 3)\} \right\}.$$

We seek the following variants of \mathcal{S}_{opt} :

$$\begin{aligned} \mathcal{S}_{coord} &:= \bigcap \{\mathcal{S} \subseteq \mathbb{S}^n : \mathcal{S} \text{ is admissible and a coordinate subspace}\}, \\ \mathcal{S}_{part} &:= \bigcap \{\mathcal{S} \subseteq \mathbb{S}^n : \mathcal{S} \text{ is admissible and a partition subspace}\}. \end{aligned}$$

The families of admissible, coordinate, and partition subspaces are all closed under intersection. Hence, \mathcal{S}_{coord} is both admissible and a coordinate subspace. Similar remarks apply for \mathcal{S}_{part} .

Though coordinate subspaces are a highly restricted family, our conference paper [35] illustrates \mathcal{S}_{coord} can have small dimension for SDPs arising in polynomial optimization. Partition subspaces also arise naturally in symmetry reduction. Indeed, the fixed-point subspace (Sect. 2.1.3)

$$\mathcal{M}_{\mathcal{G}} = \{X \in \mathbb{R}^{n \times n} : PXP^T = X \quad \forall P \in \mathcal{G}\}$$

is a partition subspace (of $\mathbb{R}^{n \times n}$) and $\mathcal{M}_{\mathcal{G}} \cap \mathbb{S}^n$ a partition subspace of \mathbb{S}^n when \mathcal{G} is a group of permutation matrices. The partition \mathcal{P} of $[n] \times [n]$ that induces $\mathcal{M}_{\mathcal{G}}$ arises from the orbits $\{PE_{ij}P^T : P \in \mathcal{G}\}$ of the standard basis matrices $E_{ij} \in \mathbb{R}^{n \times n}$; precisely, (i, j) and (k, l) are in the same class of \mathcal{P} if the orbit $\{PE_{ij}P^T : P \in \mathcal{G}\}$ contains E_{kl} . (Such a partition is called a *Schurian coherent configuration* [23].)

5.1 Modified algorithms

To find \mathcal{S}_{part} or \mathcal{S}_{coord} , we modify the Theorem 3.2 algorithm line-by-line to operate on partitions or relations instead of subspaces. These modified algorithms first represent the image of a coordinate/partition subspace \mathcal{S} under the maps $X \mapsto X^2$ and $P_{\mathcal{L}} : \mathbb{S}^n \rightarrow \mathbb{S}^n$ with a polynomial matrix, an idea inspired by [43]; see also [2, Section 5]. They then refine/grow a partition/relation based on the unique/nonzero entries of this polynomial matrix. These algorithms are explicitly given in Fig. 2. They leverage the following notation (Definitions 5.2–5.3).

Definition 5.2 For a finite set $\mathcal{B} \subset \mathbb{S}^n$, let $f_{X^2}(\mathcal{B})$ and $f_{\mathcal{L}}(\mathcal{B})$ denote the polynomial matrices

$$f_{X^2}(\mathcal{B}) := \left(\sum_{B \in \mathcal{B}} t_B B \right)^2 \quad f_{\mathcal{L}}(\mathcal{B}) := \sum_{B \in \mathcal{B}} t_B P_{\mathcal{L}}(B),$$

where $[t_B]_{B \in \mathcal{B}}$ is a vector of commuting² indeterminates indexed by \mathcal{B} .

If \mathcal{S} is the span of \mathcal{B} , then the set of point evaluations of $f_{X^2}(\mathcal{B})$ equals $\{X^2 : X \in \mathcal{S}\}$, i.e.,

$$\{X^2 : X \in \mathcal{S}\} = \left\{ f_{X^2}(\mathcal{B})|_{t_B=t^*} : t^* \in \mathbb{R}^{|\mathcal{B}|} \right\},$$

and similarly for $f_{\mathcal{L}}(\mathcal{B})$. The following example illustrates this notation.

Example 5.2 For $\mathcal{B} = \{U, V, W\}$, where

$$U = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad V = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad W = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

² Note that the related algorithm [2, Section 5] uses *noncommuting* indeterminates.

$\mathcal{R} \leftarrow \text{Supp}(C_{\mathcal{L}}) \cup \text{Supp}(Y_{\mathcal{L}^\perp})$ repeat $\mathcal{R} \leftarrow \mathcal{R} \cup \text{Supp}(f_{\mathcal{L}}(\mathcal{B}_{\mathcal{R}}))$ $\mathcal{R} \leftarrow \mathcal{R} \cup \text{Supp}(f_{X^2}(\mathcal{B}_{\mathcal{R}}))$ until ascending chain of relations \mathcal{R} stabilizes.	$\mathcal{P} \leftarrow \text{Part}(C_{\mathcal{L}}) \vee \text{Part}(Y_{\mathcal{L}^\perp})$ repeat $\mathcal{P} \leftarrow \mathcal{P} \vee \text{Part}(f_{\mathcal{L}}(\mathcal{B}_{\mathcal{P}}))$ $\mathcal{P} \leftarrow \mathcal{P} \vee \text{Part}(f_{X^2}(\mathcal{B}_{\mathcal{P}}))$ until ascending chain of partitions \mathcal{P} stabilizes.
---	--

Fig. 2 Algorithms for finding bases $\mathcal{B}_{\mathcal{R}} \subset \mathbb{S}^n$ and $\mathcal{B}_{\mathcal{P}} \subset \mathbb{S}^n$ of \mathcal{S}_{coord} and \mathcal{S}_{part} , respectively. One algorithm grows a relation $\mathcal{R} \subseteq [n] \times [n]$ and the other refines a partition \mathcal{P} of $[n] \times [n]$. (Here, $\mathcal{P}_1 \vee \mathcal{P}_2$ denotes the coarsest common refinement of partitions \mathcal{P}_1 and \mathcal{P}_2 .) The inputs are $C_{\mathcal{L}}, Y_{\mathcal{L}^\perp} \in \mathbb{S}^n$ and the linear subspace $\mathcal{L} \subseteq \mathbb{S}^n$

we have $f_{X^2}(\mathcal{B}) := (t_U U + t_V V + t_W W)^2$. Expanding (and using the identities $t_U t_V = t_V t_U$ and $t_U t_W = t_W t_U$) shows that

$$f_{X^2}(\mathcal{B}) = \begin{pmatrix} t_U^2 + t_W^2 & 0 & t_U t_W & t_U t_V \\ 0 & t_U^2 + t_W^2 & t_U t_V & t_U t_W \\ t_U t_W & t_U t_V & t_U^2 + t_V^2 & 0 \\ t_U t_V & t_U t_W & 0 & t_U^2 + t_V^2 \end{pmatrix}.$$

Definition 5.3 For an $n \times n$ polynomial matrix X , let $\text{Supp}(X)$ denote the subset of $(i, j) \in [n] \times [n]$ for which X_{ij} is not the zero polynomial. Similarly, let $\text{Part}(X)$ denote the partition of $[n] \times [n]$ induced by the unique polynomial entries of X , i.e., (i, j) and (k, l) are in the same class of $\text{Part}(X)$ if and only if X_{ij} and X_{kl} are the same polynomial.

Example 5.2 (continued) For the polynomial matrix $f_{X^2}(\mathcal{B})$ of the previous example, the relation $\text{Supp}(f_{X^2}(\mathcal{B}))$ is the complement of $\{(1, 2), (2, 1), (3, 4), (4, 3)\} \subseteq [n] \times [n]$ (where $n = 4$.) The partition $\text{Part}(f_{X^2}(\mathcal{B}))$ has characteristic matrices

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (16)$$

$$t_U^2 + t_W^2 \quad 0 \quad t_U t_W \quad t_U t_V \quad t_U^2 + t_V^2,$$

where we've labeled each matrix by the associated polynomial entry of $f_{X^2}(\mathcal{B})$.

5.2 Randomization via sampling

Explicitly constructing symbolic representations of $f_{\mathcal{L}}(\mathcal{B})$ and $f_{X^2}(\mathcal{B})$ is not necessary for finding the partitions and relations they induce. One can instead evaluate the maps $X \mapsto X^2$ and $\mathcal{P}_{\mathcal{L}} : \mathbb{S}^n \rightarrow \mathbb{S}^n$ at a random combination of 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

$$f_{X^2}(\mathcal{B}_{\mathcal{P}})|_{t=t^*} := \left(\sum_{B \in \mathcal{B}_{\mathcal{P}}} t_B^* B \right)^2.$$

The supports of $f_{X^2}(\mathcal{B}_{\mathcal{P}})$ and $f_{X^2}(\mathcal{B}_{\mathcal{P}})|_{t=t^*}$ are the same for almost all t^* . Similarly, the partitions induced by $f_{X^2}(\mathcal{B}_{\mathcal{P}})$ and $f_{X^2}(\mathcal{B}_{\mathcal{P}})|_{t=t^*}$ are the same, i.e.,

$$\text{Part} \left(f_{X^2}(\mathcal{B}_{\mathcal{P}}) \right) = \text{Part} \left(f_{X^2}(\mathcal{B}_{\mathcal{P}})_{|t=t^*} \right),$$

for almost all t^* . The following illustrates this equality for a particular t^* .

Example 5.2 (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 the partition $\text{Part} \left(f_{X^2}(\mathcal{B})_{|t=t_{\mathcal{B}}^*} \right)$ is the same as the partition $\text{Part} \left(f_{X^2}(\mathcal{B}_{\mathcal{P}}) \right)$ given by (16).

5.3 Generalization of partition and coordinate subspaces

Coordinate and partition subspaces have a trivial common generalization: subspaces with an orthogonal basis of 0/1 matrices, or, equivalently, a basis of 0/1 matrices with disjoint support. This motivates the following definition:

$$\mathcal{S}_{0/1} := \bigcap \{ \mathcal{S} \subseteq \mathbb{S}^n : \mathcal{S} \text{ is admissible and has an orthogonal basis of 0/1 matrices} \}.$$

A procedure for finding $\mathcal{S}_{0/1}$ combines features of the algorithms presented for finding $\mathcal{S}_{\text{part}}$ and $\mathcal{S}_{\text{coord}}$: specifically, it iteratively grows a relation \mathcal{R} and refines a partition \mathcal{P} of this relation.

```

Initialize  $\mathcal{R}$  to  $\text{Supp}(C_{\mathcal{L}}) \cup \text{Supp}(Y_{\mathcal{L}^\perp})$ 
Initialize  $\mathcal{P}$  to  $\text{Part}_{\mathcal{R}}(C_{\mathcal{L}}) \vee \text{Part}_{\mathcal{R}}(Y_{\mathcal{L}^\perp})$ 
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  $\mathcal{P} \vee \text{Part}_{\mathcal{R}}(f(\mathcal{B}_{\mathcal{P}}))$ 
  end
until ascending chain of subspaces  $\text{span}(\mathcal{B}_{\mathcal{P}})$  stabilizes.

```

Here $\text{Part}_{\mathcal{R}}(T)$ 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} .

6 Examples

We now apply our techniques to SDPs rising in applications. We do all computation on an Intel 3GHz desktop with 128 gigabytes of RAM. The algorithms of Theorem 3.2 and Fig. 2, used to identify an admissible subspace \mathcal{S} , and the algorithm of

[34, Chapter 6], used to find the linear map Ψ and cone \mathcal{C} satisfying $\Psi(\mathcal{C}) = \mathbb{S}_+^n \cap \mathcal{S}$, were all implemented in MATLAB. We use the solver SeDuMi [40] to solve the SDPs. *Format of original SDPs* Each primal–dual pair is originally expressed in either SeDuMi [40] or SDPA [18] format and may have a mix of free and conic variables, where the cones are either nonnegative orthants or cones of psd matrices.³ From these formats, we eliminate free variables, reformatting the primal problem as

$$\begin{aligned} & \text{minimize} \quad \langle C, X \rangle \\ & \text{subject to} \quad \langle A_i, X \rangle = b_i \quad \forall i \in [m] \\ & \quad \quad \quad X \in \mathbb{S}_+^{n_1} \times \cdots \times \mathbb{S}_+^{n_r}, \end{aligned} \quad (17)$$

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 \mathbb{S}^{n_i} with the trace inner-product. (This reformatting amounts to eliminating free variables and relabeling nonnegative orthants as products of psd cones of order one.) We will in some cases report the number of non-zero (nnz) entries in a description of (17); this equals the number of non-zero floating-point numbers needed to store C and $\{A_i\}_{i=1}^m$. We also report a tuple of ranks for (17), which is simply the tuple (n_1, \dots, n_r) .

Format of reformulations We reformulate each SDP by finding an admissible subspace \mathcal{S} , simple algebras \mathcal{J}_i , and an injective homomorphism $\Psi : \oplus_{i=1}^q \mathcal{J}_i \rightarrow \mathbb{S}^n$ satisfying

$$\mathbb{S}_+^n \cap \mathcal{S} = \Psi(\mathcal{C}_1 \times \cdots \times \mathcal{C}_q),$$

where \mathcal{C}_i is the cone-of-squares of \mathcal{J}_i . The reformulation is as in Proposition 2.6:

$$\begin{aligned} & \text{minimize} \quad \langle \Psi^*(C), \hat{X} \rangle \\ & \text{subject to} \quad \langle \Psi^*(A_i), \hat{X} \rangle = b_i \quad \forall i \in T \subseteq [m] \\ & \quad \quad \quad \hat{X} \in \mathcal{C}_1 \times \cdots \times \mathcal{C}_q, \end{aligned} \quad (18)$$

where T indexes a maximal linearly-independent subset of equations. We will in some cases report the number of non-zero (nnz) entries in a description of (18); this equals the number of non-zero floating-point numbers needed to store $\Psi^*(C)$ and $\{\Psi^*(A_i)\}_{i \in T}$. We also report the tuple (r_1, \dots, r_q) , where r_i is the rank of \mathcal{J}_i .

Remark 3 For most examples, $\mathcal{C}_1 \times \cdots \times \mathcal{C}_q$ is a product of psd cones $\mathbb{S}_+^{r_1} \times \cdots \times \mathbb{S}_+^{r_q}$ and the tuple (r_1, \dots, r_q) indicates their orders. We discuss the only exception in Sect. 6.1.3. 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 (17). As discussed in [8], an SDP can be restricted to the *-algebra generated by its data matrices; see also Proposition 2.5. To compare with this restriction, we let

$$\mathcal{S}_{data} := \mathcal{M}_{data} \cap (\mathbb{S}^{n_1} \times \cdots \times \mathbb{S}^{n_r}),$$

³ These formats also allow for Lorentz cones. None of the examples presented, however, use this type of cone.

where $\mathcal{M}_{data} \subseteq \mathbb{R}^{n \times n}$ is the $*$ -subalgebra of $\mathbb{R}^{n \times n}$ 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

$$\mathcal{S}_{opt} \subseteq \mathcal{S}_{0/1} \subseteq \mathcal{S}_{coord} \subseteq \mathcal{S}_{full}.$$

We'll see that different inclusions hold strictly for different examples. By definition, we also have that

$$\mathcal{S}_{opt} \subseteq \mathcal{S}_{data} \subseteq \mathcal{S}_{full}.$$

Examples will show that \mathcal{S}_{opt} can be (much) smaller than \mathcal{S}_{data} .

6.1 Libraries of problem instances

The first set of SDPs come from three public sources: the parser SOSTOOLS [31], the DIMACS library [33] and a set of structured SDP instances from [9]. Table 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 1 indicates, different ones hold strictly for different instances. For a large fraction, \mathcal{S}_{full} equals \mathcal{S}_{data} , implying generating a $*$ -subalgebra from the problem data [8] does not simplify these instances.

Remark 4 We note the libraries [9,33] have additional instances on which our method was not effective ($\mathcal{S}_{opt} = \mathcal{S}_{full}$); we do not report results for these instances.

Remark 5 The kissing number and copositivity instances of Table 1 (denoted `kissing_x_y_z` and `coposxy`) can also be simplified using group theoretic techniques [7,11] (related to Proposition 2.4) that are tailored to these specific SDP families.

6.1.1 The Lovász number

We give special attention to the Table 2 instances denoted `hamming_m_x` and `hamming_m_x_y`, taken from [33]. For a specific graph G with vertices $\{1, \dots, n\}$ and edge set E , each instance has the following form

$$\begin{aligned} & \text{maximize } \text{Tr } 11^T X \\ & \text{subject to } \text{Tr } X = 1, X \in \mathbb{S}_+^n \\ & \quad \text{Tr}(E_{ij} + E_{ji})X = 0 \quad \forall (i, j) \in E, \end{aligned} \tag{19}$$

where $11^T \in \mathbb{S}^n$ 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 the Hamming graph $H(m, d)$, whose nodes are the Boolean vectors of length m that are adjacent iff their Hamming distance is at least d . The graphs of `hamming_m_x` and `hamming_m_x_y` are modifications of such graphs: nodes are adjacent iff their Hamming distance is x or is x or y . When G is a Hamming graph, it is well known that one can convert SDP (19) into a

Table 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}	\mathcal{S}_{data}	\mathcal{S}_{full}	References
sosdemo2	25	25	28	103	103	Instances from [31]
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 [33]
hamming_8_3_4	5	5	32,896	32,896	32,896	
hamming_9_5_6	6	6	131,328	131,328	131,328	
hamming_9_8	6	6	131,328	131,328	131,328	
hamming_10_2	7	7	524,800	524,800	524,800	
copo14	73	73	1834	1834	1834	
copo23	188	188	8119	8119	8119	
copos68	1576	1576	209,644	209,644	209,644	Instances from [9]
ThetaPrimeER23_red	86	762	777	101	1712	
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	18,577	72,630	479	72,630	
kissing_3_5_5	811	811	3796	3796	3796	
kissing_4_7_7	3723	3723	19,760	19,760	19,760	

linear program using the theory of association schemes [36]. Unsurprisingly, we find similar simplifications for the modified graphs; precisely, $\mathbb{S}_+^n \cap \mathcal{S}_{opt}$ is isomorphic to a non-negative orthant of order equal to the dimension of \mathcal{S}_{opt} , i.e.,

$$\mathbb{S}_+^n \cap \mathcal{S}_{opt} = \Psi(\mathbb{R}_+^{\dim \mathcal{S}_{opt}})$$

for an injective map $\Psi : \mathbb{R}^{\dim \mathcal{S}_{opt}} \rightarrow \mathbb{S}^n$.

As reported [27], these instances are challenging for a wide array of solvers due to their size; indeed, we are only able to solve two of them directly (Table 3). Constructing the reformulation over \mathcal{S}_{opt} , however, converts each SDP into a trivial linear program. Further, finding \mathcal{S}_{opt} and constructing the reformulation takes negligible effort compared to original solver time (Table 3). Note the other automated approach—generating a $*$ -algebra from the data matrices 11^T , I , and $\{E_{ij} + E_{ji}\}_{(i,j) \in E}$ —fails for these instances, i.e., $\mathcal{S}_{data} = \mathcal{S}_{full}$ (Table 2).

Table 2 Tuple of ranks for select examples after restricting to indicated subspace. Here, $s_{t \times}$ means s repeated t times, i.e., $3_{2 \times} := (3, 3)$

Instance	S_{opt}	$S_{0/1}$	S_{coord}	S_{full}
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	$(3, 2_{16 \times}, 1_{56 \times})$	$(35, 33, 5, 1_{56 \times})$	$(35, 33, 5, 1_{75 \times})$	$(73, 1_{75 \times})$
crossing_K_7n	$(36 \times, 2_{4 \times}, 1_{65 \times})$	$(134 \times, 1_{57 \times})$	$(79, 1_{57 \times})$	$(79, 1_{57 \times})$
crossing_K_8n	$(72 \times, 5_{2 \times}, 4_{9 \times}, 3_{7 \times}, 2_{4 \times}, 1_{249 \times})$	$(105, 97, 92, 86, 1_{240 \times})$	$(380, 1_{240 \times})$	$(380, 1_{240 \times})$

Table 3 The original solver time t_{orig} (in seconds) and a list t_{pre}, t_{solve} of preprocessing and solver times for restrictions to indicated subspaces. Failures were due to insufficient memory

Instance	t_{orig}	S_{opt}	$S_{0/1}$	S_{coord}	Instance	t_{orig}	S_{opt}
ThetaPrimeER23_red	0.21	0.16, 0.09	0.12, 0.12	0.02, 0.13	hamming_7_5_6	10.12	0.09, 0.04
ThetaPrimeER29_red	0.19	0.14, 0.10	0.11, 0.15	0.02, 0.14	hamming_8_3_4	4 hours	0.15, 0.02
ThetaPrimeER31_red	0.25	0.17, 0.15	0.13, 0.19	0.02, 0.19	hamming_9_5_6	Fail	0.48, 0.02
crossing_K_7n	0.33	0.25, 0.12	0.18, 0.14	0.02, 0.33	hamming_9_8	Fail	0.49, 0.02
crossing_K_8n	56.7	2.48, 0.58	2.34, 10.37	0.02, 56.7	hamming_10_2	Fail	2.29, 0.03

6.1.2 Decompositions and majorization

In Table 2 we report the tuple of ranks for the subspaces \mathcal{S}_{opt} , $\mathcal{S}_{0/1}$ and \mathcal{S}_{coord} for select examples to confirm our main theorem on optimal decompositions (Theorem 4.1). 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 4.1 predicts the ranks of $\mathcal{S}_{0/1}$ and \mathcal{S}_{coord} weakly majorize those of \mathcal{S}_{opt} in the sense of Definition 4.1. Similarly, it predicts the ranks of \mathcal{S}_{coord} weakly majorize those of $\mathcal{S}_{0/1}$.

Table 2 confirms both these predictions. 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:

$$r_1 := (3, \underbrace{2, 2, \dots, 2}_{12 \times}, \underbrace{1, 1, \dots, 1}_{44 \times}) \quad r_2 := (27, 25, 5, \underbrace{1, 1, \dots, 1}_{44 \times}).$$

It easily follows r_2 weakly majorizes r_1 , i.e., for all positive integers $q \in \mathbb{Z}$,

$$\sum_{i=1}^{\min\{q, l_2\}} [r_2]_i \geq \sum_{i=1}^{\min\{q, l_1\}} [r_1]_i.$$

As also expected, for the instances of Table 2, reformulating over \mathcal{S}_{opt} reduces solver time the most, but requires the most preprocessing (Table 3). In fact, for some instances, solver time reductions do not offset the extra preprocessing time, justifying the larger (but easier to construct) reformulations over $\mathcal{S}_{0/1}$ and \mathcal{S}_{coord} . (To further reduce preprocessing time, Sect. 6.4 introduces an alternative reformulation (21) to (18) that reduces the dimension of the feasible set but doesn't simplify the cone constraint.)

6.1.3 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 $\mu(M, \Delta)$ [30]:

$$\mu(M, \Delta) := \frac{1}{\inf\{\|\Delta\| : \Delta \in \mathbf{\Delta}, \det(I - M\Delta) = 0\}}. \quad (20)$$

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 $\mathbf{\Delta}$. After 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 $r := (\text{rank } \mathcal{S}_1, \dots, \text{rank } \mathcal{S}_{11})$ and $d := (\dim \mathcal{S}_1, \dots, \dim \mathcal{S}_{11})$, we have

$$r = (1, 1, 1, 1, 4, 4^*, 4, 6, 10, 10, 10)$$

$$d = (1, 1, 1, 1, 10, 16^*, 10, 21, 55, 55, 55).$$

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 .

6.2 Coordinate subspaces and sparse decompositions

We next consider SDPs constructed by demonstration scripts packaged with the control system analysis tools available at

<http://www.aem.umn.edu/~AerospaceControl/>,

which build upon the parser SOSOPT [37]. For these SDPs, the optimal subspace \mathcal{S}_{opt} equals the optimal coordinate subspace \mathcal{S}_{coord} . As indicated in Table 4, these SDPs illustrate we can always restrict to \mathcal{S}_{coord} without increasing the number of non-zero entries in the problem description, since restricting to \mathcal{S}_{coord} amounts to setting certain off-diagonal entries of the data to zero. Though these examples are of small size, they illustrate \mathcal{S}_{coord} is a proper subspace of \mathcal{S}_{full} for many SDPs arising in sums-of-squares optimization.

Remark 6 Note some of these scripts construct more than one SDP; reported results are for the *first* SDP constructed.

6.3 Comparison with LP method of Grohe, Kersting, Mladenov, and Selman

In [21], Grohe et al. describe a reduction method for *linear* programming (LP) and show it outperforms a symmetry reduction method of [5] on a collection of LPs; indeed, they show their method theoretically subsumes [5]. The linear programs used for comparison are relaxations of integer programs studied in [26]. 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 outperform [21]. For space reasons, Table 5 reports results for just a small subset of these LP relaxations. To match [21], we give the number of *dual* variables and inequality constraints. In terms of SDP (17) and the SDP (18), the number of dual variables and constraints equals the number of linear equations and the sum of the ranks, respectively.

6.4 Completely-positive rank, the subspace $\mathcal{S}_{0/1}$, and decomposition trade-offs

Our last 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

Table 4 Ranks and number of non-zero (nnz) entries in problem description of original instance and its restriction (18) to \mathcal{S}_{coord}

	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

The notation $r_{s\times}$ indicates r repeated s times

Table 5 Dual variables and constraints of original LP, the LP formulated via the color refinement (CR) method of [21], and the LP formulated via restriction to \mathcal{S}_{opt}

	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

Columns labeled (CR) use numbers reported in [21]

completely-positive rank, or cp-rank for short. The cp-rank of $W \in \mathbb{S}_+^n$ measures the size of the smallest non-negative factorization of W . Precisely, it is the smallest r for which $V \in \mathbb{R}_+^{n \times r}$ exists satisfying $W = VV^T$. (It is infinite if such a factorization

does not exist for any r .) As shown in [16], the cp-rank of $W \in \mathbb{S}^n$ is lower bounded by the optimal value of the following SDP:

$$\begin{aligned} & \text{minimize } t \\ & \text{subject to} \\ & \quad \begin{pmatrix} t & \text{vect } W^T \\ \text{vect } W & X \end{pmatrix} \succeq 0 \\ & \quad X_{ij,ij} \leq W_{ij}^2 \quad \forall i, j \in \{1, \dots, n\} \\ & \quad X \preceq W \otimes W \\ & \quad X_{ij,kl} = X_{il,jk} \quad \forall (1, 1) \leq (i, j) < (k, l) \leq (n, n). \end{aligned}$$

Here, $W \otimes W$ denotes the *Kronecker product* and $\text{vect } W$ 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) and (k, l) hold iff they hold element-wise. (See [16] for 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 6 reports computational savings obtained by restricting to $\mathcal{S}_{0/1}$.

Alternative reformulation For these examples, we compare (18) against an alternative reformulation that reduces the dimension of the dual feasible set, but leaves the cone constraint unchanged. It takes the following form

$$\begin{aligned} & \text{minimize } \langle P_{\mathcal{S}_{0/1}}(C), X \rangle \\ & \text{subject to } \langle P_{\mathcal{S}_{0/1}}(A_i), X \rangle = b_i \quad \forall i \in T \subseteq [m] \\ & \quad X \in \mathbb{S}_+^{n_1} \times \dots \times \mathbb{S}_+^{n_r}, \end{aligned} \quad (21)$$

where $T \subseteq [m]$ indexes a maximal subset of linearly-independent equations, and has dual

$$\begin{aligned} & \text{maximize } \sum_{i \in T} y_i b_i \\ & \text{subject to } P_{\mathcal{S}_{0/1}}(C) - \sum_{i \in T} P_{\mathcal{S}_{0/1}}(A_i) \in \mathbb{S}_+^{n_1} \times \dots \times \mathbb{S}_+^{n_r}. \end{aligned}$$

We can interpret the latter SDP as the dual of (17) restricted to the subspace $\mathcal{S}_{0/1}$, recalling by Proposition 2.1 that $\mathcal{S}_{0/1}$ contains both primal and dual solutions.

Table 6 shows solving (21) achieves computational savings and, indeed, can be preferred to solving (18). As indicated, for the largest instance, we cannot even find the homomorphism Ψ needed to construct (18) due to memory constraints. For this example, the formulation (21) also preserves sparsity.

Table 6 The first row corresponds to the original SDP (17) and the other rows to reformulations over $\mathcal{S}_{0/1}$

SDP	Ranks	num eq	nnz	t_{pre}	t_{solve}
<i>(a) Instance: Z</i>					
Orig. (17)	(10, 9, $19\times$)	37	242	–	.53
Reform. (18)	(5, 4, $24\times$, $16\times$)	14	859	0.34	0.13
Reform. (21)	(10, 9, $19\times$)	14	242	0.11	0.11
<i>(b) Instance: $Z \otimes Z$</i>					
Orig. (17)	(82, 81, $181\times$)	2026	15,752	–	24.34
Reform. (18)	(12, 11, $104\times$, $64\times$, $48\times$, $22\times$, $111\times$)	167	158,199	.98	.90
Reform. (21)	(82, 81, $181\times$)	167	15,752	0.11	2.54
<i>(c) Instance: $Z \otimes Z \otimes Z$</i>					
Orig. (17)	(730, 729, $1729\times$)	142,885	1,182,290	Out of memory	
Reform. (18)	Out of memory			Out of memory	
Reform. (21)	(730, 729, $1729\times$)	1883	1,182,290	6.5	1113

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

Acknowledgements We thank Etienne de Klerk for useful discussions during the beginning stages of this work. We also thank anonymous referees for comments that improved our presentation.

7 Appendix

7.1 Proof of Theorem 2.1

We now prove Theorem 2.1, which stated that a subspace $\mathcal{S} \subseteq \mathbb{S}^n$ is a Jordan subalgebra if and only if its orthogonal projection $P_{\mathcal{S}}$ is unital and positive. Analogues for complex Jordan algebras are well known; see [38,39] and also the thesis [24]. One direction is also shown in [28]. The converse direction is shown in part by translating an argument of [38] from the complex to real case. Since they are short and self-contained, we give full proofs of both directions.

To begin, we need the following lemma relating invariance under squaring to eigenvalue decompositions.

Lemma 7.1 *For a non-zero $X \in \mathbb{S}^n$, let $E_X \subset \mathbb{S}^n$ be the set of pairwise orthogonal idempotent matrices for which*

$$X = \sum_{E \in E_X} \lambda_E E,$$

where the range of $E \in E_X$ is an eigenspace of X and $\{\lambda_E\}_{E \in E_X}$ is the set of non-zero (distinct) eigenvalues of X . For a subspace $\mathcal{S} \subseteq \mathbb{S}^n$, the following are equivalent.

1. \mathcal{S} contains the set E_X for all non-zero $X \in \mathcal{S}$.

2. \mathcal{S} is invariant under squaring, i.e., $\mathcal{S} \supseteq \{X^2 : X \in \mathcal{S}\}$.

Proof That statement one implies two is immediate given that $X^2 = \sum_{E \in E_X} \lambda_E^2 E$. Conversely, suppose X has non-zero eigenvalue λ of maximum magnitude. Then, if statement two holds, the idempotent $\hat{E} = \lim_{n \rightarrow \infty} (|\lambda|^{-1} X)^{2n}$ is contained in \mathcal{S} and has range equal to an eigenspace or, if $\pm|\lambda|$ are both eigenvalues, the sum of two eigenspaces. Replacing X with $X - \lambda \hat{E}$ and iterating yields a set of idempotents whose span contains E_X ; moreover, this set is contained in \mathcal{S} . \square

We now use this lemma and the mentioned argument of [38] to prove Theorem 2.1

To prove $(2 \Rightarrow 1)$, consider $X \succeq 0$ and suppose $P_{\mathcal{S}}(X)$ is non-zero. For a non-zero eigenvalue λ_E of $P_{\mathcal{S}}(X)$, let $E \in \mathbb{S}^n$ denote the idempotent with range equal to the associated eigenspace. If (2) holds, then Lemma 7.1 implies $P_{\mathcal{S}}(E) = E$. Hence,

$$0 \leq E \cdot X = P_{\mathcal{S}}(E) \cdot X = E \cdot P_{\mathcal{S}}(X) = \lambda_E \|E\|^2.$$

We conclude the eigenvalues of $P_{\mathcal{S}}(X)$ are non-negative, i.e., that $P_{\mathcal{S}}(X) \succeq 0$. To show the unitality condition, let Z be a matrix in \mathcal{S} of maximum rank and let

$$\hat{E} = \sum_{E \in E_Z} E.$$

For all $X \in \mathcal{S}$, it holds that $t\hat{E} \succeq X^2$ for some $t > 0$. This shows the range of \hat{E} contains the range of X^2 and hence the range of X . It follows $\hat{E}X = X$.

To prove $(1 \Rightarrow 2)$, suppose the unit element E has rank r . Then we can find an orthogonal matrix $Q = (Q_1, Q_2) \in \mathbb{R}^{n \times n}$ for which $E = Q_1 Q_1^T$ and

$$\mathcal{S} = \left\{ Q \begin{pmatrix} X & 0 \\ 0 & 0 \end{pmatrix} Q^T : X \in \hat{\mathcal{S}} \subseteq \mathbb{S}^r \right\},$$

where $\hat{\mathcal{S}} := Q_1^T \mathcal{S} Q_1$. Further, the projection $P_{\mathcal{S}}$ satisfies

$$P_{\mathcal{S}}(X) = Q_1 Q_1^T P_{\hat{\mathcal{S}}}(X) Q_1 Q_1^T$$

where $P_{\hat{\mathcal{S}}} : \mathbb{S}^r \rightarrow \mathbb{S}^r$ is the orthogonal projection onto $\hat{\mathcal{S}}$. It follows that if $\hat{\mathcal{S}}$ is invariant under squaring, so is \mathcal{S} , and if $P_{\mathcal{S}}$ is positive, so is $P_{\hat{\mathcal{S}}}$. Hence, Statement 2 follows by showing $\hat{\mathcal{S}}$ is invariant under squaring.

We show this applying the argument from [38, 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 $P(X^2) - P(X)P(X) \succeq 0$ for all $X \in \mathbb{S}^r$ (e.g., Theorem 2.3.4 of [4]). Hence, for X in the range of P

$$P(X^2) - X^2 \succeq 0.$$

Letting $Z = P(X^2) - X^2$ and taking the trace shows

$$\begin{aligned}\operatorname{Tr} Z &= I \cdot Z = P(I) \cdot Z = I \cdot P(Z) = \operatorname{Tr} (P^2(X^2) - P(X^2)) \\ &= \operatorname{Tr} (P(X^2) - P(X^2)) = 0.\end{aligned}$$

Since $Z \succeq 0$, we conclude $Z = 0$, i.e., that $P(X^2) = X^2$. Therefore X^2 is in the range of P .

7.2 Invariant affine sets of projections

Recall Condition 2.1-(b) and Condition 2.1-(c) require invariance of the affine sets $Y + \mathcal{L}$ and $C + \mathcal{L}^\perp$ under the projection P_S . We now prove the characterization of these conditions provided by Lemma 3.1.

Lemma 7.2 *For an affine set $Y + \mathcal{L}$, let $Y_{\mathcal{L}^\perp} \in \mathbb{S}^n$ denote the projection of $Y \in \mathbb{S}^n$ onto the subspace \mathcal{L}^\perp . Let $P_S : \mathbb{S}^n \rightarrow \mathbb{S}^n$ denote the orthogonal projection onto a subspace \mathcal{S} of \mathbb{S}^n . The following statements are equivalent.*

1. $P_S(Y + \mathcal{L}) \subseteq Y + \mathcal{L}$
2. $P_S(Y_{\mathcal{L}^\perp}) = Y_{\mathcal{L}^\perp}$ and $P_S(\mathcal{L}) \subseteq \mathcal{L}$

Proof To begin, first note P_S —being an orthogonal projection—is a contraction with respect to the Frobenius norm $\|X\|_F$ (recalling our use of the trace inner-product); further, $Y_{\mathcal{L}^\perp}$ is the unique minimizer of this norm over $Y + \mathcal{L}$. Hence, if $P_S(Y + \mathcal{L}) \subseteq Y + \mathcal{L}$, then $P_S(Y_{\mathcal{L}^\perp}) = Y_{\mathcal{L}^\perp}$; in addition, since $Y + \mathcal{L} = Y_{\mathcal{L}^\perp} + \mathcal{L}$,

$$Y_{\mathcal{L}^\perp} + P_S(\mathcal{L}) = P_S(Y_{\mathcal{L}^\perp} + \mathcal{L}) \subseteq Y_{\mathcal{L}^\perp} + \mathcal{L},$$

which implies $P_S(\mathcal{L}) \subseteq \mathcal{L}$. The converse direction is obvious given that $Y + \mathcal{L} = Y_{\mathcal{L}^\perp} + \mathcal{L}$. \square

If we apply the previous lemma to both the primal and dual affine sets we obtain the conditions $P_S(\mathcal{L}) \subseteq \mathcal{L}$ and $P_S(\mathcal{L}^\perp) \subseteq \mathcal{L}^\perp$. However, Lemma 3.1 only contains one of these conditions, since they turn out to be equivalent. Consider the following.

Lemma 7.3 [15, Proposition 3.8] *Let $P_{\mathcal{L}} : \mathbb{S}^n \rightarrow \mathbb{S}^n$ and $P_S : \mathbb{S}^n \rightarrow \mathbb{S}^n$ denote the orthogonal projections onto subspaces \mathcal{L} and \mathcal{S} of \mathbb{S}^n . The following four statements are equivalent.*

- \mathcal{L} is an invariant subspace of P_S
- \mathcal{L}^\perp is an invariant subspace of P_S
- \mathcal{S} is an invariant subspace of $P_{\mathcal{L}}$
- \mathcal{S}^\perp is an invariant subspace of $P_{\mathcal{L}}$

Combining these two lemmas proves Lemma 3.1.

7.3 Linear images of self-dual cones

The following was used to prove Proposition 2.6.

Lemma 7.4 *Let \mathcal{W} and \mathcal{V} be inner-product spaces and $\mathcal{C} \subseteq \mathcal{V}$ and $\mathcal{K} \subseteq \mathcal{W}$ self-dual convex cones. Let $T : \mathcal{V} \rightarrow \mathcal{W}$ be an injective linear map with adjoint $T^* : \mathcal{W} \rightarrow \mathcal{V}$. If $\mathcal{K} = T(\mathcal{C})$, then $T^*T(\mathcal{C}) = \mathcal{C}$.*

Proof For all $x, y \in \mathcal{C}$,

$$\langle T^*T(x), y \rangle = \langle T(x), T(y) \rangle \geq 0$$

by self-duality of \mathcal{K} . By self-duality of \mathcal{C} , we conclude $T^*T(x) \in \mathcal{C}$. On the other hand, since T^* is surjective, we have for any $x \in \mathcal{C}$ existence of $w \in \mathcal{V}$ for which $x = T^*w$. Further, for all $y \in \mathcal{C}$,

$$0 \leq \langle T^*w, y \rangle = \langle w, Ty \rangle$$

which, since $\mathcal{K} = T(\mathcal{C})$, shows $w \in \mathcal{K}$. Hence, $w = Tz$ for $z \in \mathcal{C}$, showing $x = T^*Tz$. \square

References

1. Alizadeh, F., Schmieta, S.: Symmetric cones, potential reduction methods and word-by-word extensions. In: Wolkowicz, H., Saigal, R., Vandenbergh, L. (eds.) *Handbook of Semidefinite Programming*, pp. 195–233. Springer, Berlin (2000)
2. Babel, L., Chuvaeva, I.V., Klin, M., Pasechnik, D.V.: Algebraic combinatorics in mathematical chemistry. Methods and algorithms. II. Program implementation of the Weisfeiler–Leman algorithm. arXiv preprint [arXiv:1002.1921](https://arxiv.org/abs/1002.1921) (2010)
3. Bachoc, C., Gijswijt, D.C., Schrijver, A., Vallentin, F.: Invariant semidefinite programs. In: Anjos, M.F., Lasserre, J.B. (eds.) *Handbook on Semidefinite, Conic and Polynomial Optimization*, pp. 219–269. Springer, Berlin (2012)
4. Bhatia, R.: *Positive Definite Matrices*. Princeton University Press, Princeton (2009)
5. Bödi, R., Grundhöfer, T., Herr, K.: Symmetries of linear programs. *Note di Matematica* **30**(1), 129–132 (2011)
6. Borwein, J., Wolkowicz, H.: Regularizing the abstract convex program. *J. Math. Anal. Appl.* **83**(2), 495–530 (1981)
7. Caluza Machado, F., de Oliveira Filho, F.M.: Improving the semidefinite programming bound for the kissing number by exploiting polynomial symmetry. *Exp. Math.* **27**(3), 362–369 (2018)
8. de Klerk, E.: Exploiting special structure in semidefinite programming: a survey of theory and applications. *Eur. J. Oper. Res.* **201**(1), 1–10 (2010)
9. de Klerk, E., Sotirov, R.: A new library of structured semidefinite programming instances. *Optim. Methods Softw.* **24**(6), 959–971 (2009)
10. de Klerk, E., Dobre, C., Pasechnik, D.V.: Numerical block diagonalization of matrix*-algebras with application to semidefinite programming. *Math. Program.* **129**(1), 91–111 (2011)
11. Dobre, C., Vera, J.: Exploiting symmetry in copositive programs via semidefinite hierarchies. *Math. Program.* **151**(2), 659–680 (2015)
12. Drusvyatskiy, D., Wolkowicz, H.: The many faces of degeneracy in conic optimization. arXiv preprint [arXiv:1706.03705](https://arxiv.org/abs/1706.03705) (2017)
13. Eberly, W., Giesbrecht, M.: Efficient decomposition of associative algebras. In: *Proceedings of the 1996 International Symposium on Symbolic and Algebraic Computation*, pp. 170–178. ACM (1996)
14. Faraut, J., Korányi, A.: *Analysis on Symmetric Cones*. Oxford University Press, Oxford (1994)

15. Farenick, D.: *Algebras of Linear Transformations*. Universitext. Springer, New York (2012). ISBN 9781461300977
16. Fawzi, H., Parrilo, P.A.: Self-scaled bounds for atomic cone ranks: applications to nonnegative rank and cp-rank. arXiv preprint [arXiv:1404.3240](https://arxiv.org/abs/1404.3240) (2014)
17. Faybusovich, L.: Linear systems in Jordan algebras and primal–dual interior-point algorithms. *J. Comput. Appl. Math.* **86**(1), 149–175 (1997)
18. Fujisawa, K., Kojima, M., Nakata, K., Yamashita, M.: SDPA (semidefinite programming algorithm) user's manual—version 6.2. 0. Department of Mathematical and Computing Sciences, Tokyo Institute of Technology. Research Reports on Mathematical and Computing Sciences Series B: Operations Research (2002)
19. Gatermann, K., Parrilo, P.A.: Symmetry groups, semidefinite programs, and sums of squares. *J. Pure Appl. Algebra* **192**(1–3), 95–128 (2004). <https://doi.org/10.1016/j.jpaa.2003.12.011>
20. Gijswijt, D.: Matrix algebras and semidefinite programming techniques for codes. arXiv preprint [arXiv:1007.0906](https://arxiv.org/abs/1007.0906) (2010)
21. Grohe, M., Kersting, K., Mladenov, M., Selman, E.: Dimension reduction via colour refinement. In: *Algorithms-ESA 2014*, pp. 505–516. Springer (2014)
22. Hanche-Olsen, H., Størmer, E.: *Jordan Operator Algebras*, vol. 21. Pitman Advanced Publishing Program, Edinburgh (1984)
23. Higman, D.: Coherent algebras. *Linear Algebra Appl.* **93**, 209–239 (1987)
24. Idel, M.: On the Structure of Positive Maps. Technical University of Munich, Munich (2013)
25. Machara, T., Murota, K.: A numerical algorithm for block-diagonal decomposition of matrix*-algebras with general irreducible components. *Jpn. J. Ind. Appl. Math.* **27**(2), 263–293 (2010)
26. Margot, F.: Exploiting orbits in symmetric ILP. *Math. Program.* **98**(1–3), 3–21 (2003)
27. Mittelman, H.D.: An independent benchmarking of SDO and SOCP solvers. *Math. Program.* **95**(2), 407–430 (2003)
28. Németh, A., Németh, S.: Lattice-like subsets of Euclidean Jordan algebras. arXiv preprint [arXiv:1401.3581](https://arxiv.org/abs/1401.3581) (2014)
29. Nesterov, Y., Nemirovskii, A., Ye, Y.: *Interior-Point Polynomial Algorithms in Convex Programming*, vol. 13. SIAM, Philadelphia (1994)
30. Packard, A., Doyle, J.: The complex structured singular value. *Automatica* **29**(1), 71–109 (1993)
31. Papachristodoulou, A., Anderson, J., Valmorbida, G., Prajna, S., Seiler, P., Parrilo, P.: SOSTOOLS version 3.0.0 sum of squares optimization toolbox for MATLAB. arXiv preprint [arXiv:1310.4716](https://arxiv.org/abs/1310.4716) (2013)
32. Pataki, G.: Strong duality in conic linear programming: facial reduction and extended duals. *Comput. Anal. Math.*, pp. 613–634 (2013)
33. Pataki, G., Schmieta, S.: The DIMACS library of semidefinite-quadratic-linear programs. <http://dimacs.rutgers.edu/Challenges/Seventh/Instances> (1999). Accessed Dec 2018
34. Permenter, F.: Reduction methods in semidefinite and conic optimization. Ph.D. thesis, MIT. <http://hdl.handle.net/1721.1/114005> (2018). Accessed Dec 2018
35. Permenter, F., Parrilo, P.A.: Finding sparse, equivalent SDPs via minimal-coordinate-projections. In: *IEEE 54th Annual Conference on Decision and Control (CDC)*. IEEE (2015)
36. Schrijver, A.: A comparison of the Delsarte and Lovász bounds. *IEEE Trans. Inf. Theory* **25**(4), 425–429 (1979)
37. Seiler, P.: SOSOPT: A toolbox for polynomial optimization. arXiv preprint [arXiv:1308.1889](https://arxiv.org/abs/1308.1889) (2013)
38. Størmer, E.: *Positive Linear Maps of Operator Algebras*. Springer, Berlin (2013)
39. Størmer, E., Effros, E.G.: Positive projections and Jordan structure in operator algebras. *Math. Scand.* **45**, 127–138 (1979)
40. Sturm, J.F.: Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optim. Methods Softw.* **11**(1–4), 625–653 (1999)
41. Vallentin, F.: Symmetry in semidefinite programs. *Linear Algebra Appl.* **430**(1), 360–369 (2009)
42. Wedderburn, J.M.: On hypercomplex numbers. *Proc. Lond. Math. Soc.* **2**(1), 77–118 (1908)
43. Weisfeiler, B.: *On Construction and Identification of Graphs*. Springer, Berlin (1977)