Automatica 137 (2022) 110125

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

# Decomposed structured subsets for semidefinite and sum-of-squares optimization☆

Jared Miller [a,*], Yang Zheng [b], Mario Sznaier [a], Antonis Papachristodoulou [c]

[a] ECE Department, Northeastern University, Boston, MA 02115, US
[b] Department of Electrical and Computer Engineering, UC San Diego, La Jolla, CA, 92093, US
[c] Department of Engineering Science, University of Oxford, Oxford, OX1 3PJ, UK

ARTICLE INFO

*Article history:*
Received 4 December 2020
Received in revised form 27 July 2021
Accepted 15 November 2021
Available online 6 January 2022

*Keywords:*
Semidefinite programming
Chordal decomposition
Sum of squares
Approximation

ABSTRACT

Semidefinite programs (SDPs) are standard convex problems that are frequently found in control and optimization applications. Interior-point methods can solve SDPs in polynomial time up to arbitrary accuracy, but scale poorly as the size of matrix variables and the number of constraints increases. To improve scalability, SDPs can be approximated with lower and upper bounds through the use of structured subsets (e.g., diagonally-dominant and scaled-diagonally dominant matrices). Meanwhile, any underlying sparsity or symmetry structure may be leveraged to form an equivalent SDP with smaller positive semidefinite constraints. In this paper, we present a notion of *decomposed structured subsets* to approximate an SDP with structured subsets after an equivalent conversion. The lower/upper bounds found by approximation after conversion become tighter than the bounds obtained by approximating the original SDP directly. We apply decomposed structured subsets to semidefinite and sum-of-squares optimization problems with examples of $\mathcal{H}_\infty$ norm estimation and constrained polynomial optimization. An existing basis pursuit method is adapted into this framework to iteratively refine bounds.

© 2021 Elsevier Ltd. All rights reserved.

## 1. Introduction

Semidefinite programs (SDPs) are a class of convex optimization problems that include Linear Programs (LP) and Second-order Cone Programs (SOCP). SDPs are characterized by a symmetric positive semidefinite (PSD) matrix variable subject to affine constraints and a linear cost. Let $\mathbb{S}^n$ be the set of symmetric matrices of size $n$. A more general conic program has a cost matrix $C \in \mathbb{S}^n$, constraint matrices $A_1, \ldots, A_m \in \mathbb{S}^n$, and a constraint vector $b \in \mathbb{R}^m$. Variables are restricted to a proper cone $K$ and dual cone $K^*$, where $\langle \cdot, \cdot \rangle$ denotes the canonical inner product between elements in cones. A conic program has the following primal and dual forms:

$$\begin{aligned} p^* = \min_X \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i, \ i = 1, \ldots, m, \\ & X \in K, \end{aligned} \tag{1}$$

$$\begin{aligned} d^* = \max_{y,Z} \quad & \langle b, y \rangle \\ \text{subject to} \quad & Z + \textstyle\sum_{i=1}^m y_i A_i = C, \\ & y \in \mathbb{R}^m, \ Z \in K^*. \end{aligned} \tag{2}$$

The objectives in (1) and (2) are related by $p^* \geq d^*$, known as weak duality (Boyd & Vandenberghe, 2004). Strong duality, where $p^* = d^*$, may hold under appropriate constraint qualification conditions. SDPs are conic programs with $K = K^* = \mathbb{S}^n_+$, where $\mathbb{S}^n_+$ denotes the set of PSD matrices. The dual form (2) of an SDP is also known as a Linear Matrix Inequality (LMI) (Boyd, El Ghaoui, Feron, & Balakrishnan, 1994).

Classical interior-point methods (IPMs) can solve an SDP to $\epsilon$-accuracy in polynomial time with complexity $\mathcal{O}(n^2m^2+n^3\,m+m^3)$ per iteration (Alizadeh, 1995). Constraint processing methods may be able to remove linearly dependent constraints and therefore reduce the value of $m$. When $m$ is fixed, reducing the cone dimension $n$ can greatly speed up the computational efficiency

☆ J. Miller and M. Sznaier were partially supported by National Science Foundation grants CNS–1646121, CMMI–1638234, IIS–1814631 and ECCS–1808381 and AFOSR grant FA9550-19-1-0005. A. Papachristodoulou was supported in part by the Engineering and Physical Sciences Research Council under project EP/M002454/1. The material in this paper was partially presented at the 21st IFAC World Congress (IFAC 2020), July 12–17, 2020, Berlin, Germany. This paper was recommended for publication in revised form by Associate Editor Kok Lay Teo under the direction of Editor Ian R Petersen.

* Corresponding author.
*E-mail addresses:* miller.jare@northeastern.edu (J. Miller), zhengy@eng.ucsd.edu (Y. Zheng), msznaier@coe.neu.edu (M. Sznaier), antonis@eng.ox.ac.uk (A. Papachristodoulou).

https://doi.org/10.1016/j.automatica.2021.110125
0005-1098/© 2021 Elsevier Ltd. All rights reserved.

of IPMs. Structured subset methods restrict (1) to simple subsets $K_{\text{inner}} \subset \mathbb{S}^n_+ \subset K_{\text{outer}}$ to develop inner and outer approximations resulting in optima $p^*_{\text{outer}} \leq p^* \leq p^*_{\text{inner}}$. These simple subsets include (scaled-) diagonally dominant (DD or SDD) cones (Barker & Carlson, 1975; Boman, Chen, Parekh, & Toledo, 2005). Solving (1) where $X$ is DD is an LP, and the scenario where $X$ is SDD is an SOCP. These simplified formulations yielding possibly conservative bounds are often much faster to solve than the original SDP. We note that polynomial optimization problems can be approximated by a hierarchy of sum-of-squares (SOS) programs, which can be cast as structured SDPs (Parrilo, 2000). The method of structured subsets has also been used to find bounds on polynomial optimization problems when the standard SOS method leads to prohibitively large SDPs; see Majumdar, Ahmadi, and Tedrake (2014).

Structured subset techniques (DD/SDD matrices) ignore any underlying sparsity and reducible properties in the original SDP. For example, a diagonally dominant constraint imposes that each diagonal element is greater than the sum of all absolute values on its row/column. Even if the original problem is sparse (comparatively few elements appear in cost or constraints), the problem approximated by standard DD/SDD constraints will still be dense and may have a slower runtime than the sparse-converted SDP (Vandenberghe, Andersen, et al., 2015). On the other side, an SDP with sparse/symmetric structure may still have overly large blocks after conversion, and these PSD blocks may dominate the computational performance.

This paper presents the notion of *Decomposed Structured Subsets* to find improved lower/upper bounds to semidefinite programs, which exploits problem properties (e.g. sparsity and symmetry) before approximating it with a structured subset. The cones in the decomposition may be mixed, such that large PSD blocks are approximated with structured subsets while small blocks remain PSD to yield tighter bounds than a uniform cone approximation. Some preliminary results were presented in Miller, Zheng, Sznaier, and Papachristodoulou (2020). This paper additionally explores SDPs with multiple kinds of structure, including SDPs that simultaneously have sparsity and symmetry. Decomposed structured subsets are applied to polynomial optimization problems through sum-of-squares approximations. The analysis of the iterative change of basis algorithm to our framework is extended.

The rest of this paper is organized as follows. Section 2 reviews preliminaries of chordal sparsity and approximation by structured subsets. Section 3 merges these two topics into decomposed structured subsets. Section 4 applies decomposed structured subsets to semidefinite programming including an $\mathcal{H}_\infty$ norm estimation problem. Section 5 discusses the application of decomposed structured subsets in SOS optimization. Section 6 summarizes our results and provides directions for future research. The appendices contain details of sparse DD/SDD decompositions and an analysis of problems with both sparsity and symmetry. An extended version of this paper is available at arxiv:1911.12859.

## 2. Preliminaries

### 2.1. Structured subsets

For a fixed matrix size $n$, the nonnegative diagonal, DD, and SDD cones may be described as,

$$\begin{aligned}
\mathcal{D}^n &= \{A \in \mathbb{S}^n : A = \operatorname{diag}(a_1, \ldots, a_n),\, a_i \geq 0\},\\
\mathcal{DD}^n &= \{A \in \mathbb{S}^n :\ a_{ii} \geq \textstyle\sum_{j\neq i} |a_{ij}|,\, i = 1, 2, \ldots, n\}, \qquad (3)\\
\mathcal{SDD}^n &= \{A \in \mathbb{S}^n :\ \exists D \in \mathcal{D}^n,\ DAD \in \mathcal{DD}^n\}.
\end{aligned}$$

These subsets satisfy the following containment relation

$$\mathcal{D}^n \subset \mathcal{DD}^n \subset \mathcal{SDD}^n \subset \mathbb{S}^n_+. \tag{4}$$

Optimizing conic program (1) by setting $K$ equal to these cones with a minimization objective will find bounds:

$$p_{\mathcal{D}} \geq p_{\mathcal{DD}} \geq p_{\mathcal{SDD}} \geq p^*. \tag{5}$$

Another class of structured subsets is the set of factor width matrices. A matrix $M \in \mathbb{S}^n$ has factor-width at most $k$ (denoted as $M \in \mathcal{FW}^n_k$) if there exists a matrix $U$ such that $M = UU^\top$, and each column of $U$ has at most $k$ nonzero entries (Boman et al., 2005). A block factor-width $k$ matrix given a partition of indices is a matrix $M = UU^T$ where each column of $U$ has nonzero elements in at most $k$ sets in the partition (Zheng, Sootla, & Papachristodoulou, 2019).

#### 2.1.1. Change of basis

The change of basis method is an iterative algorithm that sharpens bounds from structured subsets (Ahmadi & Hall, 2017). Given a basis-change matrix $B \in \mathbb{R}^{n\times n}$ and a structured subset cone $K \subset \mathbb{S}^n_+$, the basis-changed cone is $K(B) = \{BQB^\top \mid Q \in K\}$. PSD matrices $X \notin K$ can be made $X \in K(B)$ for some appropriate basis $B$. To start the iterative refinement process, we first solve a conic optimization problem over a structured subset $K$ such as in (1), leading to the iterate $X_0$. The Cholesky decomposition $X_0 = L_0 L_0^\top$ can be used to find the next optimal solution $X_1$:

$$\begin{aligned}
X_1 = \operatorname*{argmin}_X \quad & \langle C, X\rangle \\
\text{subject to} \quad & \langle A_i, X\rangle = b_i,\, i = 1, \ldots, m, \\
& X \in K(L_0).
\end{aligned} \tag{6}$$

Use $X_1 = L_1 L_1^T$, and solve the same problem over $K^n(L_1)$ to find optimal point $X_2$. The cost $\langle C, X\rangle$ of iterate $t$ is upper bounded by the cost at iterate $t-1$, because both $X_{t-1}$ and $X_t$ are members of the feasible set $K(L_{t-1})$ at iteration $t$. We note that this procedure might not converge to the true SDP optimum.

### 2.2. Chordal decomposition

An SDP is sparse if only a few entries of $X \in \mathbb{S}^n_+$ are involved in the cost and constraints. For example, if $C_{jk} = (A_i)_{jk} = 0,\ \forall i = 1, \ldots, m$, the values of $X_{jk}$ and $X_{kj}$ are simply present to ensure that $X \succeq 0$. The sparsity structure of $C$ and $A_i$ can be represented by an undirected graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, termed as the *aggregate sparsity pattern*. The number of nodes $|\mathcal{V}| = n$, and an edge $(j, k) \in \mathcal{E}$ if $C$ or $A_i$ has a nonzero value at index $(j, k)$. We need a few graph-theoretic notions. A cycle of length $k \geq 3$ is a set of vertices $(v_1, v_2, \ldots, v_{k-1}, v_k, v_1)$ with edges $(v_k, v_1)$ and $(v_i, v_{i+1}),\ \forall i = 1, \ldots, k-1$. A chord in a cycle is an edge between two non-consecutive vertices in the cycle. A chordal graph is a graph where each cycle of length 4 or more has a chord (Vandenberghe et al., 2015). A clique $\mathcal{C}$ with cardinality $|\mathcal{C}|$ is a complete subgraph of $\mathcal{G}$, and a maximal clique is a clique that is not contained in another clique. Nonchordal graphs can be rendered chordal by adding new edges, a process known as *chordal extension* (Vandenberghe et al., 2015).

Following notation from Kakimura (2010), the set of sparse symmetric matrices with pattern $\mathcal{G}$ forms a cone $\mathbb{S}^n(\mathcal{E}, 0) = \{X \in \mathbb{S}^n \mid X_{ij} = 0,\ \forall i \neq j,\ (i,j) \notin \mathcal{E}\}$. The sparse PSD cone defined by $\mathcal{E}$ is $\mathbb{S}^n_+(\mathcal{E}, 0) = \mathbb{S}^n(\mathcal{E}, 0) \cap \mathbb{S}^n_+$. The dual cone $[\mathbb{S}^n_+(\mathcal{E}, 0)]^* = \mathbb{S}^n_+(\mathcal{E}, ?)$ is the set of sparse symmetric matrices that admit a PSD completion. For a vector $x \in \mathbb{R}^n$ and a clique $\mathcal{C} \subseteq \mathcal{V}$, there exists a vector $x_{\mathcal{C}} \in \mathbb{R}^{|\mathcal{C}|}$ that selects the entries of $x$ with indices $\mathcal{C}$. Let $E_{\mathcal{C}} \in \mathbb{R}^{|\mathcal{C}|\times n}$ be 0/1 entry selector matrices such that $x_{\mathcal{C}} = E_{\mathcal{C}} x,\ \forall x \in \mathbb{R}^n$. The cones $\mathbb{S}^n_+(\mathcal{E}, ?)$, and $\mathbb{S}^n_+(\mathcal{E}, 0)$ have a decomposable structure if $\mathcal{G}$ is chordal:

**Theorem 1** (*Grone's Theorem (Grone, Johnson, Sá, & Wolkowicz, 1984)*). *Let $\mathcal{G}(\mathcal{V},\mathcal{E})$ be a chordal graph with maximal cliques $\{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_p\}$. Then, $X \in \mathbb{S}^n_+(\mathcal{E}, ?)$ if and only if $X_k = E_{\mathcal{C}_k} X E^{\mathsf{T}}_{\mathcal{C}_k} \in \mathbb{S}^{|\mathcal{C}_k|}_+$ for $k = 1, \ldots, p$.*

**Theorem 2** (*Agler's Theorem (Agler, Helton, et al., 1988)*). *Let $\mathcal{G}(\mathcal{V},\mathcal{E})$ be a chordal graph with maximal cliques $\{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_p\}$. Then, $Z \in \mathbb{S}^n_+(\mathcal{E}, 0)$ if and only if $\exists Z_k \in \mathbb{S}^{|\mathcal{C}_k|}_+, k = 1, \ldots, p$ such that $Z = \sum_{k=1}^p E^{\mathsf{T}}_{\mathcal{C}_k} Z_k E_{\mathcal{C}_k}$.*

Grone's theorem can split problem (1) with $K = \mathbb{S}^n_+(\mathcal{E}, ?)$,

$$\begin{aligned} \min_X \quad & \langle C, X\rangle \\ \text{subject to} \quad & \langle A_i, X\rangle = b_i, i = 1, \ldots, m, \\ & E_{\mathcal{C}_k} X E^{\mathsf{T}}_{\mathcal{C}_k} \in \mathbb{S}^{|\mathcal{C}_k|}_+, k = 1, \ldots, p. \end{aligned} \tag{7}$$

Problem (7) reduces the size of the maximum PSD constraint. In numerical SDP solvers, equality constraints are introduced to enforce consistency between variables on clique overlaps (Vandenberghe et al., 2015). Criteria for merging cliques to improve computational performance are discussed in Garstka, Cannon, and Goulart (2020).

## 3. Decomposed structured subsets

This section combines decomposition methods and structured subsets into *decomposed structured subsets*. As a motivating example, consider a matrix $M(a, b)$:

$$M(a, b) = \begin{bmatrix} 1 & \frac{1}{2}+a & ? & ? \\ \frac{1}{2}+a & 2 & -2a & a+b \\ ? & -2a & 5 & \frac{b}{2} \\ ? & a+b & \frac{b}{2} & 2 \end{bmatrix}, \tag{8}$$

where '?' denotes unspecified entries. Since the sparsity pattern of this matrix is chordal, Grone's Theorem guarantees that $M(a, b) \in \mathbb{S}^4_+(\mathcal{E}, ?)$ if and only if

$$M_1(a, b) = \begin{bmatrix} 2 & -2a & a+b \\ -2a & 5 & b/2 \\ a+b & b/2 & 2 \end{bmatrix} \succeq 0,$$

$$M_2(a, b) = \begin{bmatrix} 1 & 1/2+a \\ 1/2+a & 2 \end{bmatrix} \succeq 0.$$

The blocks $M_1(a, b)$ and $M_2(a, b)$ correspond to the maximal cliques $\mathcal{C}_1 = \{2, 3, 4\}$ and $\mathcal{C}_2 = \{1, 2\}$. The structure of (8) defines an affine slice of the PSD completion cone $\mathbb{S}^4_+(\mathcal{E}, ?)$, as visualized in the black set in each panel of Fig. 1. The red region of the left panel is the set where '?' may be chosen to ensure that $M(a, b) \in \mathcal{DD}^4$. The blue set in the left panel of is the region where $M_1(a, b) \in \mathcal{DD}^3$ and $M_2(a, b) \in \mathcal{DD}^3$. The blue set contains the red set, because imposing that $M(a, b)$ is $\mathcal{DD}^4$-completable adds a constraint on the previously arbitrary entries '?', yielding $1 \geq |1/2+a| + |?_{13}| + |?_{14}|$. The same behavior is observed for the SDD cone in the right panel of Fig. 1, with the yellow set where $M(a, b)$ is $\mathcal{SDD}^4$-completable is contained within in the brown set describing $M_1(a, b) \in \mathcal{SDD}^3, M_2(a, b) \in \mathcal{SDD}^2$. This notion of structured subsets over cliques and containment will be made rigorous in this section.

### 3.1. Definition of decomposed structured subsets

A clique edge cover of a graph $\mathcal{G}(\mathcal{V},\mathcal{E})$ is a set of subsets $\{\mathcal{C}_k\}_{k=1}^p$ such that every maximal clique of $\mathcal{G}$ is contained in at least one clique $\mathcal{C}_k$. Clique edge covers allow for clique extensions and merges for possibly non-chordal graphs.

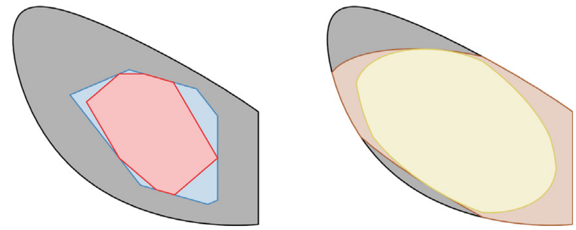


**Fig. 1.** Regions of the $(a, b)$ plane for which the matrix $M(a, b)$ from Eq. (8) is $\mathcal{DD}^4$-completable, $\mathcal{DD}^4(\mathcal{E}, ?)$, or $\mathbb{S}^4_+(\mathcal{E}, ?)$ (Left), and $\mathcal{SDD}^4$-completable, $\mathcal{SDD}^4(\mathcal{E}, ?)$, or $\mathbb{S}^4_+(\mathcal{E}, ?)$ (Right). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Definition 1.** We define sparse DD and SDD matrices as:

$$\mathcal{DD}^n(\mathcal{E}, 0) = \mathbb{S}^n(\mathcal{E}, 0) \cap \mathcal{DD}^n,$$
$$\mathcal{SDD}^n(\mathcal{E}, 0) = \mathbb{S}^n(\mathcal{E}, 0) \cap \mathcal{SDD}^n.$$

These sparse matrices obey the containment: $\mathcal{DD}^n(\mathcal{E}, 0) \subset \mathcal{SDD}^n(\mathcal{E}, 0) \subset \mathbb{S}^n_+(\mathcal{E}, 0)$. The following result is proven in Appendix A:

**Proposition 1.** *Let $\mathcal{G}(\mathcal{V},\mathcal{E})$ be a graph with a clique edge cover $\{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_p\}$. Then,*

*(1) $Z \in \mathcal{DD}^n(\mathcal{E}, 0)$ if and only if*
$$Z = \textstyle\sum_{k=1}^p E^{\mathsf{T}}_{\mathcal{C}_k} Z_k E_{\mathcal{C}_k}, \quad Z_k \in \mathcal{DD}^{|\mathcal{C}_k|}, \; k = 1, \ldots, p.$$

*(2) $Z \in \mathcal{SDD}^n(\mathcal{E}, 0)$ if and only if*
$$Z = \textstyle\sum_{k=1}^p E^{\mathsf{T}}_{\mathcal{C}_k} Z_k E_{\mathcal{C}_k}, \quad Z_k \in \mathcal{SDD}^{|\mathcal{C}_k|}, \; k = 1, \ldots, p.$$

Motivated by Theorems 1 and 2, and Proposition 1, we let $\mathcal{E}$ be a sparsity pattern, $\mathcal{K} = \{K_k\}_{k=1}^p$ be a set of cones corresponding to a clique edge cover $\mathcal{C}_1, \ldots, \mathcal{C}_p$, where each individual cone $K_k$ is some structured subset in $\mathbb{S}^{|\mathcal{C}_k|}$.

**Definition 2. Decomposed Structured Subsets.** Given clique cones $\mathcal{K}$, we define:

$$\mathcal{K}(\mathcal{E}, 0) := \Big\{ Z \in \mathbb{S}^n \mid Z = \textstyle\sum_{k=1}^p E^{\mathsf{T}}_{\mathcal{C}_k} Z_k E_{\mathcal{C}_k}, \; Z_k \in K_k, \; k = 1, \ldots, p \Big\}, \tag{9}$$

$$\mathcal{K}(\mathcal{E}, ?) := \Big\{ X \in \mathbb{S}^n \mid E_{\mathcal{C}_k} X E^{\mathsf{T}}_{\mathcal{C}_k} \in K_k, \; k = 1, \ldots, p \Big\}.$$

Note that the sets $\mathcal{DD}^n(\mathcal{E}, 0)$ and $\mathcal{SDD}^n(\mathcal{E}, 0)$ are special decomposed structured subsets for any sparsity pattern, and the sets $\mathbb{S}^n_+(\mathcal{E}, 0)$, and $\mathbb{S}^n_+(\mathcal{E}, ?)$ are special decomposed structured subsets for chordal sparsity patterns.

### 3.2. Containment analysis

**Definition 3.** A sparse matrix $X \in \mathbb{S}^n(\mathcal{E}, 0)$ has a $K$-**completion** for a structured subset $K \subseteq \mathbb{S}^n_+$ if there exists an $\bar{X} \in K$ such that $X_{ij} = \bar{X}_{ij}, \; \forall (i, j) \in \mathcal{E}$.

For a structured subset $K \subseteq \mathbb{S}^n_+$ and a sparsity pattern $\mathcal{G}(\mathcal{V},\mathcal{E})$, the set of $K$-completable matrices with pattern $\mathcal{E}$ is contained within $K(\mathcal{E}, ?)$. This containment is illustrated in Fig. 1 for the cones $\mathcal{DD}^4$ (left) and $\mathcal{SDD}^4$ (right). The notion of decomposition structured subsets (9) gives more freedom to choose the individual cones $K_k$.

**Definition 4.** For a graph $\mathcal{G}(\mathcal{V},\mathcal{E})$ with clique cover $\mathcal{C}_1, \ldots, \mathcal{C}_p$, let $\mathcal{K} = \{K_k\}_{k=1}^p$ and $\tilde{\mathcal{K}} = \{\tilde{K}_k\}_{k=1}^p$ be two sets of clique-cones. The partial ordering $\subseteq$ is defined as:

$$\mathcal{K} \subseteq \tilde{\mathcal{K}} \quad \text{if and only if} \quad K_k \subseteq \tilde{K}_k \quad \forall\, k = 1 \ldots p.$$

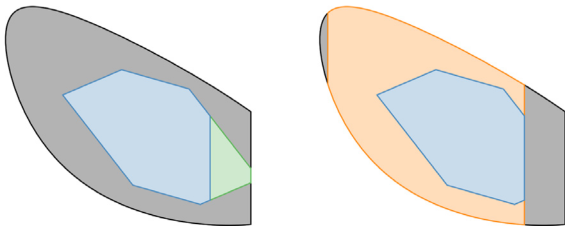


**Fig. 2.** Mixing cones expands feasibility regions of the $(a, b)$ plane for which the matrix $M(a, b) \in \mathcal{K}(\mathcal{E}, ?)$ in (8). Left: $\mathcal{K}^1 = \{\mathcal{DD}^3, \mathbb{S}^2_+\}$; Right: $\mathcal{K}^2 = \{\mathbb{S}^3_+, \mathcal{DD}^2\}$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Remark 1.** If clique-cones $\mathcal{K} \subseteq \tilde{\mathcal{K}}$ for a clique edge cover $\mathcal{C}_1, \dots, \mathcal{C}_p$ of a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, then by definition, we have

$$\mathcal{K}(\mathcal{E}, 0) \subseteq \tilde{\mathcal{K}}(\mathcal{E}, 0) \quad \text{and} \quad \mathcal{K}(\mathcal{E}, ?) \subseteq \tilde{\mathcal{K}}(\mathcal{E}, ?).$$

The relationship above is simple, yet has useful implications in semidefinite optimization. In particular, the cone $\mathcal{DD}^n(\mathcal{E}, 0)$ has the same cone on each clique $\mathcal{C}_k$ with $E_{\mathcal{C}_k} X E^{\mathsf{T}}_{\mathcal{C}_k} = X_k \in \mathcal{DD}^{|\mathcal{C}_k|}$. Mixing cones with $\mathcal{K} : K_k \supseteq \mathcal{DD}^{|\mathcal{C}_k|}, \forall k = 1 \dots p$ where not all $K_k = \mathcal{DD}^{|\mathcal{C}_k|}$ will form a cone $\mathcal{K}(\mathcal{E}, 0) \supset \mathcal{DD}(\mathcal{E}, 0)$. Mixing cones therefore results in cones closer to $\mathbb{S}^n_+(\mathcal{E}, 0)$. Similar statements hold for $\mathbb{S}^n_+(\mathcal{E}, ?)$. This allows us to get better lower and upper bounds for problem (1). In Fig. 2, the green set on the left panel and the orange set on the right panel illustrate mixed cones. The green set has $M_1(a, b) \in \mathcal{DD}^3$, $M_2(a, b) \in \mathbb{S}^2_+$, and the feasibility set is slightly expanded to the right. The orange set on the right panel with $M_1(a, b) \in \mathbb{S}^3_+$, $M_2(a, b) \in \mathcal{DD}^2$ cuts a large leftward swath of the PSD feasibility region.

Decomposed structured subsets can be posed over dual cones that are larger than the PSD cone.

**Proposition 2.** *Let $\mathcal{K} = \{K_k\}_{k=1}^p$ be a set of cones with dual $\mathcal{K}^* = \{K_k^*\}_{k=1}^p$ where each $K_k \subseteq \mathbb{S}^{|\mathcal{C}_k|}_+$, and $\mathcal{C}_k$ is a set of max. cliques for $\mathcal{E}$. Then, we have*

$$[\mathcal{K}(\mathcal{E}, 0)]^* = \mathcal{K}^*(\mathcal{E}, ?), \tag{10}$$

$$[\mathcal{K}(\mathcal{E}, ?)]^* \supseteq \mathcal{K}^*(\mathcal{E}, 0). \tag{11}$$

**Proof.** For the equivalence in (10), recall the definition of $\mathcal{K}(\mathcal{E}, 0)$ and $\mathcal{K}(\mathcal{E}, ?)$ in (9). We verify that $[\mathcal{K}(\mathcal{E}, 0)]^*$ is,

$$\begin{aligned}
\mathcal{K}(\mathcal{E}, 0)t^* &= \{M \in \mathbb{S}^n \mid \langle M, N \rangle \geq 0, \forall N \in \mathcal{K}(\mathcal{E}, 0)\} \\
&= \left\{M \in \mathbb{S}^n \mid \left\langle M, \textstyle\sum_{k=1}^p E^{\mathsf{T}}_{\mathcal{C}_k} N_k E_{\mathcal{C}_k} \right\rangle \geq 0, \ \forall N_k \in K_k \right\} \\
&= \left\{M \in \mathbb{S}^n \mid \textstyle\sum_{k=1}^p \left\langle E_{\mathcal{C}_k} M E^{\mathsf{T}}_{\mathcal{C}_k}, N_k \right\rangle \geq 0, \ \forall N_k \in K_k \right\} \\
&= \left\{M \in \mathbb{S}^n \mid E_{\mathcal{C}_k} M E^{\mathsf{T}}_{\mathcal{C}_k} \in K_k^*, \ k = 1, \dots, p \right\},
\end{aligned}$$

where the last set is the definition of $\mathcal{K}^*(\mathcal{E}, ?)$. The last equality used the following fact: if any $E_{\mathcal{C}_t} M E^{\mathsf{T}}_{\mathcal{C}_t} \notin K_t^*$ for some $t$, we can choose $N_t \in K_t$ such that $\langle E_{\mathcal{C}_t} M E^{\mathsf{T}}_{\mathcal{C}_t}, N_t \rangle < 0$. Now, by choosing $N_k = 0 \in K_k, k \neq t$, we have

$$\textstyle\sum_{k=1}^p \left\langle E_{\mathcal{C}_k} M E^{\mathsf{T}}_{\mathcal{C}_k}, N_k \right\rangle = \langle E_{\mathcal{C}_t} M E^{\mathsf{T}}_{\mathcal{C}_t}, N_t \rangle < 0,$$

which contradicts line 4. Thus, we must have $E_{\mathcal{C}_k} M E^{\mathsf{T}}_{\mathcal{C}_k} \in K_k^*, k = 1, \dots, p$. We now prove the containment in (11). Given any $M \in \mathcal{K}^*(\mathcal{E}, 0)$, we show $M \in [\mathcal{K}(\mathcal{E}, ?)]^*$. By definition (6), there exists $M_k \in K_k^*, k = 1, \dots, p$, such that $M = \sum_{k=1}^p E^{\mathsf{T}}_{\mathcal{C}_k} M_k E_{\mathcal{C}_k}$. We now verify that $\forall N \in \mathcal{K}(\mathcal{E}, ?)$

$$\begin{aligned}
\langle M, N \rangle &= \left\langle \textstyle\sum_{k=1}^p E^{\mathsf{T}}_{\mathcal{C}_k} M_k E_{\mathcal{C}_k}, N \right\rangle \\
&= \textstyle\sum_{k=1}^p \left\langle M_k, E_{\mathcal{C}_k} N E^{\mathsf{T}}_{\mathcal{C}_k} \right\rangle \\
&= \textstyle\sum_{k=1}^p \langle M_k, N_k \rangle \geq 0,
\end{aligned}$$

where the last inequality used the definition of $N \in \mathcal{K}(\mathcal{E}, ?)$. Therefore, $M \in [\mathcal{K}(\mathcal{E}, ?)]^*$ and $\mathcal{K}^*(\mathcal{E}, 0) \subseteq [\mathcal{K}(\mathcal{E}, ?)]^*$. □

**Table 1**
Cost vs. subset and pattern in Fig. 3.

| | $K$ | $\mathcal{K}(\mathcal{E}_F, ?)$ | $\mathcal{K}(\mathcal{E}, ?)$ |
|---|---|---|---|
| $\mathcal{DD}$ | Inf. | Inf. | Inf. |
| $B_1$ | 64.5 | 34.7 | 19.4 |
| $B_2$ | 51.4 | 27.1 | 13.9 |
| $B_5$ | 32.1 | 15.0 | 5.34 |
| $B_{10}$ | 20.8 | 7.10 | −1.23 |
| $\mathbb{S}_+$ | −1.23 | −1.23 | −1.23 |

**Remark 2.** The equality $\mathcal{K}^*(\mathcal{E}, 0) = [\mathcal{K}(\mathcal{E}, ?)]^*$ will hold if the clique cover $\mathcal{C}_k$ of $\mathcal{E}$ are disjoint: $\mathcal{C}_k \cap \mathcal{C}_{k'} = \varnothing, \ \forall k \neq k'$.

## 4. Applications to semidefinite optimization

In this section, we develop inner and outer approximations of SDPs using the notion of decomposed structured subsets, and apply $\mathcal{H}_\infty$ norm estimation to networked systems. All code is publicly available at https://github.com/soc-ucsd/SDPfw within the folder `decomposed_structured_subsets`.

### *4.1. Decomposed structured subsets in SDPs*

Let $K \subset \mathbb{S}^n_+$ be a structured subset. As discussed in the introduction, bounds $p^*_{\text{inner}}$ and $p^*_{\text{outer}}$ can be found by constraining $X$ in (1) to $K$ and $K^*$ respectively. This procedure can also be done for decomposed structured subsets. Let $\mathcal{G}(\mathcal{V}, \mathcal{E})$ be a sparsity pattern, $\mathcal{C} = \{\mathcal{C}_k\}_{k=1}^p$ be a clique cover of $\mathcal{E}$, and $\mathcal{K}$ be a cone-set over $\mathcal{C}$. The relationship $\mathcal{K} \subseteq \{\mathbb{S}^{|\mathcal{C}_k|}_+\}_{k=1}^p \subseteq \mathcal{K}^*$ holds with respect to the partial ordering $\subseteq$ from Definition 4. Inner and outer approximations for problem (1) over $X \in \mathbb{S}^n_+(\mathcal{E}, ?)$ may be found by constraining $X$ to $\mathcal{K}(\mathcal{E}, ?)$ and $\mathcal{K}^*(\mathcal{E}, ?)$ respectively. The same procedure can be done to approximate $\mathbb{S}^n_+(\mathcal{E}, 0)$ with $\mathcal{K}(\mathcal{E}, 0)$ from inside and $\mathcal{K}^*(\mathcal{E}, 0)$ from outside.

Fig. 3 visualizes the aggregate sparsity pattern of a random SDP with 80 equality constraints, where each of the 15 blocks has size 10 and the arrowhead has width 10. The original SDP has $X \in \mathbb{S}^{160}_+$, and a chordal decomposition has $X_k \in \mathbb{S}^{20}_+$ for each clique where cliques overlap in the $10 \times 10$ bottom right corner (blue pattern $\mathcal{E}$). A coarser chordal decomposition is the union of blue and magenta blocks in Fig. 3 (fill-in $\mathcal{E}_F$), which has block sizes $(\mathbb{S}^{50}_+)^2 \times \mathbb{S}^{40}_+ \times (\mathbb{S}^{30}_+)^2$ that are each equal in the bottom right corner. Clique consistency for $\mathcal{E}_F$ and $\mathcal{E}$ adds 220 and 770 equality constraints respectively.

Cost values of this SDP and its approximants are recorded in Table 1. Rows are different structured subsets and the columns apply the structured subsets: imposing that $X \in K$ and that the cliques of $X_k \in K_k$, where cliques are set based on the graphs $\mathcal{E}_F$ and $\mathcal{E}$. We introduce shorthand $B_k$ as a cone of block factor-width 2 matrices where each block has $k$ components (so $B_1 = \mathcal{SDD}$) and membership constraints in $\mathbb{S}^{2k}_+$ are imposed. By Agler's theorem, all entries of $K = \mathbb{S}^{160}_+$ have the same optima. All entries $K = \mathcal{DD}^{160}$ are infeasible, and objectives decrease towards the bottom right corner of the table as expected in the above containment analysis.

Table 1 demonstrates that merging blocks together may degrade the resultant approximation quality. Even well-chosen merges that speed up program execution such as in Garstka et al. (2020) may worsen the approximated SDP bound.

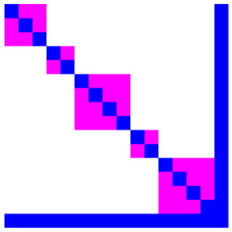


**Fig. 3.** Block arrow sparsity pattern and extension. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

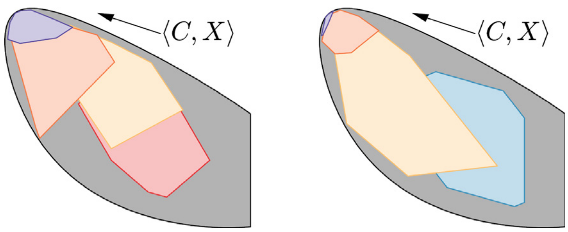



**Fig. 4.** Decomposed vs. Standard Change of Basis on (8). Left: Start from $\mathcal{DD}^4$-completable; Right: Start from $\mathcal{DD}^4(\mathcal{E}, ?)$.

**Remark 3** (*Certifying Optimality*)**.** A structured subset approximation yields the same optimum as the original SDP if an eigenvalue test is satisfied. Assume that a structured subset approximation with $K \subseteq \mathbb{S}^n_+$ yields an optimal primal–dual triple $(X, y, Z) \in (K, \mathbb{R}^m, K^*)$(Ahmadi, Dash, & Hall, 2017). The approximation is tight if the dual matrix satisfies $Z \in \mathbb{S}^n_+$ (dual feasibility of the original SDP). For SDP lower bound programs with decomposed structured subsets, the clique cones $\mathcal{K}$ have $K_k \supseteq \mathbb{S}^{|\mathcal{C}_k|}_+$. Tightness is certified if $X_k \in \mathbb{S}^{|\mathcal{C}_k|}_+$, $\forall k = 1, \dots, p$. Upper bounds have $K_k \subseteq \mathbb{S}^{|\mathcal{C}_k|}_+$. For each clique $\mathcal{C}_k$ in the $\mathcal{K}$, check if the corresponding dual block $Z_k \in \mathbb{S}^{|\mathcal{C}_k|}_+$. The dual clique blocks $Z_k$ can be obtained by computing $Z = C - \sum_{i=1}^m y_i A_i$.

### 4.2. Decomposed change of basis

Decomposed structured subsets are compatible with the change of basis algorithm as reviewed in Section 2.1.1. Assume that $X_0 \in \mathcal{K}(\mathcal{E}, ?)$ is a solution to Problem (1). Define Cholesky factorization matrices $L_k^0$ for each clique $k = 1 \dots p$ such that $L_k^0 L_k^{0\intercal} = E_{\mathcal{C}_k} X_0 E_{\mathcal{C}_k}^\intercal$. The next iteration of the change of basis algorithm will solve

$$\begin{aligned} X_1 = \arg\min_X \quad & \langle C, X\rangle \\ \text{subject to} \quad & \langle A_i, X\rangle = b_i, \ i = 1, 2, \dots, m, \\ & E_{\mathcal{C}_k} X E_{\mathcal{C}_k}^\intercal \in K_k(L_k^0), \ k = 1, \dots, p. \end{aligned} \tag{12}$$

The solution to (12) can be used to find a new set of factor matrices $L_k^1$ by finding $L_k^1 L_k^{1\intercal} = E_{\mathcal{C}_k} X_1 E_{\mathcal{C}_k}^\intercal$. Each clique $\mathcal{C}_k$ is described by basis $L_k$, and different bases may describe the same elements of $X$ on clique-overlaps.

**Remark 4.** Performing a decomposed change-of-basis over $K^n(\mathcal{E}, ?)$ will result in a lower cost as compared to applying change-of-basis over $K^n$ at the first iteration. No conclusions can be drawn after the first iteration. In experiments, the cost sequence obtained from performing change-of-basis over $K^n$ remains above $K^n(\mathcal{E}, ?)$'s cost sequence.

Fig. 4 illustrates the change of basis technique on the example in 3 optimizing a cost function $\langle C, X\rangle$ (direction of black arrow).

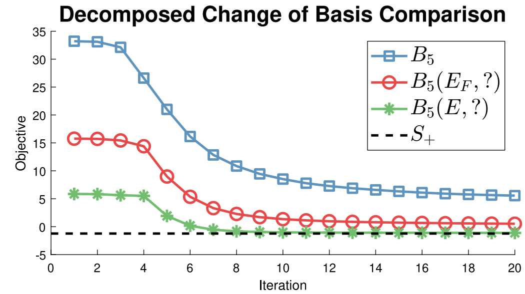



**Fig. 5.** Change of basis on Block Arrow SDP. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

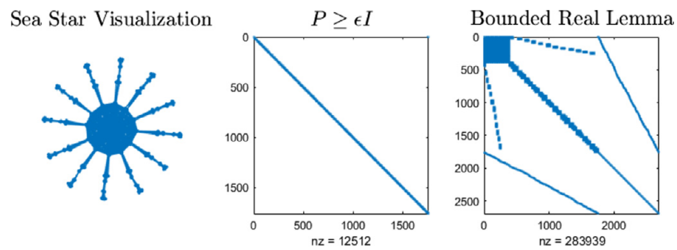



**Fig. 6.** Sea Star topology and LMI constraint sparsity.

The left plot is a change of basis starting from $\mathcal{DD}^4$, while the right plot is a decomposed change of basis on $\mathcal{DD}^4(\mathcal{E}, ?)$. At the second iteration, the change-of-basis on the $\mathcal{DD}$ set (left) is $-2.50$, while the $\mathcal{DD}(\mathcal{E}, ?)$ set (right) is $-3.02$. Fig. 5 shows the output of the change of basis algorithm for the cone $B_5^{160}$ on the block arrow system shown in Fig. 3. Over the course of 20 iterations, the basis-changed cone starting with $B_5^{160}(\mathcal{E}, ?)$ (green curve) eventually matches the SDP optimum.

### 4.3. $\mathcal{H}_\infty$ Norm estimation for networked systems

Consider a dynamical system $G(s)$ with state-space form,

$$\dot{x} = Ax + Bu, \qquad y = Cx + Du.$$

The $\mathcal{H}_\infty$ norm of $G(s)$ is the supremum over frequencies $\omega$ of the maximum singular value of $G(j\omega)$. The norm $\|G\|_\infty$ is finite when $A$ is Hurwitz. The Bounded Real Lemma can be used to find upper bounds on $\|G\|_\infty$:

**Theorem 3** (*Bounded Real Lemma (Boyd et al., 1994)*)**.** *The following statements are equivalent:*

*(1)* $\|G\|_\infty < \gamma$;
*(2) There exists a* $P \succ 0$ *such that*

$$\begin{bmatrix} PA + A^\intercal P + C^\intercal C & P^\intercal B + C^\intercal D \\ B^\intercal P + D^\intercal C & -\gamma^2 I + D^\intercal D \end{bmatrix} \prec 0. \tag{13}$$

We propose the 'sea star' networked system as a testbed for applying decomposed structured subsets to $\mathcal{H}_\infty$ norm estimation. Each agent obeys linear dynamics with $n_i$ states, $m_i$ inputs, $d_i$ outputs. Agents in the sea star network either lie in the densely connected head or in one of the arms. The left panel of Fig. 6 shows a sea star network with 12 arms and 70 agents in the head. Each arm has 2 densely connected 'knuckles', where in each knuckle 4 agents connect to the previous knuckle (or head) and 4 agents connect to the next knuckle. Agent dynamics are chosen such that the global dynamics are open-loop stable.

The $\mathcal{H}_\infty$-norm of $G(s)$ can be estimated through minimization of $\gamma^2$ on the Bounded Real Lemma. The LMI enforcing the

**Table 2**
Time to find $\gamma$ by upper bound $K$ (minutes).

Upper bound time (min.) $\gamma = 1.137$

| PSD threshold | $\mathcal{DD}$ | $B_1$ | $B_3$ | $B_5$ | $B_8$ | $B_{15}$ | $B_{30}$ | $B_{55}$ |
|---|---|---|---|---|---|---|---|---|
| 0 | | | 2.27 | 1.84 | 2.17 | 3.10 | 9.49 | 35.57 |
| 11 | | | 2.33 | 1.84 | 2.18 | 3.11 | 9.62 | 35.02 |
| 60 | | 9.54 | 2.71 | 2.33 | 2.01 | 3.95 | 9.50 | 34.75 |
| 100 | | 11.88 | 5.21 | 4.64 | 4.49 | 5.02 | 9.60 | 33.99 |

Cone Complexity

**Table 3**
Minutes to find constrained LR lower bound over $\mathcal{K}^*$.

Constrained $f^* = 4939.1$ time (min.)

| PSD threshold | $\mathcal{DD}$ | $\mathcal{SDD}$ | $B_2$ | $B_3$ | $B_5$ | $B_6$ | $B_{10}$ | $B_{15}$ | $B_{20}$ | $B_{30}$ | $B_{40}$ | $\mathbb{S}_+$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | | | | | 35.10 | 36.05 | 38.79 | 45.42 | 66.53 | 63.39 | 61.00 | 59.12 |
| 5 | | | | | 35.45 | 36.14 | 38.40 | 45.82 | 59.84 | 63.45 | 59.87 | 59.48 |
| 12 | | | | | 35.03 | 36.11 | 38.40 | 45.33 | 57.90 | 63.35 | 61.09 | 58.96 |
| 45 | 38.19 | 60.81 | 63.15 | 31.79 | 43.64 | 43.10 | 45.94 | 45.42 | 60.07 | 63.44 | 60.01 | 59.62 |
| 100 | 43.79 | 63.87 | 31.30 | 61.12 | 40.15 | 57.75 | 42.25 | 41.88 | 61.03 | 63.29 | 59.91 | 59.43 |

Cone Complexity

Bounded Real Lemma has two PSD variables corresponding to the matrix $P \succeq 0$ and Eq. (13) with size 1760 and 2691 respectively. A block-diagonal choice of $P$ such that each agent's block in $P$ has a size equal to order of that agent's dynamics will be compatible with the sparsity of the network, but may yield conservative $\mathcal{H}_\infty$ norm bounds (Zheng, Kamgarpour, Sootla, & Papachristodoulou, 2018). The block-diagonal $P$ is displayed on the center panel of Fig. 6, and the Bounded Real Lemma network constraint is on the right panel. The chordal decomposition $\mathcal{H}_\infty$ sea-star LMI's aggregate sparsity pattern possesses a giant clique of size 387 (agents in head), a set of cliques of size 37–90 (knuckles), and a run of cliques of sizes 1–11. For the rest of this section, cone dimensions when referring to cones such as $\mathbb{S}^n_+$ and $K(\mathcal{E}, ?)$ may be omitted for ease of explanation.

Table 2 displays the time required to find $\mathcal{H}_\infty$-norm upper bounds for the sea star system. The columns of each table are structured subsets $K$ which includes $\mathcal{DD}$, $\mathbb{S}_+$, and block-factor-width 2 matrices $B_q$ for integer $q$. The rows are size thresholds that define the decomposed structured subset $\mathcal{K}$. For an example size threshold of 11, all cliques with $|\mathcal{C}| \leq 11$ remain PSD while cliques with $|\mathcal{C}| > 11$ are restricted to $K$. Displayed values in the tables achieve the same bound as the $\mathbb{S}_+$ case with a certification from Section 4.1. Non-displayed values on Table 2 did not achieve the optimal $\gamma = 1.137$: the cone $\mathcal{DD}$ with thresholds 0, 11 was primal infeasible. The fastest time to obtain the optimal $\gamma$ was the program with cone $B_5$ and size threshold of 0 ($B_5(\mathcal{E}, 0)$). In this case, $B_5(\mathcal{E}, 0)$ offered the best compromise between cone complexity and number of new equality constraints. All experiments were written in Matlab R2018a and performed on Mosek (Andersen & Andersen, 2000) on a Intel i7 CPU with a clock frequency of 2.7 GHz and 16.0 GB of RAM.

## 5. Applications to polynomial optimization

### 5.1. Preliminaries for polynomial optimization

A polynomial optimization problem (POP) may be approximated by semidefinite programming (Lasserre, 2010). A basic semialgebraic set $\mathbb{K}$ is defined by a finite number of bounded-degree polynomial constraints:

$$\mathbb{K} = \{x \in \mathbb{R}^N \mid g_i(x) \geq 0, \quad h_j(x) = 0\}. \tag{14}$$

The task of minimizing a polynomial $p(x) \in \mathbb{R}[x]_{\leq d}$ of bounded degree $d$ with $x \in \mathbb{K}$ is equivalent to solving:

$$p^* = \max_{\gamma \in \mathbb{R}} \gamma, \qquad \text{subject to} \quad p(x) - \gamma \geq 0, \quad \forall x \in \mathbb{K}. \tag{15}$$

Polynomial nonnegativity constraints are generically NP hard, and sum-of-squares (SOS) methods offer convex relaxations by solving SDPs. A polynomial $q(x)$ is SOS if there exists a monomial map $v(x)$ and a Gram matrix $Q \succeq 0$ such that $q(x) = v(x)^\mathsf{T} Q v(x)$. The SOS cone $\Sigma[x]$ is a subcone of all nonnegative polynomials. The Putinar Positivstellensatz (Putinar, 1993) yields an SOS relaxation to Eq. (15):

$$p^* = \max_{\gamma \in \mathbb{R},\ \sigma(x) \in \Sigma[x],\ \zeta_i(x) \in \Sigma[x],\ \phi_j(x) \in \mathbb{R}[x]} \gamma \tag{16}$$
$$\text{subject to} \quad p(x) - \gamma = \sigma(x) + \textstyle\sum_i \zeta_i(x) g_i(x) + \sum_j \phi_j(x) h_j(x).$$

Eq. (16) is an SDP when restricted to polynomials $\sigma(x)$, $\zeta_i(x)$, $\phi_j(x)$ of bounded degree $2d$. Solving (16) at degree $d$ will result in a bound $p^*_d \leq p^*$, and the sequence $p^*_d \leq p^*_{d+1} \leq p^*_{d+2} \leq \cdots$ is an increasing sequence of lower bounds to $p^*$ (Theorem 5.6 and 4.1 of Lasserre, 2010). If $\mathbb{K}$ satisfies an Archimedean condition, then $p^*$ will be reached at a finite degree $d$. The size of the Gram matrix Q scales as $O(N^d)$.

Two methods of defining and exploiting sparsity in POPs include Correlative Sparsity (Waki, Kim, Kojima, & Muramatsu, 2006) and Term Sparsity (Wang, Magron, & Lasserre, 2021b). Term sparsity may be combined with Correlative Sparsity to maximize performance (CS-TSSOS) (Wang, Magron, & Lasserre, 2021a).

### 5.2. Decomposed structured subsets for POPs

Decomposed structured subsets can be integrated into polynomial optimization. A polynomial $p(x) \in SOS$ may exhibit sparsity in which the Gram matrix $Q \in \mathbb{S}^n_+(\mathcal{E}, 0)$ for a pattern $\mathcal{G}(\mathcal{V}, \mathcal{E})$. The matrix $Q$ may be restricted to a decomposed structured subset $\mathcal{K}^*(\mathcal{E}, 0)$ to form approximations of the SOS lower bound. An example of decomposed structured subsets for polynomial optimization is the minimization of the $N$-variate polynomial $f(x) = f_Q(x) + f_R(x)$, where $f_Q(x) = x^\mathsf{T}_{1:N/6} A x_{1:N/6}$ and,

$$f_R(x) = \textstyle\sum_{i=1}^{N-3} 10(x_{i+2} + 2x_{i+1} - x_i^2)^2 + (1 - x_i - x_{i+3})^2.$$

$f_Q(x)$ is a quadric where $A$ is the Lehmer matrix defined as $A_{ij} = \min(i/j, j/i)$, and $f_R(x)$ is a Rosenbrock function. The Rosenbrock function $f_R$ is highly sparse and multiplies together variables $(x_i, x_j)$ only when $|i - j| \leq 3$, while the LR function has a block containing $(x_{1:N/6})$ from $f_Q$.

The cliques of largest size for minimizing the LR function $f(x)$ with $N = 120$ over $\mathbb{K} = [1, 2]^{120}$ in the $d = 4$ CS-TSSOS relaxation with option 'clique' are 124, 80, 50, 50, 50. All other cliques are of size 40 or less, and there are 10647 cliques in total (including 5473 cliques of size 6 and 5727 cliques of size 1). Table 3 displays the time taken in minutes to provide tight lower bounds. The SDP optimum of 4939.1 is attained in 31.8 min with the cone $B^*_3$ and size threshold 45, compared to 63.4 min on the full SDP $\mathbb{S}_+(\mathcal{E}, ?)$. The $\mathbb{S}_+$ column timing is reported for multiple runs.

## 6. Conclusions

Structured subsets can be used to find upper and lower bounds of SDP optima. Decomposition methods may be able to convert large PSD constraints into smaller PSD blocks. This paper combines the two methods into decomposed structured subsets. Properties of these subsets are analyzed with their bound quality, and the facility to mix cones adds additional flexibility. Improved approximations are demonstrated on $\mathcal{H}_\infty$ norm and polynomial optimization problems. Future directions include applying these techniques to network $\mathcal{H}_\infty$-optimal control and more POPs. It would also be valuable to investigate compromises between cone complexity, additional consistency constraints, and approximation quality.

## Appendix A. Proof of Proposition 1

In this Appendix, we provide the proof for Proposition 1 where $\mathcal{E}$ is not necessarily a chordal graph.

$\Leftarrow$: If $Z_k \in \mathcal{DD}^{|\mathcal{C}_k|}$ then $E_{\mathcal{C}_k}^{\mathsf{T}} Z_k E_{\mathcal{C}_k} \in \mathcal{DD}^n$ because diagonal elements of $Z_k$ will remain on the diagonal of $E_{\mathcal{C}_k}^{\mathsf{T}} Z_k E_{\mathcal{C}_k}$, and the diagonal dominance relation is preserved in the embedded matrix. $\mathcal{DD}^n$ is a cone, so $Z = \sum_{k=1}^{p} E_{\mathcal{C}_k}^{\mathsf{T}} Z_k E_{\mathcal{C}_k} \in \mathcal{DD}^n$ for multiple cliques in $\mathcal{C}$. Let $\mathcal{E}^c \supseteq \mathcal{E}$ be a chordal completion of $\mathcal{E}$ with a clique cover $\mathcal{C}^c$. All of the induced submatrices of $Z$ by cliques in $\mathcal{C}^c$ are *DD*, so from Agler's theorem we have $Z \in \mathbb{S}^n(\mathcal{E}, 0)$. Thus, $Z \in \mathcal{DD}^n(\mathcal{E}, 0)$.

$\Rightarrow$: Let $e_i \in \mathbb{R}^n$, $e_j \in \mathbb{R}^n$ where $i \neq j$ be standard basis vectors, we define the following *DD* basis matrices:

$$v_i = e_i e_i^{\mathsf{T}}, \qquad v_{ij}^{\pm} = (e_i \pm e_j)(e_i \pm e_j)^{\mathsf{T}}.$$

Given a symmetric $Z \in \mathcal{DD}^n(\mathcal{E}, 0)$ for a (not-necessarily chordal) sparsity pattern $\mathcal{E}$, define the slack quantities $\Delta_i = Z_{ii} - \sum_{j \neq i} |Z_{ij}| \geq 0$. Such a $Z$ can be decomposed as:

$$Z = \sum_{i=1}^{n} \Delta_i v_i + \sum_{(i,j)\in\mathcal{P}} Z_{ij} v_{ij}^{+} + \sum_{(i,j)\in\mathcal{N}} |Z_{ij}| v_{ij}^{-}$$
$$\mathcal{P} := \{(i,j) \mid Z_{ij} > 0,\ i < j\}$$
$$\mathcal{N} := \{(i,j) \mid Z_{ij} < 0,\ i < j\}.$$

By this characterization, $Z \in \mathcal{DD}^n(\mathcal{E}, 0)$ can be represented as the sum of *DD* matrices with the same pattern $\mathcal{E}$ and clique cover $\mathcal{C}$. The terms $v_{ij}^{+}$, $v_{ij}^{-}$ with $Z_{ij} \neq 0$ can be uniquely assigned to some clique $(i,j) \in \mathcal{C}_k$, and the slack terms $\Delta_i v_i$ can be distributed among all cliques in $\mathcal{C}$ that include $i$. Grouping summands into cliques yields

$$Z = \sum_{k=1}^{p} E_{\mathcal{C}_k}^{\mathsf{T}} Z_k E_{\mathcal{C}_k}, \quad Z_k \in \mathcal{DD}^{|\mathcal{C}_k|}, \quad k = 1, \ldots, p.$$

Let $D$ be a positive definite (PD) diagonal matrix, and define matrices $D_{\mathcal{C}_k} = E_{\mathcal{C}_k} D E_{\mathcal{C}_k}^{\mathsf{T}}$ with inverses $D_{\mathcal{C}_k}^{-1} = E_{\mathcal{C}_k} D^{-1} E_{\mathcal{C}_k}^{\mathsf{T}}$. By definition (3) there exists a PD diagonal matrix $D$ for a $Z \in \mathcal{SDD}^n(\mathcal{E}, 0)$ such that $DZD \in \mathcal{DD}^n$. Since pre and post-multiplying by a diagonal matrix does not change the sparsity pattern, $DZD \in \mathcal{DD}^n(\mathcal{E}, 0)$. By the decomposition of $\mathcal{DD}^n$ matrices:

$$DZD = \sum_{k=1}^{p} E_{\mathcal{C}_k}^{\mathsf{T}} \tilde{Z}_k E_{\mathcal{C}_k},$$

where $\tilde{Z}_k \in \mathcal{DD}^{|\mathcal{C}_k|}$, $k = 1, \ldots, p$. This leads to

$$Z = \sum_{k=1}^{p} D^{-1} E_{\mathcal{C}_k}^{\mathsf{T}} \tilde{Z}_k E_{\mathcal{C}_k} D^{-1} = \sum_{k=1}^{p} E_{\mathcal{C}_k}^{\mathsf{T}} Z_k E_{\mathcal{C}_k},$$

where $Z_k = D_{\mathcal{C}_k}^{-1} \tilde{Z}_k D_{\mathcal{C}_k}^{-1} \in \mathcal{SDD}^{|\mathcal{C}_k|}$, since $E_{\mathcal{C}_k} D^{-1} = D_{\mathcal{C}_k}^{-1} E_{\mathcal{C}_k}$ and $D_{\mathcal{C}_k}^{-1} = E_{\mathcal{C}_k} D^{-1} E_{\mathcal{C}_k}^{\mathsf{T}}$, completing the proof.

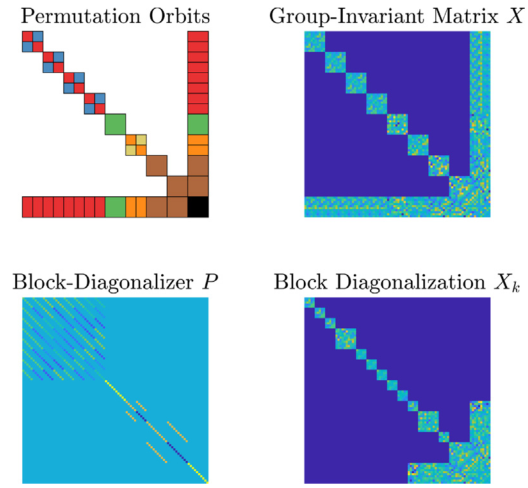



**Fig. B.1.** Sparse and Symmetric *-algebra.

## Appendix B. Combining decompositions

### *B.1. Symmetry/*-algebra decomposition*

An additional form of structure occurs when all constraint and cost matrices $(C, A_i) \in \mathbb{S}^n$ can be simultaneously block diagonalized by a unitary matrix $P \in \mathbb{R}^{n\times n}$:

$$\langle C, X\rangle = \langle P^T C P, \tilde{X}\rangle = \textstyle\sum_k \langle \tilde{C}_k, \tilde{X}_k\rangle.$$

Application of $P$ together with Agler's theorem breaks the large PSD variable $X = P\tilde{X}P^T$ into a set of smaller PSD variables $\tilde{X}_k$, and the SDP in $(\tilde{C}_k, \tilde{A}_{ik})$ will have an equivalent optimum as $(C, A_i)$ (Vallentin, 2009). This block diagonalization can occur if all matrices $(C, A_i)$ lie in a common *-algebra $\mathcal{A} \subset \mathbb{S}^n$ , which is closed under addition, products, and transposition. The structure of the *-algebra's block-diagonalization into blocks of size $n_i$ with multiplicity $m_i$ such that $n = \sum_i n_i m_i$ is expressed in the Wedderburn decomposition (Wedderburn, 1934) of $\mathcal{A} \cong \oplus_{i=1}^{p} I_{m_i} \otimes \mathbb{S}^{n_i}$, with inner products $\langle C, X\rangle = \sum_i m_i \langle \tilde{C}_i, \tilde{X}_i\rangle$. The set of matrices in $\mathbb{S}_+^n$ invariant under action by a group $G$ forms a *-algebra $[\mathbb{S}_+^n]^G$.

### *B.2. Symmetry and sparse structure*

Decomposed structured subsets can be applied to symmetric SDPs in the same manner as chordally sparse problems. In the *-algebra framework, the cone-set $\mathcal{K}$ refers to the cone of each symmetric block in the program. Fig. B.1 shows block-arrow (sparse) matrices in the invariant ring $[\mathbb{S}^{90}]^G$ under block-permutation action. The permutation structure is illustrated in the top-left pane of Fig. B.1. Matrices $X \in [\mathbb{S}^{90}]^G$ are invariant under swapping blocks of the same color, and the split blocks are additionally invariant under swapping the top and bottom halves. The permutation group acting on these matrices is $G = (S_4 \times \mathbb{Z}_2) \times Id \times \mathbb{Z}_2 \times \mathbb{Z}_2$. The top right pane of Fig. B.1 shows a one such matrix $X \in [\mathbb{S}^{90}]^G$. All such matrices can be block diagonalized (bottom right) under unitary action by the matrix $P$ (bottom left, based on the Discrete Cosine Transform). The block-sizes and multiplicities are:

$$\{m_i, n_i\}_{i=1}^{5} = \{(3, 5), (1, 10), (4, 5), (1, 5), (1, 40)^{sparse}\}.$$

Table B.1 shows the cost and time of running a randomly generated $G$-invariant SDP with 80 equality constraints under the cone $\mathcal{SDD}$. The rows show if the matrix has been block-diagonalized before applying $\mathcal{SDD}$, and the columns indicate if Grone's theorem has been applied to apply $\mathcal{SDD}$ to cliques. Whole blocks (green, brown orbits) contribute PSD constraints with $n = 10$ while split blocks yield $n = 5$.

**Table B.1**
$\mathcal{SDD}$ block arrow with symmetry.

| Cost | Full | Sym. | Time (s) | Full | Sym. |
|---|---|---|---|---|---|
| Dense | 12.96 | 10.86 | Dense | 124.5 | 19.3 |
| Sparse | 9.49 | 8.44 | Sparse | 38.2 | 12.0 |

## References


Agler, J., Helton, W., et al. (1988). Positive semidefinite matrices with a given sparsity pattern. *Linear Algebra and Its Applications*, *107*, 101–149.

Ahmadi, A. A., Dash, S., & Hall, G. (2017). Optimization over structured subsets of positive semidefinite matrices via column generation. *Discrete Optimization*, *24*, 129–151.

Ahmadi, A. A., & Hall, G. (2017). Sum of squares basis pursuit with linear and second order cone programming. In *Algebraic and geometric methods in discrete mathematics* (pp. 27–54). Oxford: AMS.

Alizadeh, F. (1995). Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM Journal on Optimization*, *5*(1), 13–51.

Andersen, E. D., & Andersen, K. D. (2000). The mosek interior point optimizer for linear programming: An implementation of the homogeneous algorithm. In *High performance optimization* (pp. 197–232). Springer.

Barker, G., & Carlson, D. (1975). Cones of diagonally dominant matrices. *Pacific Journal of Mathematics*, *57*(1), 15–32.

Boman, E. G., Chen, D., Parekh, O., & Toledo, S. (2005). On factor width and symmetric H-matrices. *Linear Algebra and Its Applications*, *405*, 239–248.

Boyd, S., El Ghaoui, L., Feron, E., & Balakrishnan, V. (1994). *Linear matrix inequalities in system and control theory, vol. 15*. SIAM.

Boyd, S., & Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press.

Garstka, M., Cannon, M., & Goulart, P. (2020). A clique graph based merging strategy for decomposable SDPs. *IFAC-PapersOnLine*, *53*(2), 7355–7361.

Grone, R., Johnson, C. R., Sá, E. M., & Wolkowicz, H. (1984). Positive definite completions of partial Hermitian matrices. *Linear Algebra and its Applications*, *58*, 109–124.

Kakimura, N. (2010). A direct proof for the matrix decomposition of chordal-structured positive semidefinite matrices. *Linear Algebra and its Applications*, *433*(4), 819–823.

Lasserre, J. -B. (2010). *Moments, positive polynomials and their applications, vol. 1*. World Scientific.

Majumdar, A., Ahmadi, A. A., & Tedrake, R. (2014). Control and verification of high-dimensional systems with DSOS and SDSOS programming. In *53rd IEEE conference on decision and control* (pp. 394–401). IEEE.

Miller, J., Zheng, Y., Sznaier, M., & Papachristodoulou, A. (2020). Decomposed structured subsets for semidefinite optimization. *IFAC-PapersOnLine*, *53*(2), 7374–7379. http://dx.doi.org/10.1016/j.ifacol.2020.12.1262, 21th IFAC World Congress.

Parrilo, P. A. (2000). *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization* (Ph.D. thesis), California Institute of Technology.

Putinar, M. (1993). Positive polynomials on compact semi-algebraic sets. *Indiana University Mathematics Journal*, *42*(3), 969–984.

Vallentin, F. (2009). Symmetry in semidefinite programs. *Linear Algebra and its Applications*, *1*(430), 360–369.

Vandenberghe, L., Andersen, M. S., et al. (2015). Chordal graphs and semidefinite optimization. *Foundations and Trends in Optimization*, *1*(4), 241–433.

Waki, H., Kim, S., Kojima, M., & Muramatsu, M. (2006). Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity. *SIAM Journal on Optimization*, *17*(1), 218–242.

Wang, J., Magron, V., & Lasserre, J. -B. (2021a). Chordal-TSSOS: A moment-SOS hierarchy that exploits term sparsity with chordal extension. *SIAM Journal on Optimization*, *31*(1), 114–141.

Wang, J., Magron, V., & Lasserre, J. -B. (2021b). TSSOS: A moment-sos hierarchy that exploits term sparsity. *SIAM Journal on Optimization*, *31*(1), 30–58.

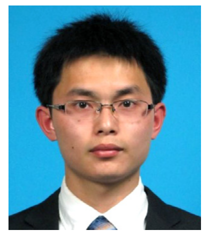
Wedderburn, J. H. M. (1934). *Lectures on matrices, vol. 17*. American Mathematical Society.

Zheng, Y., Kamgarpour, M., Sootla, A., & Papachristodoulou, A. (2018). Scalable analysis of linear networked systems via chordal decomposition. In *2018 European control conference* (pp. 2260–2265). IEEE.

Zheng, Y., Sootla, A., & Papachristodoulou, A. (2019). Block factor-width-two matrices and their applications to semidefinite and sum-of-squares optimization. arXiv preprint arXiv:1909.11076.



**Jared Miller** is currently a Research Assistant (4th-year Ph.D. student) at the Robust Systems Laboratory in Northeastern University. He received his B.S. and M.S. degrees in Electrical Engineering from Northeastern University in 2018. He is a recipient of the 2020 Chateaubriand Fellowship from the Office for Science Technology of the Embassy of France in the United States. He received an Outstanding Student Paper Award at the 60th Conference on Decision and Control in 2021. His research interests include large-scale convex optimization, semi-algebraic geometry, measure theory, and analysis of nonlinear systems.



**Yang Zheng** received the B.E. and M.S. degrees from Tsinghua University, Beijing, China, in 2013 and 2015, respectively, and the D.Phil. (Ph.D.) degree in Engineering Science from the University of Oxford, U.K., in 2019. He is currently an assistant professor with the Department of Electrical and Computer Engineering, UC San Diego. He was a research associate at Imperial College London and was a postdoctoral scholar in SEAS and CGBC at Harvard University. His research interests focus on learning, optimization, and control of network systems, and their applications to autonomous vehicles and traffic systems.

Dr. Zheng was a finalist (co-author) for the Best Student Paper Award at the 2019 ECC. He received the Best Student Paper Award at the 17th IEEE ITSC in 2014, and the Best Paper Award at the 14th Intelligent Transportation Systems Asia-Pacific Forum in 2015. He was a recipient of the National Scholarship, Outstanding Graduate in Tsinghua University, the Clarendon Scholarship at the University of Oxford, and the Chinese Government Award for Outstanding Self-financed Students Abroad. Dr. Zheng won the 2019 European Ph.D. Award on Control for Complex and Heterogeneous Systems.

**Mario Sznaier** is currently the Dennis Picard Chaired Professor at the Electrical and Computer Engineering Department, Northeastern University, Boston. Prior to joining Northeastern University, Dr. Sznaier was a Professor of Electrical Engineering at the Pennsylvania State University and also held visiting positions at the California Institute of Technology. His research interest includes robust identification and control of hybrid systems, robust optimization, and dynamical vision. Dr. Sznaier is currently serving as Editor in Chief of the section on AI and Control of the journal Frontiers in Control Engineering, and chair of IFAC's Technical Committee TC 2.5, Additional recent service includes Associate Editor for the Journal Automatica (2005–2021), chair of the CSS Technical Committee on Computational Aspects of Control System Design (2012–2016), General co-Chair of the 2016 MSC and Program Chair of the 2017 CDC. He is a distinguished member of the IEEE Control Systems Society and an IEEE Fellow.

**Antonis Papachristodoulou** (Fellow, IEEE) received the M.A./M.Eng. degree in electrical and information sciences from the University of Cambridge, Cambridge, U.K., in 2000, and the Ph.D. degree in control and dynamical systems (with a minor in aeronautics) from the California Institute of Technology, Pasadena, CA, USA, in 2005. He is currently a Professor of engineering science with the University of Oxford, Oxford, U.K., and a Tutorial Fellow with Worcester College, Oxford. His research interests include large-scale nonlinear systems analysis, sum-of-squares programming, synthetic and systems biology, networked systems, and flow control. Professor Papachristodoulou was the recipient of the 2015 European Control Award for his contributions to robustness analysis and applications to networked control systems and systems biology. In the same year, he received the O. Hugo Schuck Best Paper Award.