Degeneracy in Maximal Clique Decomposition for Semidefinite Programs

Arvind U. Raghunathan and Andrew V. Knyazev
Mitsubishi Electric Research Laboratories
201 Broadway, Cambridge, MA 02139
Email:raghunathan@merl.com,knyazev@merl.com

September 29, 2015

Abstract

Exploiting sparsity in Semidefinite Programs (SDP) is critical to solving large-scale problems. The chordal completion based maximal clique decomposition is the preferred approach for exploiting sparsity in SDPs. In this paper, we show that the maximal clique-based SDP decomposition is primal degenerate when the SDP has a low rank solution. We also derive conditions under which the multipliers in the maximal clique-based SDP formulation is not unique. Numerical experiments demonstrate that the SDP decomposition results in the schur-complement matrix of the Interior Point Method (IPM) having higher condition number than for the original SDP formulation.

1 Introduction

Semidefinite programming (SDP) is a subfield of convex optimization concerned with the optimization of a linear objective function over the intersection of the cone of symmetric positive semidefinite matrices with an affine space. Many problems in operations research and combinatorial optimization can be modeled or approximated as SDPs [3, 12]. For a SDP defined over the set of $n \times n$ symmetric matrices and affine subspace of dimension m, the number of unknowns in the problem grows as $O(n^2)$. Since the seminal work of Nesterov and Nemirovskii [9], Interior Point Methods (IPMs) have becomes the preferred approach for solving SDPs. The complexity of the step computation in IPM is typically $O(mn^3 + m^2n^2)$ [10].

Given the quadratic growth in m, n of the computational cost, it is imperative to exploit problem structure in solving large-scale SDPs. For SDPs modeling practical applications, the data matrices involved are typically *sparse*. Denote by, $\mathsf{N} = \{1, \ldots, n\}$ and by $\mathsf{E} = \{(i, j) \mid i \neq j,\ (i, j)-$th entry of some data matrix is non-zero$\}$. The set E, also called the *aggregate sparsity pattern*, represents the non-zero entries in the objective and constraint matrices, that is the sparsity in the problem data. Consequently, only the entries of the matrix variable corresponding to the aggregate sparsity pattern are involved in the problem. From the computational stand-point it is desirable to work only with such entries to reduce the number of unknowns in the problem from $O(n^2)$ to $O(|E|)$. However, the semidefinite constraint couples all of the entries of the symmetric matrix. Fukuda et al [4] exploit the result of Grone et al [6] to decompose the SDP defined on $n \times n$ symmetric matrices into smaller sized matrices. Grone et al [6, Theorem 7] states that for a graph $\mathsf{G}(\mathsf{N}, \mathsf{E})$ that is chordal: the positive semidefinite condition on $n \times n$ matrix is equivalent to positive semidefinite condition on submatrices corresponding to the maximal cliques that cover all the nodes and edges in the graph $\mathsf{G}(\mathsf{N}, \mathsf{E})$. Nakata et al [8] implemented the decomposition within a SDP software package SDPA [13] and demonstrated the scalability of the approach. More recently, the authors of SDPA have also extended the implementation to take advantage of multi-core architectures [14]. More recently, Kim and Kojima [7] extended this approach for solving semidefinite relaxations of polynomial optimization problems.

1.1 Our Contribution

In this paper, we study the properties of the *conversion approach* of [4, 8] which converts the original SDP into an SDP with multiple sub-matrices and additional inequality constraints. We show that the SDP resulting from the conversion approach is *primal degenerate* when the SDP solution has low-rank. We show that this can occur even when the solution to the original SDP is *primal non-degenerate*. Thus, this degeneracy is a consequence of the conversion approach. We also derive conditions under which the dual multipliers are not unique. We demonstrate through numerical experiments that condition numbers of schur-complement matrix of IPM are much higher for the conversion approach as compared with the original SDP formulation. To the best of our knowledge, this is the first result describing the degeneracy of the conversion approach.

The rest of the paper is organized as follows. § 2 introduces the SDP formulation and the maximal clique decomposition. The conversion of approach of [4] is described in § 3. § 4 proves the primal degeneracy and dual non-uniqueness of the conversion approach. Numerical experiments validating the results are presented in § 5, followed by conclusions in § 6.

1.2 Notation

In the following, $\mathbb{R}$ denotes the set of reals and $\mathbb{R}^n$ is the space of n dimensional column vectors. For a vector $x \in \mathbb{R}^n$, $[x]_i$ denotes the i-th component of x and $0_n \in \mathbb{R}^n$ denotes the zero vector, $e_i \in \mathbb{R}^n$ the vector with 1 for the i-th component and 0 otherwise. The notation $\text{diag}(\lambda_1, \ldots, \lambda_n)$ denotes a diagonal matrix with the values λ_i on the diagonal. Given a vector $v \in \mathbb{R}^n$ and subset $\mathsf{C} \subseteq \{1, \ldots, n\}$, v_{C} denotes the subvector from $[v]_i$ for $i \in \mathsf{C}$. $\mathbb{S}^n$ denotes the set of $n \times n$ real symmetric matrices and $\mathbb{S}^n_+$ ($\mathbb{S}^n_{++}$) denotes the set of $n \times n$ real symmetric positive semi-definite (definite) matrices. Further, $A \succeq (\succ) 0$ denotes that $A \in \mathbb{S}^n_+ (\mathbb{S}^n_{++})$. For a matrix $A \in \mathbb{S}^n$, $[A]_{ij}$ denotes the (i,j)-th entry of the matrix A and $\text{rank}(A)$ denotes the rank of A. For a matrices $A_1, A_2 \in \mathbb{S}^n$, $\text{range}(A_1, A_2)$ denotes the subspace of symmetric matrices spanned by A_1, A_2. Denote by $\mathsf{N} = \{1, \ldots, n\}$. The notation $\bullet$ denotes the standard trace inner product between symmetric matrices $A \bullet B = \sum_{i=1}^n \sum_{j=1}^n [A]_{ij}[B]_{ij}$ for $A, B \in \mathbb{S}^n$. For sets $\mathsf{C}_s, \mathsf{C}_t \subseteq \mathsf{N}$ and $A \in \mathbb{S}^n$, $A_{\mathsf{C}_s \mathsf{C}_t}$ is a $|\mathsf{C}_s| \times |\mathsf{C}_t|$ submatrix A formed by removing rows and columns of A that are not in C_s, C_t respectively.

1.3 Background on Graph Theory [2]

In this paper we only consider undirected graphs. Given a graph $\mathsf{G}(\mathsf{N}, \mathsf{F})$, a *cycle in* F is a sequence of vertices $\{i_1, i_2, \ldots, i_q\}$ such that $i_j \neq i_{j'}$, $(i_j, i_{j+1}) \in \mathsf{F}$ and $(i_q, i_1) \in \mathsf{F}$. The cycle in F with q vertices is called a cycle of length q. Given a cycle $\{i_1, \ldots, i_q\}$ in a F, a *chord* is an edge $(i_j, i_{j'})$ for $|j - j'| > 1$. A graph $\mathsf{G}(\mathsf{N}, \mathsf{F})$ is said to be *chordal* if every cycle of length greater than 3 has a chord. Given $\mathsf{G}(\mathsf{N}, \mathsf{F})$, $\mathsf{F}' \supseteq \mathsf{F}$ is called a *chordal extension* if the graph $\mathsf{G}'(\mathsf{N}, \mathsf{F}')$ is chordal. Given a graph $\mathsf{G}(\mathsf{N}, \mathsf{F})$, $\mathsf{C} \subset \mathsf{F}$ is called a *clique* if it satisfies the property that $(i,j) \in \mathsf{F}$ for all $i, j \in \mathsf{C}$. A clique C is *maximal* if there does not exist clique $\mathsf{C}' \supset \mathsf{C}$. For a chordal graph, the maximal cliques can be arranged as a tree, called *clique tree*, $\mathcal{T}(\mathcal{N}, \mathcal{E})$ in which $\mathcal{N} = \{\mathsf{C}_1, \ldots, \mathsf{C}_\ell\}$ and $(\mathsf{C}_s, \mathsf{C}_t) \in \mathcal{E}$ are edges between the cliques.

1.4 Matrix Terminology [4]

For a set $\mathsf{F} \subset \mathsf{N} \times \mathsf{N}$, $\mathbb{S}^n(\mathsf{F})$ denote the set of symmetric $n \times n$ matrices with only entries in F specified. A matrix $\bar{X} \in \mathbb{S}^n(\mathsf{F})$ is called a *symmetric partially specified matrix.* A completion $X \in \mathbb{S}^n$ is called a *completion* of $\bar{X} \in \mathbb{S}^n(\mathsf{F})$ if $[X]_{ij} = [\bar{X}]_{ij}$ for all $(i,j) \in \mathsf{F}$. A completion $X \in \mathbb{S}^n$ of $\bar{X} \in \mathbb{S}^n(\mathsf{F})$ that is positive semidefinite (definite) is said to be a *positive semidefinite (definite) completion* of $\bar{X}$.

2 Maximal Clique Decomposition in SDP

Consider the following SDP:

$$\begin{aligned} \min_{X \in \mathbb{S}^n} \ & A_0 \bullet X \\ \text{s.t.} \ & A_p \bullet X = b_p \ \forall\, p = 1, \ldots, m \\ & X \succeq 0 \end{aligned} \tag{1}$$

where $A_p \in \mathbb{S}^n$. Denote by $\mathsf{E} = \{(i,j) \mid i \neq j,\ [A_p]_{ij} \neq 0 \text{ for some } 0 \leq p \leq m\}$. The set E, also called the *aggregate sparsity pattern* [4], represents the non-zero entries in the objective and constraint matrices, that is the sparsity in the problem data. Clearly, only the entries $[X]_{jk}$ for $(j,k) \in \mathsf{E}$ feature in the objective and equality constraints in (1). In a number of practical applications, $|\mathsf{E}| << n^2$. From a computational standpoint, it is desirable to work only with $[X]_{jk}$ for $(j,k) \in \mathsf{E}$. In other words, we want to solve

$$\begin{aligned} \min_{\bar{X} \in \mathbb{S}^n(\mathsf{E})} \ & \sum_{(i,j) \in \mathsf{E}} [A_0]_{ij} [\bar{X}]_{ij} \\ \text{s.t.} \ & \sum_{(i,j) \in \mathsf{E}} [A_p]_{ij} [\bar{X}]_{ij} = b_p \ \forall\, p = 1, \ldots, m \\ & \bar{X} \text{ has a positive semidefinite completion.} \end{aligned} \tag{2}$$

The result of [6] provides the conditions under which such a completion exist. We state this below in a form convenient for further development as in [4, Theorem 2.5].

Lemma 1 ([4, Theorem 2.5])**.** *Let $\mathsf{G}(\mathsf{N}, \mathsf{F})$ be a chordal graph and let $\{\mathsf{C}_1, \ldots . \mathsf{C}_\ell\}$ be the family of all maximal cliques. Then, $\bar{X} \in \mathbb{S}^n(\mathsf{F})$ has a positive semidefinite (definite) completion if and only if $\bar{X}$ satisfies*

$$\bar{X}_{\mathsf{C}_s \mathsf{C}_s} \succeq 0 (\succ 0) \ \forall\, s = 1, \ldots, \ell. \tag{3}$$

Using Lemma 1 Fukuda et al [4] proposed the *conversion approach* which we describe next.

3 Conversion Approach

Given the graph $\mathsf{G}(\mathsf{N}, \mathsf{E})$, with E the aggregate sparsity pattern of SDP (1), the conversion approach [4] proceeds by : (a) computing a chordal extension $\mathsf{F} \supseteq \mathsf{E}$; (b) the set of maximal cliques $\mathcal{N} = \{\mathsf{C}_1, \ldots, \mathsf{C}_\ell\}$ of the graph $\mathsf{G}(\mathsf{N}, \mathsf{F})$ are identified; (c) the clique tree $\mathcal{T}(\mathcal{N}, \mathcal{E})$ is computed; and (d) a SDP is posed in terms of matrices defined on the set of maximal cliques that is quivalent to SDP in (1). Additional equality constraints are introduced to equate the overlapping entries in the maximal cliques. Prior to stating the SDP formulation we introduce notation that facilitates further development. Denote,

$$
\begin{aligned}
&\sigma_s : \mathsf{N} \to \{1, \ldots, |\mathsf{C}_s|\} \text{ mapping the original indices} \\
&\qquad \text{to the ordering in the clique } \mathsf{C}_s \\
[A_{s,p}]_{\sigma_s(i)\sigma_s(j)} &= \begin{cases} [A_p]_{ij} \text{ if } s = \min\{t \mid (i,j) \in \mathsf{C}_t\} \\ \quad 0 \text{ otherwise} \end{cases} \\
E_{s,ij} &= \frac{1}{2}\left(e_{\sigma_s(i)} e_{\sigma_s(j)}^T + e_{\sigma_s(j)} e_{\sigma_s(i)}^T\right) \ \forall\, i,j \in \mathsf{C}_s \\
(s,t) \in \mathcal{T} &\iff (\mathsf{C}_s, \mathsf{C}_t) \in \mathcal{E} \\
\mathsf{C}_{st} &= \mathsf{C}_s \cap \mathsf{C}_t
\end{aligned} \tag{4}
$$

where $e_{\sigma_s(i)} \in \mathbb{R}^{|\mathsf{C}_s|}$. The SDP in (1) can be equivalently posed using the above notation as,

$$
\begin{aligned}
\min_{X_s \in \mathbb{S}^{|\mathsf{C}_s|}} \ & \sum_{s=1}^{\ell} A_{s,0} \bullet X_s \\
\text{s.t.} \ & \sum_{s=1}^{\ell} A_{s,p} \bullet X^s = b_p && \forall\, p = 1, \ldots, m \\
& E_{s,ij} \bullet X_s = E_{t,ij} \bullet X_t && \forall\, i \leq j, i,j \in \mathsf{C}_{st}, \\
& && (s,t) \in \mathcal{E} \\
& X_s \succeq 0 && \forall\, s = 1, \ldots, \ell.
\end{aligned} \tag{5}
$$

We refer to the SDP in (5) as the *conversion SDP* for sake of brevity. In the following we analyze the conversion approach for linear programming. Formal proofs for primal degeneracy of conversion SDP are left for § 4.

3.1 Intuition for Primal Degeneracy

To motivate the primal degeneracy of conversion SDP we provide a conversion approach inspired decomposition for linear programs (LPs). Consider a LP of the form,

$$
\begin{aligned}
\min_{x \in \mathbb{R}^n} \ & a_0^T x \\
\text{s.t.} \ & a_p^T x = b_p \ \forall\, p = 1, \ldots, m \\
& x \geq 0
\end{aligned} \tag{6}
$$

where $a_i \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$. Suppose we decompose the LP (6) using the sets in $\{\mathsf{C}_1, \ldots, \mathsf{C}_\ell\}$ as,

$$
\begin{aligned}
\min_{x_s \in \mathbb{R}^{|\mathsf{C}_s|}} \ & \sum_{s=1}^{l} a_{0,s}^T x_s \\
\text{s.t.} \ & \sum_{s=1}^{l} a_{s,p}^T x_s = b_p \ \forall\, p = 1, \ldots, m \\
& [x_s]_{\sigma_s^{\mathrm{LP}}(i)} = [x_t]_{\sigma_t^{\mathrm{LP}}(i)} \ \forall\, i \in \mathsf{C}_{st}, (s,t) \in \mathcal{E} \\
& x_s \geq 0 \ \forall\, s = 1, \ldots, \ell
\end{aligned} \tag{7}
$$

where

$$\sigma_s^{\mathrm{LP}} : \mathsf{N} \to \{1, \ldots, |\mathsf{C}_s|\}$$
$$[a_{s,p}]_{\sigma_s^{\mathrm{LP}}(i)} = \begin{cases} [a_p]_i & \text{if } s = \min\{t | i \in \mathsf{C}_t\} \\ 0 & \text{otherwise} \end{cases} .$$

With the above definition of the matrices it is easy to see that the LPs in (6) and (7) are equivalent. Further, if x^* is an optimal solution to LP (6) then, $x_s^* = x_{\mathsf{C}_s}^*$ is optimal for (7). Suppose, there exists $i \in \mathsf{C}_s \cap \mathsf{C}_t$ for which $[x^*]_j = 0$ then, the set of constraints

$$[x_s]_{\sigma_s^{\mathrm{LP}}(i)} = [x_t]_{\sigma_t^{\mathrm{LP}}(i)}$$
$$[x_s]_{\sigma_s^{\mathrm{LP}}(i)} = 0, [x_t]_{\sigma_t^{\mathrm{LP}}(i)} = 0$$

are linearly dependent. The linear dependency of the constraints can be avoided if the nonnegative bounds on shared entries are enforced exactly once for each index i. For example, the non negativity constraints in (7) can be enforced for each $i \in \mathsf{N}$:

$$[x_s]_{\sigma_s^{\mathrm{LP}}(i)} \geq 0 \text{ if } s = s^{\min}(i).$$

In summary the degeneracy occurs due to a shared element activating the bound at the solution. In a direct analogy, the conversion SDP in (5) primal degenerate when, $\mathrm{rank}(X_{\mathsf{C}_{st}\mathsf{C}_{st}}) < |\mathsf{C}_{st}|$. This degeneracy is directly attributable to the duplication of the semidefinite constraints for the submatrix $X_{\mathsf{C}_{st}\mathsf{C}_{st}}$ in both $X_s \succeq 0$ and $X_t \succeq 0$ for every pair of $(s,t) \in \mathcal{T} : s \neq t$. Unfortunately, the duplication of the semidefinite constraints cannot be avoided in the case of SDP without losing the linearity. We provide formal arguments for the degeneracy and dual multiplicity of the conversion SDP (5) in the following section.

4 Primal Degeneracy & Dual Non-uniqueness of Conversion Approach

We review the conditions for primal non-degeneracy and dual uniqueness for the SDP (1) introduced by Alizadeh et al [1] in § 4.1. § 4.2 proves the primal degeneracy result for the conversion approach.

4.1 Primal Nondegeneracy and Dual Uniqueness in SDPs

Suppose $X \in \mathbb{S}^n$ with $\mathrm{rank}(X) = r$ with eigenvalue decomposition $X = Q\mathrm{diag}(\lambda_1, \ldots, \lambda_r, 0, \ldots, 0)Q^T$ then, the tangent space $\mathbb{T}_X$ of rank-r symmetric matrices is

$$\mathbb{T}_X = \left\{ Q \begin{bmatrix} U & V \\ V^T & 0 \end{bmatrix} Q^T \,\middle|\, U \in \mathbb{S}^r, V \in \mathbb{R}^{r \times (n-r)} \right\}. \tag{8}$$

The null space of equality constraints $\mathbb{N}_A$ is,

$$\mathbb{N}_A = \{ Y \in \mathbb{S}^n | A_p \bullet Y = 0 \; \forall \, p = 1, \ldots, m \} . \tag{9}$$

Definition 1 ([1]). *Suppose X is primal feasible for (1) with rank$(X) = r$ then, X is primal nondegenerate if*

$$\mathbb{T}_X + \mathbb{N}_A = \mathbb{S}^n. \tag{10}$$

Lemma 2 ([1, Theorem 2])**.** *Suppose $X^* \in \mathbb{S}^n$ is primal nondegenerate and optimal for* (1)*. Then the optimal dual multipliers (ζ^*, S^*), for the equality and positive semidefinite constraints respectively, satisfying the first order optimality conditions* (11) *are unique,*

$$\begin{aligned} A_0 + \sum_{p=1}^{m} \zeta_p^* A_p - S^* = 0 \\ A_p \bullet X^* = b_p \; \forall \; p = 1, \dots, m \\ X^*, S^* \succeq 0, X^* S^* = 0. \end{aligned} \tag{11}$$

4.2 Primal Degeneracy of Conversion SDP

Assumption 1. *The SDP* (1) *has an optimal solution X^* with rank$(X^*) < |\mathsf{C}_{st}|$ for some $(s,t) \in \mathcal{E}$.*

In the following we denote by X_s^* the optimal solution to the conversion SDP (5). The following result is immediate.

Lemma 3. *rank$(X_s^*) \leq$ rank(X^*) $\forall$ $s = 1, \dots, \ell$.*

Proof. By definition, $X_s^* = X^*_{\mathsf{C}_s \mathsf{C}_s}$ is a principal submatrix of X^*. The claim follows by noting that the rank of any principal sub-matrix cannot exceed that of the original matrix. □

The following result characterizes the eigenvectors of the matrices X_s^*, X_t^* for cliques s, t satisfying Assumption 1. Without loss of generality and for ease of presentation, we assume that

$$\sigma_s(i) = \sigma_t(i) \text{ and } 1 \leq \sigma_s(i) \leq |\mathsf{C}_{st}| \; \forall \; i \in \mathsf{C}_{st}. \tag{12}$$

Lemma 4. *Suppose Assumption 1 holds for cliques s, t. Then, there exists $u \in \mathbb{R}^{|\mathsf{C}_{st}|}$ such that $v_s = [u^T \; 0^T_{|\mathsf{C}_s \setminus \mathsf{C}_{st}|}]^T$ is a 0-eigenvector of X_s^* and $v_t = [u^T \; 0^T_{|\mathsf{C}_t \setminus \mathsf{C}_{st}|}]^T$ is a 0-eigenvector of X_t^*.*

Proof. From Lemma 3, we have that rank(X_s^*), rank$(X_t^*) <$ rank$(X^*) < |\mathsf{C}_{st}|$ where the second inequality follows from Assumption 1. Using the arguments in the proof of Lemma 3 we have that the submatrix of X_s^*, X_t^* corresponding to C_{st} must have rank smaller than $|\mathsf{C}_{st}|$. Hence, there exists a vector $u \in \mathbb{R}^{|\mathsf{C}_{st}|}$ that lies in the nullspace of the principal submatrix. Taking the right and left products with v_s of the matrix X_s^*,

$$\begin{aligned} v_s^T X^{s*} v_s = u^T X_{s1}^* u = 0 \implies \frac{v_s^T X_s^* v_s}{v_s^T v_s} = 0 \\ \implies v_s \text{ is a } 0-\text{eigenvector of } X_s^* \end{aligned}$$

where X_{s1}^* denotes the submatrix of X_s^* corresponding to the indices of C_{st}. The claim on X_t^* can be proved similarly and this completes the proof. □

The tangent space for the matrices X_s is defined as,

$$\mathbb{T}_{s,X} = \left\{ Q_s \begin{bmatrix} U & V \\ V^T & 0 \end{bmatrix} Q_s^T \,\middle|\, U \in \mathbb{S}^{r_s}, V \in \mathbb{R}^{r_s \times (|\mathsf{C}_s| - r_s)} \right\} \tag{13}$$

where $X^s = Q^s \text{diag}(\lambda_1, \dots, \lambda_{r_s}, 0, \dots, 0)\,(Q^s)^T$ and $r_s = \text{rank}(X_s)$. The tangent space for the conversion SDP 5 is denoted by $\mathbb{T}_X = \cup_{s=1}^{\ell} \mathbb{T}_{s,X}$. The null space of equality constraints for the conversion SDP (5) is,

$$\mathbb{N} = \left\{ Y_s \in \mathbb{S}^{|\mathsf{C}_s|} \;\middle|\; \begin{array}{l} \sum_{s=1}^{\ell} A_{s,p} \bullet Y_s = 0 \;\forall\, p = 1, \ldots, m \\ E_{s,ij} \bullet X_s = E_{t,ij} \bullet X_t \\ \forall\, i \leq j, i, j \in \mathsf{C}_{st}, (s,t) \in \mathcal{E} \end{array} \right\}. \tag{14}$$

Based on the definitions in (13),(14), the conversion SDP in (5) is primal non-degenerate if,

$$\begin{aligned} &\mathbb{T}_X + \mathbb{N} = \mathbb{S}^{\bar{n}} \iff \mathbb{T}_X^{\perp} \cap \mathbb{N}^{\perp} = \{0\} \text{ where } \bar{n} = \sum_{s=1}^{\ell} |\mathsf{C}_s| \\ &\mathbb{T}_X^{\perp} = \cup_{s=1}^{\ell} \mathbb{T}_{s,X}^{\perp} \\ &\mathbb{T}_{s,X}^{\perp} = \left\{ Q_s \begin{bmatrix} 0 & 0 \\ 0 & W \end{bmatrix} Q_s^T \;\middle|\; W \in \mathbb{S}^{|\mathsf{C}_s| - r_s} \right\} \\ &\mathbb{N}^{\perp} = \text{subspace spanned by the constraints in (5).} \end{aligned} \tag{15}$$

We can now state the main result on the primal degeneracy of the conversion SDP (5).

Theorem 1. *Suppose Assumption 1 holds. Then, the solution X_s^* of the conversion SDP* (5) *is primal degenerate.*

Proof. From the conditions for primal non-denegeracy in (15) , we have that X_s^* is primal degenerate if, there exist scalars $\alpha_p, \beta_{st,ij} \neq 0$ for some $s = 1, \ldots, \ell$ such that

$$\begin{aligned} &\sum_{p=1}^{m} \alpha_p A_{s,p} + \sum_{(s,t) \in \mathcal{E}} \sum_{i \leq j \in \mathsf{C}_{st}} \beta_{st,ij} E_{s,ij} \in \mathbb{T}_{s,X^*}^{\perp} \\ &\sum_{p=1}^{m} \alpha_p A_{t,p} - \sum_{(s,t) \in \mathcal{E}} \sum_{i \leq j \in \mathsf{C}_{st}} \beta_{st,ij} E_{t,ij} \in \mathbb{T}_{t,X^*}^{\perp}. \end{aligned} \tag{16}$$

In the following we show that $\alpha_p = 0$ and some $\beta_{st,ij} \neq 0$ for s, t satisfying Assumption 1 for which (16) holds.

In the rest of the proof s, t denote a pair of cliques $\mathsf{C}_s, \mathsf{C}_t$ satisfying Assumption 1. By Lemma 4 and the definition of $\mathbb{T}_{s,X^*}^{\perp}$ we have that,

$$v_s v_s^T \in \mathbb{T}_{s,X^*}^{\perp} \text{ and } v_t v_t^T \in \mathbb{T}_{t,X^*}^{\perp}.$$

Define $\hat{\beta}_{st,ij} = E_{s,ij} \bullet (v_s v_s^T) = v_s^T E_{ij}^s v_s$. By definition of $\hat{\beta}_{st,ij}$,

$$v_s v_s^T = \sum_{i \leq j \in \mathsf{C}_{st}} \hat{\beta}_{st,ij} E_{s,ij}. \tag{17}$$

From (12), we also have that $v_t v_t^T = \sum_{i \leq j \in \mathsf{C}_{st}} \hat{\beta}_{st,ij} E_{t,ij}$. Thus, the choice of

$$\alpha_p = 0, \beta_{st,ij} = \begin{cases} \hat{\beta}_{st,ij} \text{ for } s, t \text{ satisfying Assumption 1} \\ \quad 0 \text{ otherwise} \end{cases}$$

satisfies (16) for scalars $\alpha_p, \beta_{st,ij}$ not all 0. Thus, the solution to conversion SDP (5) is primal degenerate $\square$

4.3 Dual Non-uniqueness in Conversion SDP

The solution X_s^* and multipliers $\zeta_{s,p}^*, \xi_{st,ij}^*, S_s^*$ satisfy the first order optimality conditions for the conversion SDP (5),

$$
\begin{aligned}
& A_{s,0} + \sum_{p=1}^{m} \zeta_{s,p}^* A_{s,p} + \sum_{t:(s,t)\in\mathcal{E}} \sum_{i\le j\in \mathsf{C}_{st}} \xi_{st,ij}^* E_{s,ij} \\
& - \sum_{t:(t,s)\in\mathcal{E}} \sum_{i\le j\in \mathsf{C}_{ts}} \xi_{ts,ij}^* E_{s,ij} - S_s^* = 0 \\
& \sum_{s=1}^{\ell} A_{s,p} \bullet X_s^* = b_p \\
& E_{s,ij} \bullet X_s^* = E_{t,ij} \bullet X_t^* \\
& X_s^*, S_s^* \succeq 0, X_s^* S_s^* = 0.
\end{aligned}
\tag{18}
$$

Theorem 2. *Suppose Assumption 1 holds and $v_s^T S_s^* v_s > 0$ or $v_t^T S_t^* v_t > 0$. Then, the optimal multipliers for the conversion SDP* (5) *are not unique.*

Proof. Let $\zeta_{s,p}^*, \xi_{st,ij}^*, S_s^*$ satisfy the first order optimality conditions (18) for the conversion SDP (5). In the following we show by construction the existence of other multipliers satisfying the conditions in (18). Suppose, $v_s^T S_s^* v_s = \gamma > 0$. Since, v_s is a 0-eigenvector of X_s^* (Lemma 4) and $X_s^* S_s^* = 0$ (18) we have that v_s is also an eigenvector of S_s^*. Thus, for all $0 \le \delta \le \gamma$,

$$
\begin{aligned}
X_s^*(S_s^* - \delta v_s v_s^T) = 0, S_s^* - \delta v_s v_s^T \succeq 0 \\
X_t^*(S_t^* + \delta v_t v_t^T) = 0, S_t^* + \delta v_s v_s^T \succeq 0
\end{aligned}
\tag{19}
$$

Following the proof of Theorem 1 we have that there exist $\hat{\beta}_{st,ij}$ such that (17) holds. Hence,

$$
\begin{aligned}
& \sum_{i\le j\in \mathsf{C}_{st}} (\xi_{st,ij}^* - \delta\hat{\beta}_{st,ij}) E_{s,ij} - (S_s^* - \delta v_s v_s^T) \\
= & \sum_{i\le j\in \mathsf{C}_{st}} \xi_{st,ij}^* E_{s,ij} - S_s^*.
\end{aligned}
$$

Further, by Lemma 4 we also have that,

$$
\begin{aligned}
& - \sum_{i\le j\in \mathsf{C}_{st}} (\xi_{st,ij}^* - \delta\hat{\beta}_{st,ij}) E_{t,ij} - (S_t^* + \delta v_t v_t^T) \\
= & - \sum_{i\le j\in \mathsf{C}_{st}} \xi_{st,ij}^* E_{t,ij} - S_t^* \\
& + \left(\sum_{i\le j\in \mathsf{C}_{st}} \delta\hat{\beta}_{st,ij} E_{t,ij} - \delta v_t v_t^T \right) \\
= & - \sum_{i\le j\in \mathsf{C}_{st}} \xi_{st,ij}^* E_{t,ij} - S_t^* + \delta(v_t v_t^T - v_t v_t^T) \\
= & - \sum_{i\le j\in \mathsf{C}_{st}} \xi_{st,ij}^* E_{t,ij} - S_t^*.
\end{aligned}
$$

Thus, for any $0 < \delta \leq \gamma$ replacing $\xi^*_{st,ij}, S^*_s, S^*_t$ with

$$\xi^*_{st,ij} + \delta\hat{\beta}_{st,ij}, S^*_s - \delta v_s v_s^T, S^*_t + \delta v_t v_t^T$$

will also result in satisfaction of the first order optimality conditions in (18). Hence, the multipliers are not unique when $v_s^T S^*_s v_s > 0$. The proof follows in an identical fashion for $v_t^T S^*_t v_t > 0$. This completes the proof. $\square$

5 Numerical Experiments

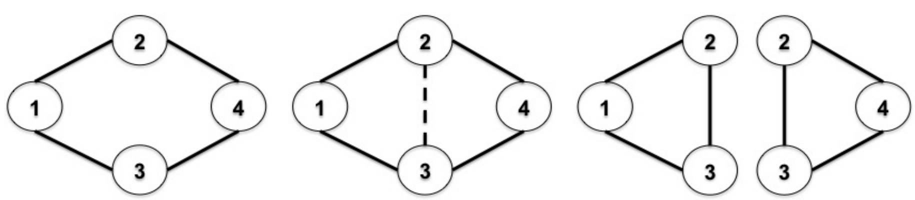
We demonstrate the results of the previous section through numerical experiments on a simple SDP. Consider the SDP with data

$$A_0 = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}, A_p = e_p e_p^T \; \forall \, p = 1, \ldots, 4. \tag{20}$$

This form of the SDP has the same structure as the SDP relaxation for MAXCUT investigated by Goemans and Williamson [5]. The eigenvalues and vectors of A_0 are,

$$\Lambda_0 = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}, Q_0 = \begin{bmatrix} -\frac{1}{2} & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{\sqrt{2}} & 0 & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{2} \\ -\frac{1}{2} & 0 & -\frac{1}{\sqrt{2}} & -\frac{1}{2} \end{bmatrix}.$$

Since A_0 has the smallest eigenvalue to be -1, the optimal solution to the SDP defined by (20) is $X^* = 4q_1 q_1^T$ where q_1 is the first column of Q_0 (the eigenvector of A_0 corresponding to eigenvalue of -1). The factor 4 ensures that the equality constraints are satisfied.



(a) $\mathsf{G}(\mathsf{N},\mathsf{E})$ (b) $\mathsf{G}(\mathsf{N},\mathsf{F})$ (c) $\mathsf{C}_1 = \{2,3,1\}, \mathsf{C}_2 = \{2,3,4\}$

Figure 1: (a) Graph of the original SDP. (b) Graph of the chordal completion. (c) Maximal clique decomposition of chordal completion.

For the data in (20), the graph of the aggregate sparsity sparsity pattern is depicted in Figure 1(a). The $\mathsf{G}(\mathsf{N},\mathsf{E})$ is a 4-cycle and not chordal. Figure 1(b) shows a chordal extension where an edge $(2,3)$ has been introduced. The maximal clique decomposition for the chordal graph $\mathsf{G}(\mathsf{N},\mathsf{F})$ is shown in Figure 1(c). Note that we have ordered the vertices such that (12) for ease of presentation. The conversion SDP is given by the data,

$$\begin{aligned}
&\mathsf{C}_1 = \{2,3,1\}, \mathsf{C}_2 = \{2,3,4\} \\
&A_{1,0} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, A_{2,0} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \\
&A_{1,1} = e_3 e_3^T, A_{2,1} = 0; A_{1,2} = e_1 e_1^T, A_{2,2} = 0 \\
&A_{1,3} = e_3 e_3^T, A_{2,3} = 0; A_{1,4} = 0, A_{2,4} = e_4 e_4^T \\
&E_{1,22} = E_{2,22} = e_1 e_1^T, E_{1,23} = E_{2,23} = \frac{1}{2}(e_1 e_2^T + e_2 e_1^T) \\
&E_{1,33} = E_{2,33} = e_2 e_2^T.
\end{aligned}$$

The solution to the conversion SDP is,

$$X_1^* = X_2^* = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}$$

Clearly, $\mathrm{rank}(X_1^*) = \mathrm{rank}(X_2^*) = 1 < |\mathsf{C}_{12}|$. Hence, Assumption 1 holds. The eigenvectors, eigenvalues of S_1^* are,

$$\Lambda_1 = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, Q_1 = \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \end{bmatrix}.$$

5.1 Primal Degeneracy

As shown in Lemma 4 we have that $u = [\frac{1}{\sqrt{2}}\ \frac{1}{\sqrt{2}}]^T$ is a 0-eigenvector of the submatrix which corresponds to the intersection of the cliques, C_{12}. As shown in Lemma 4, $v_1 = v_2 = [u^T\ 0]^T$ are 0-eigenvectors of X_1^*, X_2^* respectively.

From the definition of $\mathbb{T}_{s,X^*}$ it is easy to see that,

$$\begin{aligned}
v_1 v_1^T &= \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix} = Q_1 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} Q_1^T \in \mathbb{T}_{1,X^*} \\
v_1 v_1^T &= \frac{1}{2} E_{1,22} - E_{1,23} + \frac{1}{2} E_{1,33}.
\end{aligned}$$

Similarly, it can be shown that

$$\begin{aligned}
-v_2 v_2^T &= Q_1 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} Q_1^T \in \mathbb{T}_{1,X^*} \\
-v_2 v_2^T &= \frac{1}{2}(-E_{2,22}) - (-E_{2,23}) + \frac{1}{2}(-E_{2,33}).
\end{aligned}$$

Thus, there exists an element in $\mathbb{T}_{1,X^*}$ and $\mathbb{T}_{2,X^*}$ that is in the span of the constraints equating the elements in C_{12}. Hence, the conversion SDP is primal degenerate.

5.2 Non-unique Multipliers

For the original SDP, the optimal multipliers are,

$$\zeta^* = \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \end{bmatrix}, S^* = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \\ 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 2 \end{bmatrix}.$$

For the conversion SDP, multipliers satisfying (18) are

$$\zeta_1^* = \begin{bmatrix} -1 \\ -1 \\ -1 \\ 0 \end{bmatrix}, \zeta_2^* = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \end{bmatrix}, \xi_{12,22}^* = 1, \xi_{12,23}^* = 0, \xi_{12,33}^* = 1$$

$$S_1^* = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}, S_2^* = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}.$$

The eigenvectors of S_1^*, S_2^* are Q_1 while the eigenvalues are,

$$\Lambda_{S1} = \Lambda_{S2} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Thus it is easy to see that $X_1^* S_1^* = 0$ and they satisfy strict complementarity. The same is also true of X_2^* and S_2^*. The eigenvalue of v_1 is 1 and satisfies the conditions in Theorem 2 and hence, for all $0 \le \delta \le 1$,

$$\begin{aligned} & S_1^* - \delta v_1 v_1^T \succeq 0, X_1^*(S_1^* - \delta v_1 v_1^T) = 0 \\ & (\xi_{12,22}^* - \frac{1}{2}\delta) E_{12,22} + (\xi_{12,23}^* + \delta) E_{12,23} \\ & + (\xi_{12,33}^* - \frac{1}{2}\delta) E_{12,33} - (S_1^* - \delta v_1 v_1^T) \\ = {} & \xi_{12,22}^* E_{12,22} + \xi_{12,23}^* E_{12,23} + \xi_{12,33}^* E_{12,33} - S_1^* \end{aligned}$$

Further, it can also be shown that,

$$\begin{aligned} & (\xi_{12,22}^* - \frac{1}{2}\delta)(-E_{12,22}) + (\xi_{12,23}^* + \delta)(-E_{12,23}) \\ & + (\xi_{12,33}^* - \frac{1}{2}\delta)(-E_{12,33}) - (S_2^* + \delta v_2 v_2^T) \\ = & - \xi_{12,22}^* E_{12,22} - \xi_{12,23}^* E_{12,23} - \xi_{12,33}^* E_{12,33} - S_2^* \\ & S_2^* + \delta v_2 v_2^T \succeq 0, X_2^*(S_2^* + \delta v_2 v_2^T) = 0. \end{aligned}$$

Thus, we have that the multipliers

$$\begin{aligned} & \zeta_1^*,\ \zeta_2^*,\ \xi_{12,22}^* - \frac{1}{2}\delta,\ \xi_{12,23}^* + \delta,\ \xi_{12,33}^* - \frac{1}{2}\delta \\ & S_1^* + \delta v_1 v_1^T,\ S_2^* + \delta v_2 v_2^T \end{aligned}$$

also satisfy the first order optimality conditions for conversion SDP. This shows that there are an infinite set of multipliers for the conversion SDP.

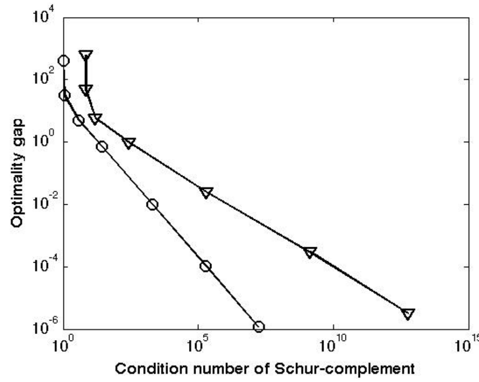


Figure 2: Plot of the condition number of the schur-complement matrix in the IPM against the optimality gap. ∘ - original SDP formulation, △ - conversion SDP.

5.3 Ill-conditioning in IPM

Since the multipliers are not unique, the matrix used in the step computation of the IPM for SDP must be singular in the limit. Figure 2 plots the condition number of the schur-complement matrix in SDPT3 [11] against the optimality gap. SDPT3 takes 7 iterations to solve either formulation. But the plot clearly shows that the condition number of the schur-complement matrix is higher for the conversion SDP. This is attributable to the non-uniqueness of the dual multipliers.

6 Conclusions & Future Work

We analyzed the conversion approach for SDP proposed by Fukuda et al [4]. The analysis showed that for SDPs with a low rank solution, the conversion SDP was primal degenerate. We also provided conditions under which the multipliers for the conversion SDP were non-unique. The theory was exemplified using a simple 4×4 SDP. In the example, the ill-conditioning in the schur-complement matrix was greater for the conversion SDP. Nevertheless, this did not affect the number of iterations to reach the said tolerance. We believe the effect of the ill-conditioning is likely to be more dramatic for larger problems and affect convergence of IPM. This will be investigated in a future study.

References

[1] F. Alizadeh, J-P. A. Haeberly, and M. L. Overton. Complementarity and nondegeneracy in semidefinite programming. *Mathematical Programming*, 77(2):111–128, 1997.

[2] J. R. S. Blair and B. Peyton. An introduction to chordal graphs and clique trees. In A. George, J. R. Gilbert, and J. W. H. Liu, editors, *Graph Theory and Sparse Matrix Computation*, volume 56 of *The IMA Volumes in Mathematics and its Applications*, pages 1–29. Springer New York, 1993.

[3] E. de Klerk. *Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications.* Kluwer Academic Publishers, 2002.

[4] M. Fukuda, M. Kojima, K. Murota, and K. Nakata. Exploiting Sparsity in Semidefinite Programming via Matrix Completion I: General Framework. *SIAM J. Optimization*, 11(3):647–674, 2000.

[5] M. X. Goemans and D. P. Williamson. .879-approximation algorithms for MAX CUT and MAX 2SAT. In *Proceedings of the twenty-sixth annual ACM symposium on Theory of computing*, STOC '94, pages 422–431, New York, NY, USA, 1994.

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

[7] S. Kim and M. Kojima. Exploiting sparsity in sdp relaxation of polynomial optimization problems. In Miguel F. Anjos and Jean B. Lasserre, editors, *Handbook on Semidefinite, Conic and Polynomial Optimization*, volume 166 of *International Series in Operations Research & Management Science*, pages 499–531. Springer US, 2012.

[8] K. Nakata, K. Fujisawa, M. Fukuda, M. Kojima, and K. Murota. Exploiting Sparsity in Semidefinite Programming Via Matrix Completion II: Implementation and numerical results. *Mathematical Programming (Series B)*, 95(2):303–327, 2003.

[9] Y. Nesterov and A. Nemirovskii. *Interior Point Polynomial Algorithms in Convex Programming*. Studies in Applied Mathematics. Society for Industrial and Applied Mathematics, 1987.

[10] M. J. Todd, K. C. Toh, and R. H. Tütüncü. On the Nesterov-Todd Direction in Semidefinite Programming. *SIAM J. on Optimization*, 8(3):769–796, 1998.

[11] K. C. Toh, M.J. Todd, and R. H. Tütüncü. SDPT3 – a MATLAB software package for semidefinite programming. *Optimization Methods and Software*, 11:545–581, 1999.

[12] L. Vandenberghe and S. Boyd. Semidefinite Programming. *SIAM Review*, 38:49–95, March 1996.

[13] M. Yamashita, K. Fujisawa, and M. Kojima. Implementation and evaluation of SDPA 6.0 (SemiDefinite Programming Algorithm 6.0). *Optimization Methods and Software*, 18:491–505, 2003.

[14] M. Yamashita, K. Fujisawa, K. Nakata, M. Nakata, M. Fukuda, K. Kobayashi, and K. Goto. A high-performance software package for semidefinite programs: SDPA 7. Technical Report Research Report B-460, Dept. of Mathematical and Computing Science, Tokyo Institute of Technology, Tokyo, Japan, September 2010.