A Bound on the Minimum Rank of Solutions to Sparse Linear Matrix Equations

Raphael Louca[1] Subhonmesh Bose[2] Eilyan Bitar[1]

***Abstract*—We derive a new upper bound on the minimum rank of matrices belonging to an affine slice of the positive semidefinite cone, when the affine slice is defined according to a system of sparse linear matrix equations. It is shown that a feasible matrix whose rank is no greater than said bound can be computed in polynomial time. The bound depends on both the number of linear matrix equations and their underlying sparsity pattern. For certain problem families, this bound is shown to improve upon well known bounds in the literature. Several examples are provided to illustrate the efficacy of this bound.**

***Index Terms*—rank minimization, semidefinite programming, sparse linear matrix equations, chordal graphs.**

I. Introduction

Let $\mathbf{R}$ (resp. $\mathbf{R}_+$) denote the set of real (resp. nonnegative) numbers and $\mathbf{S}^n$ (resp. $\mathbf{S}^n_+$) denote the space of all $n \times n$ symmetric (resp. symmetric positive semidefinite) matrices. Consider an affine subspace $\mathcal{A}$, defined by

$$\mathcal{A} := \{X \in \mathbf{S}^n \mid \text{Tr}(A_i X) = b_i, \ i = 1, \dots, m\}, \quad (1)$$

where $A_i \in \mathbf{S}^n$ and $b_i \in \mathbf{R}$ for all $i = 1, \dots, m$. Given an integer $1 \leq r \leq n$, of general interest is the derivation of conditions under which a matrix $X \in \mathcal{A} \cap \mathbf{S}^n_+$ satisfying $\text{rank}(X) \leq r$ is guaranteed to exist and can be computed in polynomial time.

A wide range of problems can be cast as rank minimization problems over an affine slice of the positive semidefinite cone. Examples include covariance estimation problems, and k-means clustering problems [1]. In such examples, the minimum rank solution corresponds to the simplest model that explains the observed data. A closely related class of problems is that of minimizing the rank of a matrix over an affine subspace of symmetric matrices. These so-called affine rank minimization problems appear in collaborative filtering [2], [3], quantum state tomography [4], cardinality minimization [5], and matrix completion problems [6], [7], to name a few. Fazel [8] shows that an affine rank minimization problem can be equivalently reformulated as one of minimizing rank over an affine slice of the positive semidefinite cone. However, computing a matrix of minimum rank in the affine slice is known to be NP-hard [9]. A common approach to the approximation of such problems entails the use of the nuclear norm as a convex surrogate for rank. The resulting nuclear norm minimization problem can be solved in polynomial time using semidefinite programming. Furthermore, when the linear map defining the affine subspace satisfies a certain null space property, nuclear norm minimization is guaranteed to find the minimum rank matrix in the affine slice. Certain random ensembles of linear systems that arise in practice satisfy the null-space property with high probability. For example, see [3], [10]–[12]. Random subspaces drawn from such ensembles are largely unstructured. A priori upper bounds on the minimum rank of matrices in $\mathcal{A} \cap \mathbf{S}^n_+$ for unstructured $\mathcal{A}$ have appeared in [13]–[16], among others.

Rank minimization problems arising in engineering applications often impose additional structures on $\mathcal{A}$. For example, system identification problems in control theory involve minimizing the rank of a Hankel matrix [17], [18]. See [19], [20] for applications that involve optimization, more generally, over patterned matrices such as Hankel, Toeplitz, Sylvester, etc. Structured subspaces $\mathcal{A}$ also appear in applications where the matrices defining $\mathcal{A}$ are sparse. The semidefinite relaxation of the optimal power flow problem in power systems [21]–[23] is one such example. More generally, there are a wide variety of nonconvex optimization problems over graphs that can be cast as sparse semidefinite programs with an additional rank constraint. See, for example, [21], [23]–[29].

When seeking a low-rank solution, X, to a semidefinite program, the authors in [30] advocate the use of first-order methods over the low-rank matrix factorizations of X. Such algorithms are demonstrably faster than interior-point methods for solving semidefinite programs, especially when the weighting matrices are sparse. Moreover, they are also proven to converge to a local minimum of the rank-constrained variants of the semidefinite program [31] with fast convergence rates [32], [33]. Knowledge of an a priori upper bound on the minimum attainable rank serves to reduce the search space of the algorithms.

Contribution and Related Work: Our primary objective in this paper is the derivation of upper bounds on the minimum rank of matrices in $\mathcal{A} \cap \mathbf{S}^n_+$. There are two streams of prior work in the literature, which are closely related. The first line of research [13], [14] provides an upper bound based solely on the dimension of $\mathcal{A}$. The second body of work [23], [24]

Supported in part by NSF grant ECCS-1351621, NSF grant CNS-1239178, US DoE under the CERTS initiative, PSERC under sub-award S-52, and a Postdoctoral Fellowship from the Atkinson Center for a Sustainable Future.

[1]R. Louca and E. Bitar are with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, 14853, USA. Emails: rl553@cornell.edu, eyb5@cornell.edu

[2]S. Bose is with the Department of Electrical and Computer Engineering, University of Illinois, Champaign, IL, 61820, USA. Email: boses@illinois.edu

leverages on the graph structure of underlying sparsity pattern of the matrices defining $\mathcal{A}$. Loosely speaking, our main result, in Theorem 1, marries these two approaches.

Organization: We begin the paper with graph theoretic definitions in Section II. In Section III, we summarize known upper bounds on the minimum rank of matrices in $\mathcal{A}\cap\mathbf{S}^n_+$. Our main result – providing a novel upper bound – is stated, derived, and compared to known bounds in Section IV. In Section V, we further investigate the bound when the collective sparsity pattern of the matrices defining $\mathcal{A}$ is given by a chordal graph. We conclude the paper with Section VI.

II. Graph Theory Preliminaries

We begin with several basic definitions from graph theory that will prove useful in the sequel. Let $G = (\mathcal{V}, \mathcal{E})$ be a simple undirected graph, where $\mathcal{V} := \{1, \ldots, n\}$ denotes its vertex set and $\mathcal{E}$ its edge set. For convenience, let $|G| := |\mathcal{V}|$. We say G_1 is a subgraph of G_2, denoted by $G_1 \subseteq G_2$, if the inclusion relation holds for both the vertex and the edge sets of the respective graphs. A *cycle* on k vertices in G is defined as a k-tuple $(n_1, n_2, \ldots, n_k)$, such that (n_1, n_2), (n_2, n_3),. . ., (n_k, n_1) are edges belonging to $\mathcal{E}$. A cycle $(n_1, n_2, \ldots, n_k)$ in G is said to be *minimal* if no strict subset of $\{n_1, n_2, \ldots, n_k\}$ defines a cycle in G.

Definition 1 (Chordal graph)**.** *A graph is said to be chordal if all its minimal cycles have at most three nodes.*[1]

For any graph G on n nodes, a graph H is said to be a *chordal extension* of G, if $H \supseteq G$, and H is a chordal graph on the same n nodes. Denote by $\mathcal{C}(G)$, the set of all chordal extensions of G.

In the sequel, it will be of notational convenience to associate with any graph G the index set $\mathcal{I}(G)$ defined as

$$\mathcal{I}(G) := \{(j, k) : j = k \in \mathcal{V} \text{ or } j < k,\ (j, k) \in \mathcal{E}\}. \tag{2}$$

With a slight abuse of notation, we also define the index set associated with a collection of graphs $G_1, \ldots, G_p$ as $\mathcal{I}(G_1, \ldots, G_p) := \cup_{i=1}^p \mathcal{I}(G_i)$.

A. Maximal cliques

Let G be a simple undirected graph. It is said to be *complete*, if every pair of nodes in the graph share an edge. A *clique* of G is a complete subgraph of G. A clique is said to be *maximal*, if it is not a subgraph of any other clique of G. Finally, a clique is said to be a *maximum* clique if it contains the most vertices among all cliques.

For a graph G, let $\mathcal{K}(G)$ denote the set of all maximal cliques in G. For any maximal clique $K \in \mathcal{K}(G)$, denote its *complement* by $K^c := \mathcal{K}(G) \setminus \{K\}$. The *clique number* of G is defined as

$$\omega(G) := \max_{K \in \mathcal{K}(G)} |K|.$$

[1]See [34]–[38] for a detailed survey on chordal graphs.

(a) Graph G_1 (b) Graph G_2

Fig. 1: G_1 is not a chordal graph, while G_2 is a chordal extension of G_1. The clique numbers of G_1 and G_2 are $\omega(G_1) = 2$ and $\omega(G_2) = 3$, respectively.

In other words, $\omega(G)$ equals the number of vertices in a maximum clique of G. Recall that $\mathcal{C}(G)$ denotes the set of all chordal extensions of G. We remark that $\min_{H \in \mathcal{C}(G)} \{\omega(H) - 1\}$ is often called the *treewidth* of G.

Example 1. We illustrate several of the preceding definitions using the chordal graph G_2 in Figure 1. The set of all maximal cliques of G_2 is given by $\mathcal{K}(G_2) = \{K_1, K_2, K_3\}$, where $K_1 = \{1, 2, 4\}$, $K_2 = \{1, 3, 4\}$, and $K_3 = \{3, 5\}$. For the maximal clique K_2 and its complement $K_2^c = \{K_1, K_3\}$, we have

$$\begin{aligned}\mathcal{I}(K_2) &= \{(1,1), (1,3), (1,4), (3,3), (3,4), (4,4)\},\\ \mathcal{I}(K_2^c) &= \{(1,1), (1,2), (1,4), (2,2),\\ &\qquad (2,4), (3,3), (3,5), (4,4), (5,5)\}.\end{aligned}$$

Moreover, G_2 is a chordal extension of G_1 with clique number $\omega(G_2) = 3$.

III. Known Bounds on Minimum Rank

We now summarize two well known results from the literature, which provide upper bounds on the minimum rank of matrices in $\mathcal{A} \cap \mathbf{S}^n_+$. These results will play a central role in the derivation of our main result in Theorem 1.

A. The Role of Convex Geometry

The authors in [13]–[16], [39] provide an upper bound on the minimum rank of matrices belonging to $\mathcal{A} \cap \mathbf{S}^n_+$, as a function only of the number of constraints m defining the affine subspace $\mathcal{A}$. We require some basic notation to present these results. For any positive integer n, let $n^{\overline{2}}$ denote the dimension of the space $\mathbf{S}^n$, given by

$$n^{\overline{2}} := n(n+1)/2.$$

Also, define the function

$$\sqrt[\overline{2}]{x} := \left\lfloor \frac{\sqrt{8x+1} - 1}{2} \right\rfloor$$

for $x \in \mathbf{R}_+$. Here, $\sqrt[\overline{2}]{x} = \lfloor y \rfloor$, where y is the positive real root of $x = y^{\overline{2}}$. It is also straightforward to verify that for any positive integer n, we have $\sqrt[\overline{2}]{n^{\overline{2}}} = n$. Equipped with this notation, we now summarize the main result of [13], [14], [39], [40] in the following proposition.

Proposition 1. *If $\mathcal{A} \cap \mathbf{S}_+^n$ is nonempty, then there exists a matrix $X \in \mathcal{A} \cap \mathbf{S}_+^n$ satisfying*

$$\textit{rank}(X) \leq \sqrt[2]{m}.$$

Moreover, such a matrix X can be computed in polynomial time.

A detailed proof for existence, relying only on the convex geometry of the positive semidefinite cone, can be found in [41]. An alternative proof based on nondegeneracy and complementarity conditions for semidefinite programs can be found in [15]. The reader may refer to [16] for the specification of a polynomial time algorithm to compute such matrices.

The upper bound in Proposition 1 is known to be tight. That is to say, for every $m \leq n^{\overline{2}}$, there exists an affine subspace $\mathcal{A}$ defined according to (1) such that $\text{rank}(X) \geq \sqrt[2]{m}$ for all $X \in \mathcal{A} \cap \mathbf{S}_+^n$. Thus, the upper bound in Proposition 1 cannot be improved without imposing additional conditions on $\mathcal{A}$. Examples of such refinements can be found in [16] and the references therein.

B. The Role of Sparsity

The bound on minimum rank presented in Proposition 1 depends only on the number of constraints used in defining the affine subspace $\mathcal{A}$. In Proposition 2, we present a complementary result that specifies an upper bound on the minimum rank of matrices belonging to $\mathcal{A} \cap \mathbf{S}_+^n$, as a function only of the sparsity pattern of the weighting matrices $A_1, \ldots, A_m$ defining $\mathcal{A}$. In the following definition, we formalize the notion of sparsity by defining a graph whose edges reflect the nonzero entries associated with a collection of symmetric matrices.

Definition 2. *The collective sparsity pattern of the weighting matrices $A_1, \ldots, A_m \in \mathbf{S}^n$ characterizing the affine subspace $\mathcal{A}$ in (1) is defined by a simple undirected graph $G_{\mathcal{A}} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} := \{1, \ldots, n\}$ and*

$$\mathcal{E} := \{(j,k) : j \neq k, \ [A_i]_{jk} \neq 0 \text{ for some } i = 1, \ldots, m\}.$$

Essentially, $G_{\mathcal{A}}$ contains only those edges which correspond to the nonzero entries of the matrices in the collection $A_1, \ldots, A_m$. Of particular interest are matrices whose collective sparsity pattern defines a chordal graph. The following proposition specifies an upper bound on the minimum rank of matrices in $\mathcal{A} \cap \mathbf{S}_+^n$, when $G_{\mathcal{A}}$ is a chordal graph. The result follows from arguments in [25], [42], [43]

Proposition 2. *If $\mathcal{A} \cap \mathbf{S}_+^n$ is nonempty and $G_{\mathcal{A}}$ is a chordal graph, then there exists a matrix $X \in \mathcal{A} \cap \mathbf{S}_+^n$ satisfying*

$$\textit{rank}(X) \leq \omega(G_{\mathcal{A}}).$$

Moreover, such a matrix X can be computed in polynomial time.

Proposition 2 is a special case of Lemma 5.2.8 in [24], where the authors show that a matrix $X \in \mathcal{A} \cap \mathbf{S}_+^n$ satisfying $\text{rank}(X) \leq \omega(G_{\mathcal{A}})$ exists for arbitrary graphs $G_{\mathcal{A}}$. The result of Proposition 2 relies critically on the assumption that the graph $G_{\mathcal{A}}$ is chordal to ensure that such a matrix can be computed in polynomial time. When $G_{\mathcal{A}}$ is not chordal, one can replace $\omega(G_{\mathcal{A}})$ in Proposition 2 with $\omega(H)$ for any $H \in \mathcal{C}(G_{\mathcal{A}})$, where recall that $\mathcal{C}(G_{\mathcal{A}})$ denotes the set of all chordal extensions of $G_{\mathcal{A}}$. The problem of finding a chordal extension with the smallest clique number is, however, NP-complete in general. One can rely instead on chordal extensions of $G_{\mathcal{A}}$ that can be computed in polynomial time. We refer the reader to [44, Chapter 18] for one such method.

IV. Main Result

We now present our main result in Theorem 1, which combines insights from both Propositions 1 and 2. Given any set of indices $\mathcal{I} \subseteq \{(j,k) : 1 \leq j \leq k \leq n\}$, define $\mu_{\mathcal{A}}(\mathcal{I})$ as

$$\mu_{\mathcal{A}}(\mathcal{I}) := |\{i \in \{1, \ldots, m\} : [A_i]_{jk} \neq 0 \text{ for some } (j,k) \in \mathcal{I}\}|.$$

Essentially, $\mu_{\mathcal{A}}(\mathcal{I})$ equals the number of constraints whose weighting matrices have nonzero entries at indices in $\mathcal{I}$.

Theorem 1. *Let H be a chordal extension of $G_{\mathcal{A}}$, the collective sparsity pattern of $\mathcal{A}$. For each maximal clique $K \in \mathcal{K}(H)$, define*

$$m_K := \mu_{\mathcal{A}}(\mathcal{I}(K) \setminus \mathcal{I}(K^c)) \ + \ |\mathcal{I}(K) \cap \mathcal{I}(K^c)|. \tag{3}$$

If $\mathcal{A} \cap \mathbf{S}_+^n$ is nonempty, then there exists a matrix $X \in \mathcal{A} \cap \mathbf{S}_+^n$ satisfying

$$\textit{rank}(X) \leq \rho_{\mathcal{A}}(H),$$

where

$$\rho_{\mathcal{A}}(H) := \max_{K \in \mathcal{K}(H)} \min\left\{|K|, \ \sqrt[2]{m_K}\right\}. \tag{4}$$

Moreover, such a matrix X can be computed in polynomial time.

A proof sketch is provided in Section IV-A. Here, we offer a brief discussion of the result. If the weighting matrices defining the affine subspace $\mathcal{A}$ are *dense* (i.e., all elements are nonzero), then the associated graph $G_{\mathcal{A}}$ equals the complete graph. For dense systems, it holds that $\rho_{\mathcal{A}}(G_{\mathcal{A}}) = \sqrt[2]{m}$. In other words, when the weighting matrices are dense, we recover the upper bound in Proposition 1. Now, consider an affine subspace $\mathcal{A}$ defined by *sparse* weighting matrices for which the associated graph $G_{\mathcal{A}}$ is chordal. It follows from Theorem 1 that $\rho_{\mathcal{A}}(G_{\mathcal{A}})$ is a valid upper bound on the minimum rank of matrices in $\mathcal{A} \cap \mathbf{S}_+^n$. And, it is easy to check that $\rho_{\mathcal{A}}(G_{\mathcal{A}}) \leq \omega(G_{\mathcal{A}})$. In essence, we recover the upper bound in Proposition 2, in this case. In general, however, one cannot compare $\sqrt[2]{m}$ and $\rho_{\mathcal{A}}(H)$ given an arbitrary chordal extension $H \in \mathcal{C}(\mathcal{G}_{\mathcal{A}})$.

Corollary 1. *If $\mathcal{A} \cap \mathbf{S}_+^n$ is nonempty, then there exists $X \in \mathcal{A} \cap \mathbf{S}_+^n$, such that $\textit{rank}(X) \leq \rho_{\mathcal{A}}^*$, where*

$$\rho_{\mathcal{A}}^* := \min_{H \in \mathcal{C}(G_{\mathcal{A}})} \rho_{\mathcal{A}}(H).$$

Corollary 1 generalizes both Propositions 1 and 2 in that $\rho_{\mathcal{A}}^* \leq \min\{\omega(G_{\mathcal{A}}), \sqrt[2]{m}\}$ for all affine subspaces $\mathcal{A}$ having a graph $G_{\mathcal{A}}$ that is chordal. However, the gain in generality is offset by the loss of a polynomial time guarantee for construction, as calculating an optimal chordal extension $H \in \mathcal{C}(G_{\mathcal{A}})$ – one that minimizes $\rho_{\mathcal{A}}(H)$ – appears to be an intractable problem. Intuitively, one expects that computing $\rho_{\mathcal{A}}^*$ should be as hard as minimizing $\omega(H) = \max_{K \in \mathcal{K}(H)} |K|$ over $H \in \mathcal{C}(G_{\mathcal{A}})$. The optimal value of the latter problem is precisely equal to one plus the treewidth of $G_{\mathcal{A}}$. And, computing the treewidth of an arbitrary graph is known to be NP-complete [45]. This does not, however, preclude the possibility of identifying natural families of systems for which optimizing $\rho_{\mathcal{A}}$ over $\mathcal{C}(G_{\mathcal{A}})$ is tractable. We discuss such issues in Section V.

A. A Proof Sketch of Theorem 1

A detailed proof of the result is not included in this paper, but we outline the key steps involved in the remainder of this section. Central to our proof is a technical result presented in Lemma 1. Its statement requires the following additional notation.

For a simple undirected graph $G = (\mathcal{V}, \mathcal{E})$ having vertices $\mathcal{V} = \{1, \ldots, n\}$, define a *$G$-partial matrix* X_G as a collection of real numbers, indexed by the set $\mathcal{I}(G)$, where $\mathcal{I}(G)$ is defined in (2). With a slight abuse of notation, also identify X_G with a partially filled symmetric matrix, where $[X_G]_{kj} = [X_G]_{jk}$ for each $(j,k) \in \mathcal{I}(G)$. When G is the complete graph on n nodes, a G-partial matrix is an $n \times n$ symmetric matrix. For any graph $H \subseteq G$, let $X_G(\mathcal{I}(H))$ denote the restriction of X_G to the indices in $\mathcal{I}(H)$. If K is a maximal clique, i.e., $K \in \mathcal{K}(G)$, then $X_G(\mathcal{I}(K))$ can be identified with a $|K| \times |K|$ symmetric matrix.

Definition 3 (Matrix completion)**.** *A matrix $X \in \mathbf{S}^n$ is said to be a completion of a G-partial matrix X_G, if $X(\mathcal{I}(G)) = X_G$.*

The proof of Theorem 1 relies on the following result, a formal proof of which is omitted due to space constraints. We remark that [24, Lemma 2.3.11], [25], [43] form the crux of the proof of this lemma.

Lemma 1. *Let G be a chordal graph and X_G a G-partial matrix, such that $X_G(\mathcal{I}(K)) \in \mathbf{S}_+^{|K|}$ and rank$(X_G(\mathcal{I}(K)) \leq r$ for each $K \in \mathcal{K}(G)$. Then, there exists a completion X of X_G, such that $X \in \mathbf{S}_+^n$ and rank$(X) \leq r$. Moreover, X can be computed from X_G in polynomial time.*

Equipped with this result, we now outline the proof of Theorem 1. First, let $Z \in \mathcal{A} \cap \mathbf{S}_+^n$. One can compute such a matrix in polynomial time by solving a semidefinite program. We use Z to construct an H-partial matrix X_H, where finding $X_H(K)$ amounts to solving a semidefinite feasibility problem for each maximal clique $K \in \mathcal{K}(H)$. The semidefinite program associated with each clique K has m_K linear constraints, where m_K is defined as in (3). Invoking Proposition 1, a positive semidefinite matrix $X_H(\mathcal{I}(K))$ satisfying the m_K linear constraints with

$$\text{rank}(X_H(\mathcal{I}(K))) \leq \sqrt[2]{m_K}$$

exists and can be computed in polynomial time. Also, rank$(X_H(\mathcal{I}(K))) \leq |K|$. Furthermore, H has at most n maximal cliques, which can be listed in polynomial time. Hence, X_H can be constructed in polynomial time. Lemma 1 then implies that X_H can be completed to $X \in \mathbf{S}_+^n$ in polynomial time, where rank(X) is bounded from above by $\rho_{\mathcal{A}}(H)$ in (4). Finally, our construction of X_H from Z is such that its completion satisfies $X \in \mathcal{A}$.

V. Chordally sparse linear systems

We devote this section to linear systems for which the graph $G_{\mathcal{A}}$ associated with the affine subspace $\mathcal{A}$ is chordal. For such systems, it follows from Theorem 1 that $\rho_{\mathcal{A}}(G_{\mathcal{A}})$ stands as an upper bound on the minimum rank of matrices in $\mathcal{A} \cap \mathbf{S}_+^n$. Also, it is straightforward to show that $\rho_{\mathcal{A}}(G_{\mathcal{A}}) \leq \omega(G_{\mathcal{A}})$. While it is not true, in general, that $\rho_{\mathcal{A}}(G_{\mathcal{A}}) \leq \sqrt[2]{m}$, we provide examples of subspaces $\mathcal{A}$ in Section V-A, for which it holds that $\rho_{\mathcal{A}}(G_{\mathcal{A}}) < \sqrt[2]{m}$. In Section V-B, we provide an example of subspace where the inequality is reversed. We use this example to motivate the derivation of a simple method to compute a matrix in $\mathcal{A} \cap \mathbf{S}_+^n$, in polynomial time, whose rank is no greater than $\min\{\rho_{\mathcal{A}}(G_{\mathcal{A}}), \sqrt[2]{m}\}$.

A. The case of $H = G_{\mathcal{A}}$

Here, we offer two examples, where the bound obtained from Theorem 1 strictly improves upon the bounds obtained from both Propositions 1 and 2.

Example 2. Let $n = 12$, and suppose that $\mathcal{A}$ is defined by $m = 30$ linear equations. Further, suppose that $G_{\mathcal{A}}$ is given by the chordal graph in Figure 2. It follows that $\mathcal{K}(G_{\mathcal{A}}) = \{K_1, K_2\}$, where $|K_1| = 8$ and $|K_2| = 6$. Also, $|\mathcal{I}(K_1) \cap \mathcal{I}(K_2)| = 3$. One can suitably choose the weighting matrices to ensure $\mu_{\mathcal{A}}(\mathcal{I}(K_1) \setminus \mathcal{I}(K_2)) \leq 24$, imlpying

$$\min\left\{|K_1|, \sqrt[2]{m_{K_1}}\right\} = \min\left\{8, \ \sqrt[2]{27} = 6\right\} = 6.$$

Also, $\min\left\{|K_2|, \sqrt[2]{m_{K_2}}\right\} \leq 6$, and hence, we have $\rho_{\mathcal{A}}(G_{\mathcal{A}}) = 6$. For the same example, $\sqrt[2]{m} = \sqrt[2]{30} = 7$, and $\omega(G_{\mathcal{A}}) = 8$. Thus, Theorem 1 with $H = G_{\mathcal{A}}$ sharpens the upper bound obtained from both Propositions 1 and 2.

Example 3. Consider a tree (connected acyclic graph) T_n on n nodes, labelled $1, \ldots, n$. Let $\mathcal{N}(i)$ define the set of nodes neighboring node i in T_n. Denote by $d_{\max}(T_n)$, the maximum degree of a node in T_n, i.e., the maximum of $|\mathcal{N}(i)|$ over i in T_n. Next, consider an affine subspace $\mathcal{A}$ defined by $m = n$ linear matrix equations as in (1), where the weighting matrix A_i has the property that $[A_i]_{jk} \neq 0$ if and only if j and k both belong to $\{i\} \cup \mathcal{N}(i)$. Then, the associated graph $G_{\mathcal{A}}$ is chordal. For this linear system, it can be shown that $\rho_{\mathcal{A}}(G_{\mathcal{A}})$ satisfies

$$\rho_{\mathcal{A}}(G_{\mathcal{A}}) \ \leq \ \sqrt[2]{3d_{\max}(T_n) + 2}.$$

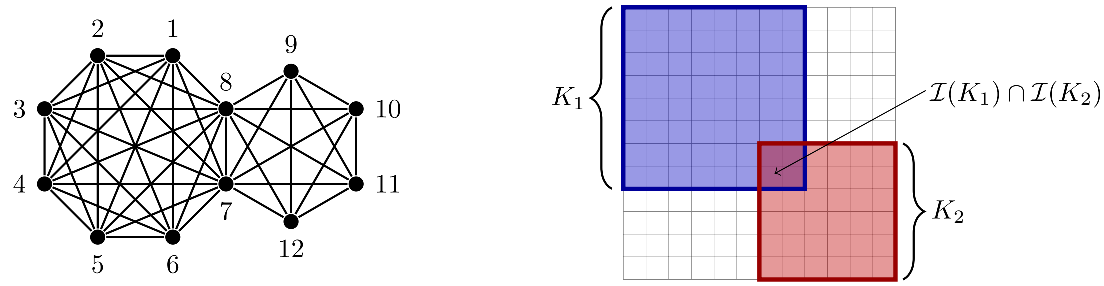

Fig. 2: The sparsity pattern of a 12-node chordal graph G for Example 2. The two maximal cliques are represented as principal minors of a G-partial matrix on the right.

The proof is omitted due to space constraints. To compare $\rho_{\mathcal{A}}(G_{\mathcal{A}})$ with the bounds from Propositions 1 and 2, notice that $\omega(G_{\mathcal{A}}) = d_{\max}(T_n) + 1$ and $\sqrt[2]{m} = \sqrt[2]{n}$. Our aim is to compare the bounds when T_n is uniformly sampled from the set of all n^{n-2} trees on n labeled nodes in the large graph limit, i.e., as $n \to \infty$. To that end, Theorem 3 in [46] implies that $d_{\max}(T_n)$ scales asymptotically almost surely as $\log n / \log \log n$. As a result, both ratios $\frac{\rho_{\mathcal{A}}(G_{\mathcal{A}})}{\omega(G_{\mathcal{A}})}$ and $\frac{\rho_{\mathcal{A}}(G_{\mathcal{A}})}{\sqrt[2]{n}}$ almost surely vanish as $n \to \infty$. In other words, the upper bound obtained using $\rho_{\mathcal{A}}(G_{\mathcal{A}})$ far outperforms the bounds from both Propositions 1 and 2 in the large graph limit for this family of problems.

B. The general case of $H \supseteq G_{\mathcal{A}}$

For a chordally sparse linear system defined by an affine subspace $\mathcal{A}$, the bound obtained from Proposition 1 may be tighter than $\rho_{\mathcal{A}}(G_{\mathcal{A}})$. However, there always exists a chordal extension H of $G_{\mathcal{A}}$ for which $\rho_{\mathcal{A}}(H)$ is at least as tight as the bound from Proposition 1, as in the following example.

Example 4. Consider the chordal graph G on $n = 26$ nodes, whose maximal cliques are described as follows. Suppose $\mathcal{K}(G) = \{K_1, K_2, K_3\}$, where $K_1 = \{1, \ldots, 13\}$, $K_2 = \{3, \ldots, 15\}$, and $K_3 = \{14, \ldots, 26\}$. Further, suppose that the affine subspace $\mathcal{A}$ is defined by $m = 15$ linear matrix equations. Choose the weighting matrices such that only the $\mathcal{I}(K_3^c) \setminus \mathcal{I}(K_3)$ entries in A_i are nonzero for $i = 1, \ldots, 5$. Also, only the $\mathcal{I}(K_1^c) \setminus \mathcal{I}(K_1)$ entries are nonzero in A_i for $i = 6, \ldots, 15$. It is easy to verify that $G_{\mathcal{A}} = G$. For this example, we have $\sqrt[2]{m} = 5 < \rho_{\mathcal{A}}(G_{\mathcal{A}}) = 12$. Now, consider the chordal extension H of $G_{\mathcal{A}}$ with two maximal cliques, comprising the nodes $\{1, \ldots, 15\}$ and $\{14, \ldots, 26\}$, respectively. One can verify that $\rho_{\mathcal{A}}(H) = 4 < \min\{\rho_{\mathcal{A}}(G_{\mathcal{A}}), \sqrt[2]{m}\}$.

Ideally, one would like to compute the optimal chordal extension H of $G_{\mathcal{A}}$ for which $\rho_{\mathcal{A}}(H) = \rho^*_{\mathcal{A}}$. Its calculation appears to be intractable, however. Hence, we resort to computing an upper bound on the minimum rank by restricting ourselves to a smaller collection of chordal extensions of $G_{\mathcal{A}}$. By doing so, we are able to compute an upper bound in polynomial time that does not exceed $\min\{\omega(G_{\mathcal{A}}), \sqrt[2]{m}\}$.

A polynomial time computable upper bound: Suppose $G_{\mathcal{A}}$ is a chordal graph on n nodes for an affine subspace $\mathcal{A}$. It follows from [38, Lemma 6], [25, Lemma 4] that there exists a finite sequence of chordal graphs $G_{\mathcal{A}} = G_0 \subseteq G_1 \subseteq \ldots \subseteq G_\ell$, where G_ℓ is the complete graph on n nodes, and each member in the sequence differs from the one preceding it by exactly one edge (cf. Appendix A). Define $\mathcal{H} := \{G_0, \ldots, G_\ell\}$. Now, $|\mathcal{H}|$ is bounded by the maximum number of edges in an n-node graph, i.e., $|\mathcal{H}| \leq \frac{1}{2}n(n+1)$, implying $\min_{H \in \mathcal{H}} \rho_{\mathcal{A}}(H)$ can be computed in polynomial time. Furthermore, $\mathcal{H}$ contains $G_{\mathcal{A}}$ and the complete graph on n nodes, and hence,

$$\min_{H \in \mathcal{H}} \rho_{\mathcal{A}}(H) \;\leq\; \min\{\rho_{\mathcal{A}}(G_{\mathcal{A}}), \sqrt[2]{m}\} \;\leq\; \min\{\omega(G_{\mathcal{A}}), \sqrt[2]{m}\}.$$

VI. Conclusion

In this paper, we derive a novel upper bound (cf. Theorem 1) on the minimum rank of matrices belonging to an affine slice of the positive semidefinite cone, when the weighting matrices defining the affine subspace are sparse. When the collective sparsity pattern of the weighting matrices defining the affine subspace is given by a chordal graph, our upper bound depends on: (1) the sizes of the maximal cliques of that graph and the maximal overlap among them, and (2) how the sparsity pattern of individual weighting matrices distributes across the maximal cliques of that graph. When the graph is not chordal, the bound relies on a chordal extension of the above graph.

References

[1] J. Peng and Y. Wei, “Approximating k-means-type clustering via semidefinite programming,” *SIAM Journal on Optimization*, vol. 18, no. 1, pp. 186–205, 2007.

[2] S. Shalev-Shwartz, A. Gonen, and O. Shamir, “Large-scale convex minimization with a low-rank constraint,” *arXiv preprint arXiv:1106.1622*, 2011.

[3] B. Recht, M. Fazel, and P. A. Parrilo, “Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization,” *SIAM review*, vol. 52, no. 3, pp. 471–501, 2010.

[4] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, “Quantum state tomography via compressed sensing,” *Physical review letters*, vol. 105, no. 15, p. 150401, 2010.

[5] M. Fazel, H. Hindi, and S. Boyd, “Rank minimization and applications in system theory,” in *American Control Conference, 2004. Proceedings of the 2004*, vol. 4. IEEE, 2004, pp. 3273–3278.

[6] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Foundations of Computational mathematics*, vol. 9, no. 6, pp. 717–772, 2009.

[7] E. J. Candès and T. Tao, "The power of convex relaxation: Near-optimal matrix completion," *Information Theory, IEEE Transactions on*, vol. 56, no. 5, pp. 2053–2080, 2010.

[8] M. Fazel, "Matrix rank minimization with applications," *Elec Eng Dept Stanford University*, vol. 54, pp. 1–130, 2002.

[9] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM journal on computing*, vol. 24, no. 2, pp. 227–234, 1995.

[10] B. Recht, W. Xu, and B. Hassibi, "Null space conditions and thresholds for rank minimization," *Mathematical programming*, vol. 127, no. 1, pp. 175–202, 2011.

[11] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 9, pp. 589–592, 2008.

[12] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, "The convex geometry of linear inverse problems," *Foundations of Computational mathematics*, vol. 12, no. 6, pp. 805–849, 2012.

[13] A. Barvinok, "A remark on the rank of positive semidefinite matrices subject to affine constraints," *Discrete & Computational Geometry*, vol. 25, no. 1, pp. 23–31, 2001.

[14] G. Pataki, "The geometry of semidefinite programming," in *Handbook of semidefinite programming*. Springer, 2000, pp. 29–65.

[15] F. Alizadeh, J.-P. A. Haeberly, and M. L. Overton, "Complementarity and nondegeneracy in semidefinite programming," *Mathematical Programming*, vol. 77, no. 1, pp. 111–128, 1997.

[16] W. Ai, Y. Huang, and S. Zhang, "On the low rank solutions for linear matrix inequalities," *Mathematics of Operations Research*, vol. 33, no. 4, pp. 965–975, 2008.

[17] P. Gahinet and P. Apkarian, "A linear matrix inequality approach to H_∞ control," *International journal of robust and nonlinear control*, vol. 4, no. 4, pp. 421–448, 1994.

[18] T. Iwasaki and R. E. Skelton, "All controllers for the general H_∞ control problem: LMI existence conditions and state space formulas," *Automatica*, vol. 30, no. 8, pp. 1307–1317, 1994.

[19] M. T. Chu, R. E. Funderlic, and R. J. Plemmons, "Structured low rank approximation," *Linear algebra and its applications*, vol. 366, pp. 157–172, 2003.

[20] M. Fazel, T. K. Pong, D. Sun, and P. Tseng, "Hankel matrix rank minimization with applications to system identification and realization," *SIAM Journal on Matrix Analysis and Applications*, vol. 34, no. 3, pp. 946–977, 2013.

[21] S. Bose, D. F. Gayme, K. M. Chandy, and S. H. Low, "Quadratically constrained quadratic programs on acyclic graphs with application to power flow," *Control of Network Systems, IEEE Transactions on*, vol. 2, no. 3, pp. 278–287, 2015.

[22] S. Low, "Convex relaxation of optimal power flow: Part I: Formulations and equivalence," *Control of Network Systems, IEEE Transactions on*, vol. 1, no. 1, pp. 15–27, 2014.

[23] R. Madani, G. Fazelnia, S. Sojoudi, and J. Lavaei, "Finding low-rank solutions of sparse linear matrix inequalities using convex optimization," *Preprint*, 2014.

[24] A. Varvitsiotis, "Combinatorial conditions for low rank solutions in semidefinite programming," School of Economics and Management, Tech. Rep., 2013.

[25] R. Grone, C. R. Johnson, E. M. Sá, and H. Wolkowicz, "Positive definite completions of partial hermitian matrices," *Linear algebra and its applications*, vol. 58, pp. 109–124, 1984.

[26] Y. Sun, M. S. Andersen, and L. Vandenberghe, "Decomposition in conic optimization with partially separable structure," *SIAM Journal on Optimization*, vol. 24, no. 2, pp. 873–897, 2014.

[27] L. Vandenberghe and M. S. Andersen, "Chordal graphs and semidefinite optimization," no. 4, pp. 241–433, 2014.

[28] V. S. Mai, D. Maity, B. Ramasubramanian, and M. C. Rotkowitz, *Convex Methods for Rank-Constrained Optimization Problems*, ch. 17, pp. 123–130. [Online]. Available: http://epubs.siam.org/doi/abs/10.1137/1.9781611974072.18

[29] Z. Luo, W. Ma, A. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 20 –34, May 2010.

[30] S. Burer and R. D. Monteiro, "A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization," *Mathematical Programming*, vol. 95, no. 2, pp. 329–357, 2003.

[31] ——, "Local minima and convergence in low rank semidefinite programming," *Mathematical Programming*, vol. 103, no. 3, pp. 427–444, 2005.

[32] S. Bhojanapalli, A. Kyrillidis, and S. Sanghavi, "Dropping convexity for faster semi-definite optimization," *arXiv preprint arXiv:1509.03917*, 2015.

[33] Y. Chen and M. J. Wainwright, "Fast low-rank estimation by projected gradient descent: General statistical and algorithmic guarantees," *arXiv preprint arXiv:1509.03025*, 2015.

[34] J. R. Blair and B. Peyton, "An introduction to chordal graphs and clique trees," in *Graph theory and sparse matrix computation*. Springer, 1993, pp. 1–29.

[35] D. Fulkerson and O. Gross, "Incidence matrices and interval graphs," *Pacific journal of mathematics*, vol. 15, no. 3, pp. 835–855, 1965.

[36] P. Buneman, "A characterisation of rigid circuit graphs," *Discrete mathematics*, vol. 9, no. 3, pp. 205–212, 1974.

[37] F. Gavril, "Algorithms for a maximum clique and a maximum independent set of a circle graph," *Networks*, vol. 3, no. 3, pp. 261–273, 1973.

[38] D. J. Rose, R. E. Tarjan, and G. S. Lueker, "Algorithmic aspects of vertex elimination on graphs," *SIAM Journal on computing*, vol. 5, no. 2, pp. 266–283, 1976.

[39] A. I. Barvinok, "Problems of distance geometry and convex properties of quadratic maps," *Discrete & Computational Geometry*, vol. 13, no. 1, pp. 189–202, 1995.

[40] G. Pataki, "Cone-LP's and semidefinite programs: Geometry and a simplex-type method," *Integer programming and combinatorial optimization*, pp. 162–174, 1996.

[41] A. Barvinok, *A course in convexity*. American Mathematical Soc., 2002, vol. 54.

[42] M. Laurent and A. Varvitsiotis, "The gram dimension of a graph," in *Combinatorial Optimization*. Springer, 2012, pp. 356–367.

[43] R. L. Smith, "The positive definite completion problem revisited," *Linear Algebra and its Applications*, vol. 429, no. 7, pp. 1442–1452, 2008.

[44] M. F. Anjos and J. B. Lasserre, *Handbook on semidefinite, conic and polynomial optimization*. Springer Science & Business Media, 2011, vol. 166.

[45] S. Arnborg, D. G. Corneil, and A. Proskurowski, "Complexity of finding embeddings in a k-tree," *SIAM Journal on Algebraic Discrete Methods*, vol. 8, no. 2, pp. 277–284, 1987.

[46] J. W. Moon *et al.*, "On the maximum degree in a random tree." *The Michigan Mathematical Journal*, vol. 15, no. 4, pp. 429–432, 1968.

[47] D. R. Shier, "Some aspects of perfect elimination orderings in chordal graphs," *Discrete Applied Mathematics*, vol. 7, no. 3, pp. 325–331, 1984.

APPENDIX

A. Constructing a sequence of chordal extensions

Let G_0 be a chordal graph on n nodes. Here, we provide a polynomial time algorithm to construct a sequence of chordal graphs $G_1, \ldots, G_\ell$, that satisfies $G_0 \subseteq G_1 \subseteq \ldots \subseteq G_\ell$, where G_ℓ is the complete graph on n nodes. Each graph in the sequence differs from the one preceding it by one edge. Construct a so-called *perfect elimination ordering (PEO)* of G_0 in polynomial time. See [47] for the definition and a polynomial time construction of a PEO for a chordal graph. Call this sequence $v_1, \ldots, v_n$. Find the highest index $i \in \{1, \ldots, n\}$, such that the induced subgraph of G_0 on the nodes $\{v_i, \ldots, v_n\}$ is not a clique. Then, find the smallest index $j \in \{i+1, \ldots, n\}$, such that (i, j) is not an edge in G_0. Add the edge (i, j) to G_0 to obtain G_1. One can verify that $v_1, \ldots, v_n$ is also a PEO of G_1. Repeat the process to add an edge to G_1, and so on.