ISSN 1342-2804

Research Reports on Mathematical and Computing Sciences

Correlative sparsity in primal-dual interior-point methods for LP, SDP and SOCP

Kazuhiro Kobayashi, Sunyoung Kim and Masakazu Kojima

September 2006, B–434

Department of Mathematical and Computing Sciences
Tokyo Institute of Technology

SERIES B: Operations Research

B-434

Correlative sparsity in primal-dual interior-point methods for LP, SDP and SOCP

Kazuhiro Kobayashi[♯], Sunyoung Kim[⋆] and Masakazu Kojima[†]

September 2006

Abstract.
Exploiting sparsity has been a key issue in solving large-scale optimization problems. The most time-consuming part of primal-dual interior-point methods for linear programs, second-order cone programs, and semidefinite programs is solving the Schur complement equation at each iteration, usually by the Cholesky factorization. The computational efficiency is greatly affected by the sparsity of the coefficient matrix of the equation that is determined by the sparsity of an optimization problem (linear program, semidefinite program or second-order program). We show if an optimization problem is *correlatively sparse,* then the coefficient matrix of the Schur complement equation inherits the sparsity, and a sparse Cholesky factorization applied to the matrix results in no fill-in.

Key words.
Correlative sparsity, primal-dual interior-point method, linear program, semidefinite program, second-order program, partial separability, chordal graph

♯ Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, 2-12-1 Oh-Okayama, Meguro-ku, Tokyo 152-8552 Japan.
`kazuhir2@is.titech.ac.jp`

⋆ Department of Mathematics, Ewha Women's University, 11-1 Dahyun-dong, Sudaemoon-gu, Seoul 120-750 Korea. The research was supported by Kosef R01-2005-000-10271-0 and KRF-2006-312-C00062.
`skim@ewha.ac.kr`

† Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, 2-12-1 Oh-Okayama, Meguro-ku, Tokyo 152-8552 Japan.
`kojima@is.titech.ac.jp`

1 Introduction

Primal-dual interior-point methods are shown to be numerically robust for solving linear programs (LPs), semidefinite programs (SDPs) and second-order cone programs (SOCPs). For large-scale LPs, many free and commercial software packages implementing primal-dual interior-point methods have been established as very efficient and powerful solvers. Challenges still remain in solving large-scale SDPs and SOCPs, although several successful software packages have been developed.

Efficient handling of large-scale LPs, SDPs and SOCPs in implementation of primal-dual interior-point methods have taken two significant approaches: solving the Schur complement equation (or larger systems that induce the Schur complement equation [26, 28]) efficiently and exploiting the sparsity of the problems. In the case of LP, these two issues have been studied widely and various practical techniques have been implemented in free and commercial software packages. For LPs and SOCPs, the sparsity of the Schur complement matrix (the coefficient matrix of the Schur complement equation) was exploited by splitting the matrix into sparse and dense parts, factorizing the sparse part, and applying low-rank update to the dense part [2, 9, 24]. For the case of SDP, the importance of exploiting the sparsity of the data matrices was recognized in [6], which proposed three types of methods for computing the elements of the Schur complement matrix depending on their sparsity. For large-scale SDPs, solving the Schur complement equation using iterative methods was proposed in [21, 25, 26]. The aggregated sparsity pattern of all data matrices of SDPs was exploited for the primal-dual interior-point methods [7, 22]. The current paper adopts some of basic ideas such as a chordal graph structure of the sparsity pattern used there. The aggregated sparsity, however, does not necessarily imply the sparsity of the Schur complement matrix. The issue of an efficient solution to the Schur complement equation was not addressed there. Instead, the focus was on an efficient handling of the primal matrix variable that becomes dense in general even when the aggregated sparsity pattern is sparse.

Sparsity can be used in various ways depending on optimization problems. The sparsity from partially separable functions was used in connection with efficient implementation of quasi-Newton methods for solving large-scale unconstrained optimization [10]. The correlative sparsity was introduced to handle the sparsity of polynomial optimization problems (POPs) [30], as a special case of sparsity described in [13, 15]. The relationship between the partial separability and the correlative sparsity was discussed in the recent paper [14]. In the primal-dual interior-point methods, exploiting the sparsity of the Schur complement matrix becomes important for efficiency because the Cholesky factorization is commonly used for the solution of the Schur complement equation. We note that many fill-ins may occur after applying a sparse Cholesky factorization to a general non-structured sparse matrix. The sparsity of the Schur complement matrix is determined by the sparsity of the data matrices of an optimization problem (LP, SDP or SOCP). Our motivation is based on finding a sparsity condition on SDPs and SOCPs that can lead to a sparse Cholesky factorization of the Schur complement matrix with no fill-in. We show that the correlative sparsity is indeed such a sparsity condition that provides a sparse Cholesky factorization of the Schur complement matrix with no fill-in. We also propose a correlatively-sparse linear optimization problem (LOP) for a unified representation of correlatively-sparse LPs, SDPs and SOCPs.

We introduce the correlative sparsity, which was originally proposed for a POP [30], for an LOP with a linear objective function in an n-dimensional variable vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ and inequality (linear matrix inequality, second-order cone inequality) constraints in $\boldsymbol{x}$. This LOP is so called the dual inequality standard form LOP unifying LPs, SDPs and SOCPs. The correlative sparsity of the LOP is defined by an $n \times n$ symmetric matrix $\boldsymbol{R}$, called the correlative sparsity pattern (csp) matrix, as follows. Each element R_{ij} of the csp matrix $\boldsymbol{R}$ is either 0 or $\star$ for a nonzero value. The symbol $\star$ was assigned to all diagonal elements of $\boldsymbol{R}$ and also to each off-diagonal element $R_{ij} = R_{ji}$ $(1 \leq i < j \leq n)$ if and only if the variables x_i and x_j appear simultaneously in a linear inequality (linear matrix inequality, second-order cone inequality) constraint. If the csp matrix $\boldsymbol{R}$ allows a symbolic sparse Cholesky factorization with no fill-in (under an appropriate simultaneous reordering of its rows and columns), we say that the LOP is correlatively-sparse. The objective of the paper is to show that if the LOP satisfies the correlative sparsity, then the sparsity pattern of the Schur complement matrix coincides with the csp matrix $\boldsymbol{R}$. This guarantees a sparse Cholesky factorization of the Schur complement matrix with no fill-in.

Although our major concern is a correlatively-sparse LOP, we deal with an almost-correlatively-sparse LOP, a slightly more practical LOP with a small-sized dense linear inequality (linear matrix inequality, second-order cone inequality) constraint and sparse constraints inducing a csp matrix $\boldsymbol{R}$. With this form of LOP, the correlatively-sparse LOP and the almost-correlatively-sparse LOP can be dealt with simultaneously because all the results for the correlatively-sparse LOP can be obtained by simply neglecting the dense constraint. The Schur complement matrix of the almost-correlatively-sparse LOP is dense in general. Its sparsity, however, can be exploited by splitting the matrix into two parts, the sparse part with the same sparsity pattern as the csp matrix $\boldsymbol{R}$ and the dense part of low-rank. A sparse Cholesky factorization can be used for the sparse part, and the well-known Sherman-Morrison-Woodbury formula for the dense part. This technique of splitting sparse and dense parts of the Schur complement matrix was used in an affine scaling variant of Karmarkar's algorithm for linear programs [1]. See also [2, 9, 24].

We also examine the link between the correlative sparsity of a POP and the sparsity of its SDP relaxation. When the sparse SDP relaxation ([30], see also [16, 18]) is applied to a POP satisfying the correlative sparsity, an equivalent polynomial SDP satisfying the same correlative sparsity is constructed as an intermediate optimization problem. It is then linearized to an SDP relaxation problem. We prove that the correlative sparsity of the POP is further maintained in the SDP relaxation problem. It was also observed through the numerical experiment in [30] that "the sparse SDP relaxation for a correlatively-sparse POP leads to an SDP that can maintain the sparsity for primal-dual interior-point methods". This observation is supported by a theoretical proof given in Section 6.

This paper is organized as follows. After describing a correlatively-sparse LOP for a unified treatment of LPs, SOCPs and SDPs, the definition of correlative sparsity for the LOP is presented in Section 2. An almost-correlatively-sparse LOP and its simple illustrative example are also given in Section 2. In Sections 3, 4 and 5, we present almost-correlatively-sparse LP, SDP and SOCP, respectively. In Section 6, we show that the correlative sparsity of a polynomial SDP is preserved in the SDP obtained as its linearization. Section 7 is devoted to concluding remarks.

2 Preliminaries

2.1 Correlatively-sparse and almost-correlatively-sparse linear optimization problems

Let $M = \{1, 2, \ldots, m\}$ and $N = \{1, 2, \ldots, n\}$. Consider an optimization problem with a linear objective function

$$\text{maximize} \quad \boldsymbol{b}^T \boldsymbol{y} \quad \text{subject to} \quad (y_i : i \in I_p) \in C_p \ (p \in M). \tag{1}$$

Here I_p denotes a nonempty subset of N, $(y_i : i \in I_p)$ a subvector of an n-dimensional vector $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ consisting of elements y_i $(i \in I_p)$, and C_p a nonempty subset of $\#I_p$-dimensional Euclidean space. With the form of (1), we can deal with the problems whose C_p represents linear inequalities, linear matrix inequalities or second-order cone inequalities. Thus, (1) represents a unified linear optimization problem including LPs, SDPs and SOCPs.

Notice that some I_p can coincide with or contained in another I_q. Let $\{I_p : p \in M_{\max}\}$ denote the family of maximal I_p's among the entire family $\{I_p : p \in M\}$; $\{I_p : p \in M_{\max}\}$ is uniquely determined although $M_{\max}$ does not have to be unique. This family plays an essential role in defining the correlative sparsity for the optimization problem (1) in the following discussion. For simplicity of notation, we assume that $M_{\max} = \{1, 2, \ldots, \ell\}$.

Define the $n \times n$ symmetric symbolic matrix $\boldsymbol{R}$ whose (i, j)th element is given by

$$R_{ij} = \begin{cases} \star & \text{if } i = j \ \text{ or } i, \ j \in I_p \text{ for some } p \in M_{\max}, \\ 0 & \text{otherwise.} \end{cases}$$

We call $\boldsymbol{R}$ *the correlative sparsity pattern (csp) matrix* of the optimization problem (1). It should be noted that we can replace $M_{\max}$ by M to define the csp matrix $\boldsymbol{R}$. According to [30], we say that the optimization problem (1) is *correlatively-sparse* if the csp matrix $\boldsymbol{R}$ allows a sparse Cholesky factorization (under an appropriate simultaneous reordering of row and column indices, which is called a perfect elimination ordering) with no fill-in. Here we implicitly assume that $\boldsymbol{R}$ is sufficiently sparse.

The csp matrix induces a graph, which is called *the csp graph* of the optimization problem (1). We refer to [4, 8, 19] or some text books on the graph theory for the background materials stated below. Let $N = \{1, 2, \ldots, n\}$ denote the node set of the csp graph, and define the edge set $E = \{\{i, j\} \subset N : i < j \text{ and } R_{ij} = \star\}$. Then it is known that there exists a Cholesky factorization of the csp matrix $\boldsymbol{R}$ with no fill-in if and only if the csp graph $G(N, E)$ is chordal. Here a graph is said to be chordal if every cycle of length ≥ 4 has a chord (an edge jointing two non-consecutive vertices of the cycle). By construction, $\{I_p : p \in M_{\max}\}$ forms the family of maximal cliques of the csp graph $G(N, E)$. The chordal graph property is also characterized as the running intersection property: there exists a permutation $\pi(1), \pi(2), \ldots, \pi(\ell)$ of $1, 2, \ldots, \ell$ such that

$$\forall p \in \{1, 2, \ldots, \ell - 1\} \ \exists r; \pi(r) \geq \pi(p+1) \text{ and } \ I_{\pi(p)} \cap \left(\cup_{q=p+1}^{\ell} I_{\pi(q)}\right) \subset I_{\pi(r)}. \tag{2}$$

(Recall that we have assumed $M_{\max} = \{1, 2, \ldots, \ell\} \subseteq M = \{1, 2, \ldots, m\}$).

If the running intersection property (2) is satisfied, then there exists a perfect elimination ordering of $1, 2, \ldots, n$ under which the matrix $\boldsymbol{R}$ is factorized into a lower triangular matrix

$\boldsymbol{L}$ and its transpose with no fill-in; $\boldsymbol{P}\boldsymbol{R}\boldsymbol{P}^T = \boldsymbol{L}\boldsymbol{L}^T$ where $\boldsymbol{P}$ stands for the permutation matrix corresponding to the perfect elimination ordering. The running intersection property implies $\ell \leq n$.

If the optimization problem (1) is not correlatively-sparse or equivalently if the csp graph $G(N, E)$ is not a chordal graph, we can extend $G(N, E)$ to a chordal graph $G(N, E')$, where E' denotes a superset of E. Let I'_p ($p \in M'_{\max}$) denote the maximal cliques of the chordal extension $G(N, E')$ of $G(N, E)$, which induces, as its adjacency matrix, an extended csp matrix $\boldsymbol{R}'$ whose (i, j)th element is given by

$$R'_{i,j} = \begin{cases} \star & \text{if } i = j \text{ or } i,\ j \in I'_p \text{ for some } p \in M'_{\max}, \\ 0 & \text{otherwise.} \end{cases}$$

Now the extended csp matrix $\boldsymbol{R}'$ allows a sparse Cholesky factorization under a perfect elimination ordering of its rows and columns. By construction, for every $q \in M_{\max}$, there exists a $p \in M'_{\max}$ such that $I_q \subseteq I'_p$; hence we may replace the constraint $(y_i : i \in I_q) \in C_q$ by

$$(y_i : i \in I'_p) \in C'_q = \{(y_i : i \in I'_p) : (y_i : i \in I_q) \in C_q\}.$$

Then the resulting optimization problem satisfies the correlative sparsity with the csp graph $G(N, E')$ and csp matrix $\boldsymbol{R}'$. This implies that all the results in the subsequent sections remain valid if we replace the csp graph $G(N, E)$ and the csp matrix $\boldsymbol{R}$ by their chordal extensions $G(N, E')$ and $\boldsymbol{R}'$, respectively. (Here we assume that $\boldsymbol{R}'$ remains sparse). Chordal extensions of $G(N, E)$ are not unique, and a minimum chordal extension (a chordal extension with the least number of edges) of the csp graph $G(N, E)$ creates the minimum number of nonzero elements in the resulting extended csp matrix $\boldsymbol{R}'$ from $\boldsymbol{R}$. Finding such a minimum chordal extension is difficult in general. Several heuristics such as the minimum degree ordering and the reverse Cuthill-Mckee ordering are known to produce good approximations. Also some software packages [3, 12] are available for a chordal extension of a given graph.

Now we mention an almost-correlatively-sparse optimization problem by adding a dense constraint to the sparse optimization problem (1).

$$\left.\begin{array}{ll} \text{minimize} & \boldsymbol{b}^T\boldsymbol{y} \\ \text{subject to} & (y_i : i \in I_p) \in C_p\ (p \in M),\ \ \boldsymbol{y} \in C_{m+1}. \end{array}\right\} \tag{3}$$

Here we assume that the two optimization problems (1) and (3) share the correlatively-sparse constraints $(y_i : i \in I_p) \in C_p$ ($p \in M$) and the csp matrix $\boldsymbol{R}$ induced from these constraints. The added dense constraint $\boldsymbol{y} \in C_{m+1}$ is also assumed to be small in the sense that it represents a few number of linear inequalities, a small-sized linear matrix inequality and/or a few number of small-sized second order cone inequalities. When primal-dual interior-point methods are applied to an almost-correlatively-sparse LP, SOCP or SDP, the Schur complement matrix becomes dense in general. But it is decomposed into two matrices, a sparse matrix with the same sparsity pattern as the csp matrix $\boldsymbol{R}$, and a dense and low-rank matrix. Then, a sparse factorization and the Scherman-Morrison-Woodbury formula can be applied to solve the Schur complement equation. We discuss more details on almost-correlatively-sparse LP, SDP and SOCP in the subsequent sections.

2.2 An illustrative example

As an example, we consider the optimization problem

$$\left.\begin{array}{ll}\text{maximize} & y_1 + 2y_2 + 3y_3 + 4y_4 \\ \text{subject to} & 5y_1 + 6y_4 \leq 17, \\ & 3y_2 + 2y_4 \leq 16, \\ & \begin{pmatrix} 11 & 2 \\ 2 & 12 \end{pmatrix} - \begin{pmatrix} 4 & -3 \\ -3 & 3 \end{pmatrix} y_3 - \begin{pmatrix} 2 & 2 \\ 2 & 3 \end{pmatrix} y_4 \succeq \boldsymbol{O}, \\ & y_1 + y_2 + y_3 + y_4 \geq 1 \\ & y_i \geq 0 \ (i = 1, 2, 3, 4). \end{array}\right\} \tag{4}$$

Note that the 3rd constraint is an LMI where $\boldsymbol{A} \succeq \boldsymbol{O}$ implies that a symmetric matrix $\boldsymbol{A}$ is positive semidefinite. Let

$$\left.\begin{array}{lcl} \ell & = & m = 3,\ n = 4,\ M_{\max} = M = \{1, 2, 3\}, \quad N = \{1, 2, 3, 4\}, \\ I_1 & = & \{1, 4\},\ C_1 = \{(y_1, y_4) : y_1 \geq 0,\ 17 - 5y_1 - 6y_4 \geq 0\}, \\ I_2 & = & \{2, 4\},\ C_2 = \{(y_2, y_4) : y_2 \geq 0,\ 16 - 3y_2 - 2y_4 \geq 0\},\ I_3 = \{3, 4\}, \\ C_3 & = & \left\{ (y_3, y_4) : \begin{array}{c} y_3 \geq 0, y_4 \geq 0, \\ \begin{pmatrix} 11 & 2 \\ 2 & 12 \end{pmatrix} - \begin{pmatrix} 4 & -3 \\ -3 & 3 \end{pmatrix} y_3 - \begin{pmatrix} 2 & 2 \\ 2 & 3 \end{pmatrix} y_4 \succeq \boldsymbol{O} \end{array} \right\}, \\ C_4 & = & \{\boldsymbol{y} = (y_1, y_2, y_3, y_4) : -1 + y_1 + y_2 + y_3 + y_4 \geq 0\}. \end{array}\right\} \tag{5}$$

Then we can rewrite the problem (4) as a special case of LOP (3). In this case, the 4×4 csp matrix turns out to be

$$\boldsymbol{R} = \begin{pmatrix} \star & 0 & 0 & \star \\ 0 & \star & 0 & \star \\ 0 & 0 & \star & \star \\ \star & \star & \star & \star \end{pmatrix}. \tag{6}$$

This csp matrix $\boldsymbol{R}$ exhibits a typical sparsity pattern that allows a sparse Cholesky factorization wit no fill-in. The dense constraint set C_4 is described as a single number of linear inequality. We will see at the end of Section 4 that the Schur complement matrix is decomposed into a 4×4 matrix of the same sparsity pattern as the csp matrix $\boldsymbol{R}$ and a 4×4 dense and rank-1 matrix.

3 LP

While primal-dual interior-point methods for LPs have been studied for many years, the issue of exploiting the sparsity in the Schur complement equation is considered important for computational efficiency in practice. Many software packages implementing primal-dual interior-point methods for solving LPs must have utilized the correlative sparsity implicitly, although the term itself has not appeared explicitly. A primal-dual interior-point method for an LP is usually described with the equality standard form in most of the literature, however, it is its dual (or the inequality standard form) that explicitly relates the sparsity of the LP to the sparsity of the Schur complement matrix. We thus adopt an inequality standard form to describe an almost-correlative-sparse LP. This enables us to link the csp matrix $\boldsymbol{R}$ induced from the correlatively-sparse constraints (or the constraints shared with the corresponding correlatively-sparse LP without the dense constraint) and the coefficient matrix of the dense constraint to the sparsity pattern of the Schur complement matrix. If we neglect the dense constraint in the almost-correlatively-sparse LP, we see that the sparsity pattern of the Schur complement matrix coincides with the sparsity pattern of the csp matrix $\boldsymbol{R}$ in the correlatively-sparse case.

Suppose that each C_p in the LOP (3) is described as a system of linear inequalities such that

$$\begin{aligned} C_p &= \{(y_i : i \in I_p) : \boldsymbol{a}_{p0} - \sum_{i \in I_p} \boldsymbol{a}_{pi} y_i \geq 0\} \ (p \in M), \\ C_{m+1} &= \{\boldsymbol{y} = (y_1, y_2, \ldots, y_n) : \boldsymbol{a}_{(m+1)0} - \sum_{i \in N} \boldsymbol{a}_{(m+1)i} y_i \geq 0\}. \end{aligned}$$

Here $\boldsymbol{a}_{pi}$ $(i = 0, 1, 2, \ldots, n,\ p \in M \cup \{m+1\})$ denote column vectors. Then we have an LP of the form

$$\begin{array}{ll} \text{maximize} & \boldsymbol{b}^T \boldsymbol{y} \\ \text{subject to} & \boldsymbol{a}_{p0} - \sum_{i \in I_p} \boldsymbol{a}_{pi} y_i \geq 0 \ (p \in M), \\ & \boldsymbol{a}_{(m+1)0} - \sum_{i \in N} \boldsymbol{a}_{(m+1)i} y_i \geq 0, \end{array}$$

and its dual

$$\begin{array}{ll} \text{minimize} & \sum_{p \in M} \boldsymbol{a}_{p0}^T \boldsymbol{x}_p \\ \text{subject to} & \sum_{p \in K_i} \boldsymbol{a}_{pi}^T \boldsymbol{x}_p + \boldsymbol{a}_{(m+1)i}^T \boldsymbol{x}_{m+1} = b_i \ (i \in N), \\ & \boldsymbol{x}_p \geq \boldsymbol{0} \ (p \in M \cup \{m+1\}). \end{array}$$

Here

$$K_i = \{p \in M : i \in I_p\} \ (i \in N). \tag{7}$$

When the primal-dual interior-point method is applied to this primal-dual pair of LPs, the (i, j)th element of the $n \times n$ Schur complement matrix $\boldsymbol{B}$ is given by

$$\begin{aligned} B_{ij} &= B_{ij}^{\text{sparse}} + B_{ij}^{\text{dense}}, \\ B_{ij}^{\text{sparse}} &= \sum_{p \in K_i \cap K_j} \boldsymbol{a}_{pi}^T \text{diag}(\boldsymbol{x}_p) \text{diag}(\boldsymbol{s}_p)^{-1} \boldsymbol{a}_{pj}, \\ B_{ij}^{\text{dense}} &= \boldsymbol{a}_{(m+1)i}^T \text{diag}(\boldsymbol{x}_{m+1}) \text{diag}(\boldsymbol{s}_{m+1})^{-1} \boldsymbol{a}_{(m+1)j} \end{aligned}$$

for some $\boldsymbol{s}_p > \boldsymbol{0}$, which corresponds to the slack variable vector for the inequality constraint $\boldsymbol{a}_{p0} - \sum_{i \in I_p} \boldsymbol{a}_{pi} y_i \geq 0$, $(p \in M \cup \{m+1\})$ and some $\boldsymbol{x}_p > \boldsymbol{0}$ $(p \in M \cup \{m+1\})$. Here $\text{diag}(\boldsymbol{w})$ denotes a diagonal matrix with the diagonal elements $w_1, w_2, \ldots, w_\ell$ for every vector $\boldsymbol{w} = (w_1, w_2, \ldots, w_\ell)$. Hence, the Schur complement matrix $\boldsymbol{B}$ is represented as the sum of the two matrices $\boldsymbol{B}^{\text{sparse}}$ and $\boldsymbol{B}^{\text{dense}}$. If we assume that all coefficient vectors $\boldsymbol{a}_{pi}$ $(i \in I_p,\ p \in M \cup \{m+1\})$ are fully dense, then $B^{\text{sparse}}_{ij} = 0$ if and only if $R_{ij} = 0$. Hence the sparsity pattern of $\boldsymbol{B}^{\text{sparse}}$ coincides with the sparsity pattern of $\boldsymbol{R}$, and $\boldsymbol{B}^{\text{dense}}$ becomes fully dense and its rank is not greater than the dimension of the coefficient vector $\boldsymbol{a}_{(m+1)i}$. If $\dim(\boldsymbol{a}_{(m+1)i})$ is small, a low-rank update can be used for the dense part. In general, both $\boldsymbol{B}^{\text{sparse}}$ and $\boldsymbol{B}^{\text{dense}}$ may have additional zeros.

4 SDP

We now consider SDPs. Suppose that each C_p in the LOP (3) is described as an linear matrix inequality (LMI) such that

$$\begin{aligned}
C_p &= \{(y_i : i \in I_p) : \boldsymbol{A}_{p0} - \sum_{i \in I_p} \boldsymbol{A}_{pi} y_i \succeq \boldsymbol{O}\}\ (p \in M), \\
C_{m+1} &= \{\boldsymbol{y} : \boldsymbol{A}_{(m+1)0} - \sum_{i \in N} \boldsymbol{A}_{(m+1)i} y_i \succeq \boldsymbol{O}\}.
\end{aligned}$$

Here $\boldsymbol{A}_{pi}$ $(i = 0, 1, 2, \ldots, n,\ p \in M \cup \{m+1\})$ denote symmetric matrices. Then we have an SDP

$$\left.\begin{array}{ll}
\text{maximize} & \boldsymbol{b}^T \boldsymbol{y} \\
\text{subject to} & \boldsymbol{A}_{p0} - \sum_{i \in I_p} \boldsymbol{A}_{pi} y_i \succeq \boldsymbol{O}\ (p \in M), \\
& \boldsymbol{A}_{(m+1)0} - \sum_{i \in N} \boldsymbol{A}_{(m+1)i} y_i \succeq \boldsymbol{O},
\end{array}\right\} \tag{8}$$

and its dual

$$\left.\begin{array}{ll}
\text{minimize} & \sum_{p \in M} \boldsymbol{A}_{p0} \bullet \boldsymbol{X}_p + \boldsymbol{A}_{(m+1)0} \bullet \boldsymbol{X}_{m+1} \\
\text{subject to} & \sum_{p \in K_i} \boldsymbol{A}_{pi} \bullet \boldsymbol{X}_p + \boldsymbol{A}_{(m+1)i} \bullet \boldsymbol{X}_{m+1} = b_i\ (i \in N), \\
& \boldsymbol{X}_p \succeq \boldsymbol{O}\ (p \in M \cup \{m+1\}),
\end{array}\right\}$$

where K_i $(i \in N)$ are given by (7). If the dense constraint C_{m+1} involves multiple LMIs

$$\boldsymbol{A}_{q0} - \sum_{i \in N} \boldsymbol{A}_{qi} y_i \succeq \boldsymbol{O}\ (q \in L),$$

where $L = \{m+1, m+2, \ldots, m^*\}$, then those LMIs can be put into an LMI by combining the data matrices and redefining larger block diagonal matrices $\boldsymbol{A}_{(m+1)i}$ $(i = 0, 1, 2, \ldots, n)$ such that

$$\boldsymbol{A}_{(m+1)i} \;=\; \operatorname{diag}\left(\boldsymbol{A}_{(m+1)i}, \boldsymbol{A}_{(m+2)i}, \ldots, \boldsymbol{A}_{(m^*)i}\right)\ (i = 0, 1, 2, \ldots, n).$$

In the case of the SDP (8), the (i,j)th element of the Schur complement matrix $\boldsymbol{B}$ for the HKM search direction [11, 17, 20] is given by

$$
\begin{aligned}
B_{ij} &= B_{ij}^{\text{sparse}} + B_{ij}^{\text{dense}}, \\
B_{ij}^{\text{sparse}} &= \sum_{p \in K_i \cap K_j} \boldsymbol{X}_p \boldsymbol{A}_{pi} \boldsymbol{S}_p^{-1} \bullet \boldsymbol{A}_{pj}, \\
&= \begin{cases} \star & \text{if } K_i \cap K_j \neq \emptyset \text{ or } i,\ j \in I_p \text{ for some } p \in M, \\ 0 & \text{otherwise}, \end{cases} \qquad (9) \\
B_{ij}^{\text{dense}} &= \boldsymbol{X}_{m+1} \boldsymbol{A}_{(m+1)i} \boldsymbol{S}_{m+1}^{-1} \bullet \boldsymbol{A}_{(m+1)j},
\end{aligned}
$$

for some $\boldsymbol{S}_p \succ \boldsymbol{O}$, which corresponds to the slack matrix variable for the LMI $\boldsymbol{A}_{p0} - \sum_{i \in I_p} \boldsymbol{A}_{pi} y_i \succeq \boldsymbol{O}$, $(p \in M \cup \{m+1\})$ and some $\boldsymbol{X}_p \succ \boldsymbol{O}$ $(p \in M \cup \{m+1\})$. Note that we have assumed that $\boldsymbol{X}_p$ and $\boldsymbol{S}_p^{-1}$ are fully dense in the identity (9), and that the identity (9) implies that the sparsity patterns of $\boldsymbol{B}^{\text{sparse}}$ and the csp matrix $\boldsymbol{R}$ coincide with each other. Thus the Schur complement matrix $\boldsymbol{B}$ has been represented with a matrix $\boldsymbol{B}^{\text{sparse}}$ with the same sparsity pattern as the csp matrix $\boldsymbol{R}$ and a dense matrix $\boldsymbol{B}^{\text{dense}}$. When we are concerned with a correlatively-sparse SDP without the dense constraint $\boldsymbol{A}_{(m+1)0} - \sum_{i \in N} \boldsymbol{A}_{(m+1)i} y_i \succeq \boldsymbol{O}$ in (8), the dense part $\boldsymbol{B}^{\text{dense}}$ vanishes. If $\boldsymbol{A}_{(m+1)i}$ $(i = 0, 1, 2, \ldots, n)$ are $r \times r$ matrices in a general almost-correlatively-sparse case, the rank of $\boldsymbol{B}^{\text{dense}} \leq r(r+1)/2$ as shown in the lemma below. Therefore, if r is small, a low-rank update technique can be used to solve the Schur complement equation with the coefficient matrix $\boldsymbol{B}$.

Lemma 4.1. *Let $\boldsymbol{F}_i$ $(i = 1, 2, \ldots, n)$ be an $r \times r$ matrix and $\boldsymbol{G}_i$ $(i = 1, 2, \ldots, n)$ an $r \times r$ symmetric matrix. Let $\boldsymbol{H}$ be an $n \times n$ matrix whose (i,j)th element is given by $H_{ij} = \boldsymbol{F}_i \bullet \boldsymbol{G}_j$ $(i = 1, 2, \ldots, n,\ j = 1, 2, \ldots, n)$. Then, rank $(\boldsymbol{H}) \leq r(r+1)/2$.*

Proof: Let

$$
\boldsymbol{f}_{k\ell} = \begin{pmatrix} [\boldsymbol{F}_1]_{k\ell} \\ [\boldsymbol{F}_2]_{k\ell} \\ \vdots \\ [\boldsymbol{F}_n]_{k\ell} \end{pmatrix}, \quad \boldsymbol{g}_{k\ell} = \begin{pmatrix} [\boldsymbol{G}_1]_{k\ell} \\ [\boldsymbol{G}_2]_{k\ell} \\ \vdots \\ [\boldsymbol{G}_n]_{k\ell} \end{pmatrix}.
$$

Then we see that

$$
\boldsymbol{H} = \sum_{k=1}^{r} \sum_{\ell=1}^{r} \boldsymbol{f}_{k\ell} \boldsymbol{g}_{k\ell}^T = \sum_{k=1}^{r} \sum_{\ell=1}^{k-1} (\boldsymbol{f}_{k\ell} + \boldsymbol{f}_{\ell k}) \boldsymbol{g}_{k\ell}^T + \sum_{\ell=1}^{r} \boldsymbol{f}_{\ell\ell} \boldsymbol{g}_{\ell\ell}^T.
$$

Now each term $(\boldsymbol{f}_{k\ell} + \boldsymbol{f}_{\ell k}) \boldsymbol{g}_{k\ell}^T$, $\boldsymbol{f}_{\ell\ell} \boldsymbol{g}_{\ell\ell}^T$ on the right hand is rank-1 matrix. Thus rank $(\boldsymbol{H}) \leq r(r+1)/2$ follows. ∎

We now apply the discussion above to the example (4) in Section 2.2. Recall that the example (4) can be described as an almost correlatively-sparse LOP of the form (3) with $m = 3$, $n = 4$, M, N, I_p $(p \in M)$ and C_p $(p \in M \cup \{4\})$ given in (5) and the csp matrix

$\boldsymbol{R}$ given in (6). We also note that a linear inequality is a special case of an LMI whose coefficients are 1×1 matrices. Hence we can treat the example as an SDP of the form (8). Specifically, the dense constraint $\boldsymbol{y} \in C_{m+1}$ is described as an LMI $-1+y_1+y_2+y_3+y_4 \geq 0$ with 1×1 coefficient matrix; hence $r = 1$ in the discussion above. Consequently, the Schur complement matrix $\boldsymbol{B}$ in this case is of the form

$$\boldsymbol{B} = \begin{pmatrix} \star & 0 & 0 & \star \\ 0 & \star & 0 & \star \\ 0 & 0 & \star & \star \\ \star & \star & \star & \star \end{pmatrix} + \text{“a } 4 \times 4 \text{ rank-1 dense matrix”}.$$

Remark 4.2. We assume that the HKM search direction [11, 17, 20] is used in the primal-dual interior-point method for the SDP (8). All the observation remains valid for the NT search direction [23] with a slight modification in the descriptions of B_{ij}^{sparse} and B_{ij}^{dense}.

5 SOCP

Let $\mathcal{K}_p$ ($p \in M \cup L$) be the second-order cones defined as

$$\mathcal{K}_p = \{\boldsymbol{x}_p = (x_{p0}, \boldsymbol{x}_{p1}) \in \mathbb{R} \times \mathbb{R}^{k_p-1} : x_{p0}^2 - \boldsymbol{x}_{p1}^T \boldsymbol{x}_{p1} \geq 0, x_{p0} \geq 0\},$$

where $L = \{m+1, m+2, \ldots, m^*\}$. We denote $\boldsymbol{x} \succeq_{\mathcal{K}_p} \boldsymbol{0}$ to mean $\boldsymbol{x} \in \mathcal{K}_p$. Suppose that each C_p in the LOP (3) is described as a second-order cone inequality such that

$$\begin{aligned} C_p &= \{(x_i : i \in I_p) : \boldsymbol{a}_{p0} - \sum_{i \in I_p} \boldsymbol{a}_{pi} y_i \succeq_{\mathcal{K}_p} \boldsymbol{0}\} \ (p \in M), \\ C_{m+1} &= \{\boldsymbol{y} : \boldsymbol{a}_{q0} - \sum_{i \in N} \boldsymbol{a}_{qi} y_i \succeq_{\mathcal{K}_p} \boldsymbol{0} \ (q \in L)\} \end{aligned}$$

for some column vectors $\boldsymbol{a}_{pi} \in \mathbb{R}^{k_p}$ $(i = 0, 1, 2, \ldots, n, p \in M \cup L)$. Then we have an SOCP of the form

$$\left.\begin{array}{ll} \text{maximize} & \boldsymbol{b}^T \boldsymbol{y} \\ \text{subject to} & \boldsymbol{a}_{p0} - \sum_{i \in I_p} \boldsymbol{a}_{pi} y_i \succeq_{\mathcal{K}_p} \boldsymbol{0} \ (p \in M), \\ & \boldsymbol{a}_{q0} - \sum_{i \in N} \boldsymbol{a}_{qi} y_i \succeq_{\mathcal{K}_q} \boldsymbol{0} \ (q \in L), \end{array}\right\} \tag{10}$$

and its dual

$$\left.\begin{array}{ll} \text{minimize} & \sum_{p \in M} \boldsymbol{a}_{p0}^T \boldsymbol{x}_p + \sum_{q \in L} \boldsymbol{a}_{q0} x_q \\ \text{subject to} & \sum_{p \in K_i} \boldsymbol{a}_{pi}^T \boldsymbol{x}_p + \sum_{q \in L} \boldsymbol{a}_{qi}^T \boldsymbol{x}_q = b_i \ (i \in N), \\ & \boldsymbol{x}_p \succeq_{\mathcal{K}_p} \boldsymbol{0} \ (p \in M \cup L), \end{array}\right\}$$

where K_i $(i \in N)$ is given by (7). The (i,j)th element of the Schur complement matrix $\boldsymbol{B}$ for the HKM search direction [27] is given by

$$
\begin{aligned}
B_{ij} &= B_{ij}^{\text{sparse}} + B_{ij}^{\text{dense}}, \\
B_{ij}^{\text{sparse}} &= \sum_{p \in K_i \cap K_j} \boldsymbol{a}_{pi}^T \boldsymbol{Q}_p \boldsymbol{a}_{pj}, \\
&= \begin{cases} \star & \text{if } K_i \cap K_j \neq \emptyset \text{ or } i, j \in I_p \text{ for some } p \in M, \\ 0 & \text{otherwise}, \end{cases} \\
B_{ij}^{\text{dense}} &= \sum_{q \in L} \boldsymbol{a}_{qi}^T \boldsymbol{Q}_q \boldsymbol{a}_{qj}.
\end{aligned}
\tag{11}
$$

Here the scaling matrix $\boldsymbol{Q}_p$ $(p \in M \cup L)$ is given by

$$
\begin{aligned}
&\boldsymbol{Q}_p = \frac{1}{\gamma(\boldsymbol{s}_p)^2} \begin{pmatrix} x_{p0}s_{p0} - \boldsymbol{x}_{p1}^T \boldsymbol{s}_{p1} & s_{p0}\boldsymbol{x}_{p1}^T - x_{p0}\boldsymbol{s}_{p1}^T \\ s_{p0}\boldsymbol{x}_{p1} - x_{p0}\boldsymbol{s}_{p1} & (\boldsymbol{x}_p^T \boldsymbol{s}_p)\boldsymbol{I} - \boldsymbol{s}_{p1}\boldsymbol{x}_{p1}^T - \boldsymbol{x}_{p1}\boldsymbol{s}_{p1}^T \end{pmatrix}, \\
&\gamma(\boldsymbol{s}_p) = \sqrt{\boldsymbol{s}_{p0}^2 - \boldsymbol{s}_{p1}^T \boldsymbol{s}_{p1}}
\end{aligned}
$$

for some $\boldsymbol{s}_p \succ_{\mathcal{K}_p} \boldsymbol{0}$, which corresponds to the slack variable for the second order-cone inequality $\boldsymbol{a}_{p0} - \sum_{i \in I_p} \boldsymbol{a}_{pi} y_i \succeq_{\mathcal{K}_p} \boldsymbol{0}$, $(p \in M \cup L)$ and some $\boldsymbol{x}_p \succ_{\mathcal{K}_p} \boldsymbol{0}$ $(p \in M \cup L)$. We have assumed that $\boldsymbol{x}_p$ and $\boldsymbol{s}_p$ are fully dense and $\boldsymbol{a}_{pi} \neq \boldsymbol{0}$, $\boldsymbol{a}_{pj} \neq \boldsymbol{0}$ for some $p \in K_i \cap K_j$ and every i, j such that $K_i \cap K_j \neq \emptyset$ to derive the identity (11). We note that (11) implies that the sparsity patterns of $\boldsymbol{B}^{\text{sparse}}$ and the csp matrix $\boldsymbol{R}$ coincide with each other. Thus the Schur complement matrix $\boldsymbol{B}$ has been splitted into a matrix $\boldsymbol{B}^{\text{sparse}}$ with the same sparsity pattern as the csp matrix $\boldsymbol{R}$ and a dense matrix $\boldsymbol{B}^{\text{dense}}$. If we assume that $\boldsymbol{a}_{qi} \neq \boldsymbol{0}$, $\boldsymbol{a}_{qj} \neq \boldsymbol{0}$ for some $q \in L$ for every $i, j \in N$, then $\boldsymbol{B}^{\text{dense}}$ becomes fully dense. The rank of the matrix $\boldsymbol{B}^{\text{dense}}$ is not greater than $\sum_{q \in L} \dim(\boldsymbol{a}_{q0})$. If $\sum_{q \in L} \dim(\boldsymbol{a}_{q0})$ is small, then a low-rank update can be used for the dense part.

We can split the scaling matrix $\boldsymbol{Q}_p$ such that

$$
\begin{aligned}
\boldsymbol{Q}_p &= \boldsymbol{Q}_p^{\text{sparse}} + \boldsymbol{Q}_p^{\text{dense}}, \\
\boldsymbol{Q}_p^{\text{sparse}} &= \frac{1}{\gamma(\boldsymbol{s}_p)^2} (\boldsymbol{x}_p \boldsymbol{s}_p) \begin{pmatrix} -1 & 0 \\ 0 & \boldsymbol{I} \end{pmatrix}, \\
\boldsymbol{Q}_p^{\text{dense}} &= \frac{1}{\gamma(\boldsymbol{s}_p)^2} \left(\begin{pmatrix} x_{p0} \\ \boldsymbol{x}_{p1} \end{pmatrix} \begin{pmatrix} s_{p0} & -\boldsymbol{s}_{p1} \end{pmatrix} + \begin{pmatrix} s_{p0} \\ -\boldsymbol{s}_{p1} \end{pmatrix} \begin{pmatrix} x_{p0} & \boldsymbol{x}_{p1} \end{pmatrix} \right).
\end{aligned}
\tag{12}
$$

Here the the matrix $\boldsymbol{Q}_p^{\text{dense}}$ is fully dense and rank-2. Goldfarb and Scheinberg [9] proposed to utilize this type of splitting to represent the Schur complement matrix $\boldsymbol{B}$ as a sum of a sparse matrix and a dense matrix of low-rank. When $\boldsymbol{a}_{pi}$ $(i \in N)$ is sparse for some $p \in M \cup L$, we can incorporate their method to explore the sparsity of the Schur complement matrix $\boldsymbol{B}$ further.

Remark 5.1. We assume that the HKM search direction is used in the primal-dual interior-point method for the SOCP (10). All the observation remains valid for the NT search direction [23, 27] with a slight modification in the scaling matrix $\boldsymbol{Q}_p$. The method in [9] was described for the NT search direction, but the method remains valid for the HKM direction. We note that $\boldsymbol{Q}_p^{\text{dense}}$ in (12) is of rank-2 for the HKM search direction while it is of rank-1 for the NT search direction.

6 Polynomial semidefinite program

Polynomial SDPs have a wide variety of applications in the system and control theory. They also describe an intermediate optimization problem equivalent to a given sparse polynomial optimization problem (POP) when the sparse SDP relaxation [16, 18, 30] is derived for the POP. In this section, we focus on a correlatively-sparse polynomial SDP, which is obtained from a sparse POP in the relaxation process of the POP to an SDP by linearization. We show that the SDP relaxation problem inherits the correlative sparsity from the polynomial SDP.

6.1 Polynomial SDP and its linearization

We restrict our attention to a sparse unconstrained POP, and briefly mention how a correlatively-sparse polynomial SDP is derived from it. See [16, 18, 30] for more discussions on general sparse constrained POPs.

Let $\mathbb{Z}_+$ denote the set of nonnegative integers. For every $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}_+^n$ and every n-dimensional vector variable $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$, we use the notation $\boldsymbol{x}^{\boldsymbol{\alpha}} = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$. Then a real-valued polynomial $\phi(\boldsymbol{x})$ in $\boldsymbol{x}$ can be written as $\phi(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} c_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}}$ for some nonempty finite set $\mathcal{A} \subset \mathbb{Z}_+^n$ and some nonzero real numbers $c_{\boldsymbol{\alpha}}$ $(\boldsymbol{\alpha} \in \mathcal{A})$. We assume that $\boldsymbol{x}^{\mathbf{0}} = 1$ for any $\boldsymbol{x}$.

We consider an unconstrained POP

$$\text{maximize} \quad -\sum_{\alpha \in \tilde{\mathcal{F}}} \boldsymbol{c}_\alpha \boldsymbol{x}^\alpha, \tag{13}$$

where $\tilde{\mathcal{F}}$ denotes some nonempty finite subset $\tilde{\mathcal{F}}$ of $\mathbb{Z}_+^n$ and $c_{\boldsymbol{\alpha}}$ $(\boldsymbol{\alpha} \in \tilde{\mathcal{F}})$ nonzero real numbers. Let $\delta_0 = \max \left\{ \sum_{i=1}^n \alpha_i : \boldsymbol{\alpha} \in \tilde{\mathcal{F}} \right\}$ denote the degree of the objective polynomial function, and assumed to be a positive even integer since otherwise the problem (13) is unbounded. We define a graph $G(N, E)$ consisting of the node set $N = \{1, 2, \ldots, n\}$ and the edge set

$$E = \{\{i, j\} \subset N : i < j,\ \alpha_i > 0 \ \text{ and } \alpha_j > 0 \text{ for some } \boldsymbol{\alpha} \in \tilde{\mathcal{F}}\}.$$

This graph $G(N, E)$ represents the sparsity of the Hessian matrix of the objective polynomial function. Let $G(N, \bar{E})$ be a chordal extension of the graph $G(N, E)$; $G(N, \bar{E})$ is a chordal graph consisting of the node set N and an edge set $\bar{E} \supseteq E$, and $I_p \subseteq N$ $(p \in M_{\max} = \{1, 2, \ldots, \ell\})$ denote the maximal cliques of $G(N, \bar{E})$, which satisfies the running intersection property (2) for some permutation $\pi(1), \pi(2), \ldots, \pi(\ell)$ of $1, 2, \ldots, \ell$. Let $\rho = \delta_0/2$. For every $I \subseteq N$, let $\boldsymbol{u}(\boldsymbol{x}; I, \rho)$ denote a column vector consisting of the monomials in the set

$$\left\{ \boldsymbol{x}^{\boldsymbol{\alpha}} : \sum_{i \in I_p} \alpha_i \leq \rho,\ \alpha_j = 0\ (j \notin I_p) \right\}.$$

For every $p \in M_{\max}$, define

$$\boldsymbol{F}_p(x_i : i \in I_p) = \boldsymbol{u}(\boldsymbol{x}; I_p, \rho)\boldsymbol{u}(\boldsymbol{x}; I_p, \rho)^T,$$

which is symmetric, rank-1 and positive semidefinite for any $\boldsymbol{x}$. Let $m = \ell$ and $M = M_{\max}$. Now we are ready to describe a polynomial SDP which is equivalent to the unconstrained POP.

$$\left.\begin{array}{ll} \text{maximize} & -\sum_{\alpha \in \tilde{\mathcal{F}}} \boldsymbol{c}_\alpha \boldsymbol{x}^\alpha \\ \text{subject to} & \boldsymbol{F}_p(x_i : i \in I_p) \succeq \boldsymbol{O}\ (p \in M). \end{array}\right\} \tag{14}$$

6.2 Correlatively-sparse polynomial semidefinite programs and their linearization

In the previous subsection, we have derived a polynomial SDP (14) which is equivalent to the POP (13). In this subsection, we begin with a general polynomial SDP of the form (14) such that each $\boldsymbol{F}_p(x_i : i \in I_p)$ represents a polynomial in the variable x_i $(i \in I_p)$ with symmetric matrix coefficients. For every $p \in M$, let

$$C_p \quad = \quad \{(x_i : i \in I_p) : \boldsymbol{F}_p((x_i : i \in I_p)) \succeq \boldsymbol{O}\}.$$

Then we rewrite the polynomial SDP (14) as

$$\begin{array}{ll} \text{maximize} & -\sum_{\alpha \in \tilde{\mathcal{F}}} \boldsymbol{c}_\alpha \boldsymbol{x}^\alpha \\ \text{subject to} & (x_i : i \in I_p) \in C_p\ (p \in M), \end{array}$$

so that the csp matrix $\boldsymbol{R}$ can be defined in the same way as in Section 2.1. We say that the polynomial SDP (14) is correlatively-sparse if the csp matrix $\boldsymbol{R}$ allows a sparse Cholesky factorization with no fill-in.

For every positive integer ω, let

$$\mathcal{F}(I_p, \omega) = \left\{ \boldsymbol{\alpha} \in \mathbb{Z}^n_+ : \sum_{i \in I_p} \alpha_i \leq \omega,\ \alpha_i = 0\ (i \notin I_p) \right\}\ (p \in M).$$

We assume that for some positive integer ω, each $\boldsymbol{F}_p((x_i : i \in I_p))$ is described as

$$\boldsymbol{F}_p((x_i : i \in I_p)) = \sum_{\boldsymbol{\alpha} \in \mathcal{F}(I_p, \omega)} \boldsymbol{A}_{p\alpha} \boldsymbol{x}^\alpha.$$

(Note that 2ρ corresponds to ω in the previous subsection's discussion). Here $\boldsymbol{A}_{p\alpha}$ denotes a symmetric matrix; some $\boldsymbol{A}_{p\alpha}$ can be zero matrices. Thus the polynomial SDP under consideration can be written as

$$\left.\begin{array}{ll} \text{maximize} & -\sum_{\alpha\in\tilde{\mathcal{F}}} \boldsymbol{c}_\alpha \boldsymbol{x}^\alpha \\ \text{subject to} & \sum_{\boldsymbol{\alpha}\in\mathcal{F}(I_p,\omega)} \boldsymbol{A}_{p\alpha}\boldsymbol{x}^\alpha \succeq \boldsymbol{O} \ (p=1,2,\ldots,m). \end{array}\right\} \tag{15}$$

We now linearize the polynomial SDP (15) by replacing each $\boldsymbol{x}^\alpha$ by a single variable y_α to obtain an SDP as its relaxation.

$$\left.\begin{array}{ll} \text{maximize} & -\sum_{\alpha\in\tilde{\mathcal{F}}} \boldsymbol{c}_\alpha y_\alpha \\ \text{subject to} & (y_\alpha : \boldsymbol{\alpha}\in\mathcal{F}(I_p,\omega)) \in \tilde{C}_p \ (p=1,2,\ldots,m), \end{array}\right\} \tag{16}$$

where

$$\tilde{C}_p \;=\; \{(y_\alpha : \boldsymbol{\alpha}\in\mathcal{F}(I_p,\omega)) : \sum_{\boldsymbol{\alpha}\in\mathcal{F}(I_p,\omega)} \boldsymbol{A}_{p\alpha} y_\alpha \succeq \boldsymbol{O}\}.$$

The following lemma demonstrates the relationship between the correlative sparsity of the polynomial SDP (15) and that of its linearized SDP (16).

Lemma 6.1. *Suppose that the polynomial SDP (14) is correlatively sparse, and that the running intersection property (2) holds; for simplicity, suppose that (2) holds with $\pi(p) = p$ ($p \in M_{\max} = \{1, 2, \ldots, \ell\}$). Then the family of index sets $\mathcal{F}(I_p, \omega)$ ($p \in M$) involved in the SDP (16) satisfies the running intersection property:*

$$\forall p \in \{1,2,\ldots,\ell-1\}\ \exists r \geq p+1;\ \mathcal{F}(I_p,\omega)\cap\left(\cup_{q=p+1}^{\ell}\mathcal{F}(I_q,\omega)\right) \subset \mathcal{F}(I_r,\omega).$$

(Hence the SDP (16) is correlatively sparse).

Proof: Let $p \in \{1, 2, \ldots, \ell-1\}$. By (2), there is an $r \geq p+1$ such that

$$I_p \cap \left(\cup_{q=p+1}^{\ell} I_q\right) \subset I_r$$

We will show that $\mathcal{F}(I_p,\omega)\cap\left(\cup_{q=p+1}^{\ell}\mathcal{F}(I_q,\omega)\right) \subset \mathcal{F}(I_r,\omega)$. Assume that $\boldsymbol{\alpha} \in \mathcal{F}(I_p,\omega)\cap\left(\cup_{q=p+1}^{\ell}\mathcal{F}(I_q,\omega)\right)$. It follows that $\boldsymbol{\alpha} \in \mathcal{F}(I_p,\omega)\cap\mathcal{F}(I_{p'},\omega)$ for some $p' \geq p+1$. By the definitions of $\mathcal{F}(I_p,\omega)$ and $\mathcal{F}(I_{p'},\omega)$ that

$$\sum_{i=1}^{n}\alpha_i \leq \omega,\ \alpha_i = 0\ (i \notin I_p),\ \alpha_i = 0\ (i \notin I_{p'}),$$

which implies that $\boldsymbol{\alpha} \in \mathcal{F}(I_p \cap I_{p'}, \omega)$. On the other hand, we know that

$$I_p \cap I_{p'} \subseteq I_p \cap \left(\cup_{q=p+1}^{\ell} I_q\right) \subset I_r;$$

hence $I_p \cap I_{p'} \subset I_r$. This implies that $\boldsymbol{\alpha} \in \mathcal{F}(I_p \cap I_{p'},\omega) \subset \mathcal{F}(I_r,\omega)$. ∎

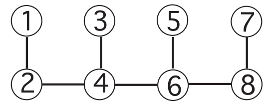

Figure 1: The graph $G(N, E)$ representing the sparsity of the Chained wood function with $n = 8$.

6.3 An example

As an example of an unconstrained POP, we consider minimizing the Chained wood function [5]

$$1 + \sum_{i \in J} \big(100(x_{i+1} - x_i^2)^2 + (1 - x_i)^2 + 90(x_{i+3} - x_{i+2}^2)^2 + (1 - x_{i+2})^2$$
$$+10(x_{i+1} + x_{i+3} - 2)^2 + 0.1(x_{i+1} - x_{i+3})^2\big),$$

where $J = \{1, 3, 5, \ldots, n-3\}$ and n is a multiple of 4. The degree of the Chained wood polynomial function is 4; hence $\delta_0 = 4$ and $\rho = \delta_0/2 = 2$. We first construct a graph $G(N, E)$ that represents the sparsity of the polynomial according to the method mentioned in Section 6.1. We then see that the graph $G(N, E)$ does not have any cycle (with length greater than 2); hence $G(N, E)$ itself is a chordal graph. Figure 1 illustrates the graph $G(N, E)$ for the case of $n = 8$. The maximal cliques of the graph $G(N, E)$ are

$$\{i, i+1\},\ \{i+1, i+3\},\ \{i+2, i+3\}\ (i \in J).$$

Using this family of maximal cliques, we construct a polynomial SDP (14) that is equivalent to the minimization of the Chained wood function. Recall that we have rewritten the polynomial SDP (14) as in (15). Figures 2 and 3 show the sparsity pattern of the csp matrix $\boldsymbol{R}$ of the polynomial SDP (15) and that of the Schur complement matrix $\boldsymbol{B}$ of the SDP (16) obtained linearizing the polynomial SDP (15), respectively. These figure are obtained by SparsePOP [29] applied to the minimization of the Chained wood function. We notice that the sparsity pattern of $\boldsymbol{R}$ of the polynomial SDP (15) is magnified in the Schur complement matrix, keeping the same sparsity pattern.

7 Concluding remarks

We have shown that the correlative sparsity of LP, SOCP, and SDP is maintained in the Schur complement matrix using a correlatively-sparse LOP and an almost correlative-sparse LOP. A sparse Cholesky factorization applied to the sparse part of the Schur complement matrix entails no fill-ins. Thus, the inherited sparsity of the Schur complement matrix can be used to increase the computational efficiency of primal-dual interior-point methods. The correlative sparsity of polynomial SDPs is also shown to lead to the correlative sparsity of the SDP relaxation, resulting in the Schur complement matrix that can be factorized with

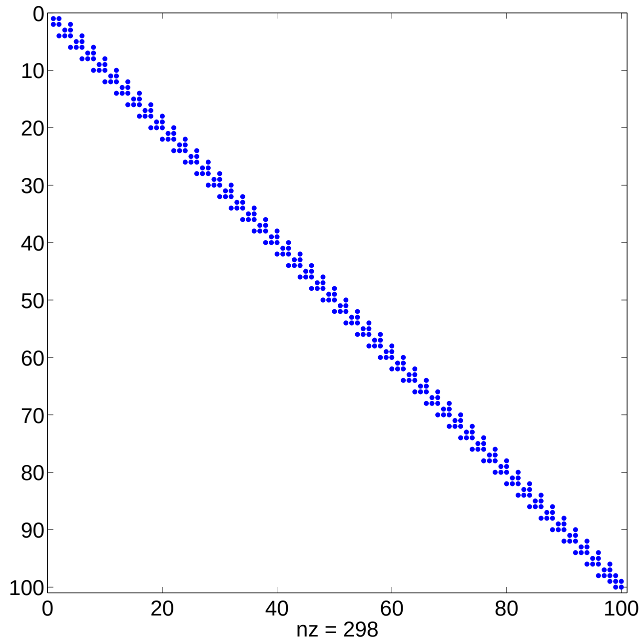

Figure 2: The csp matrix $\boldsymbol{R}$ induced from the minimization of the Chained wood function with $n = 100$

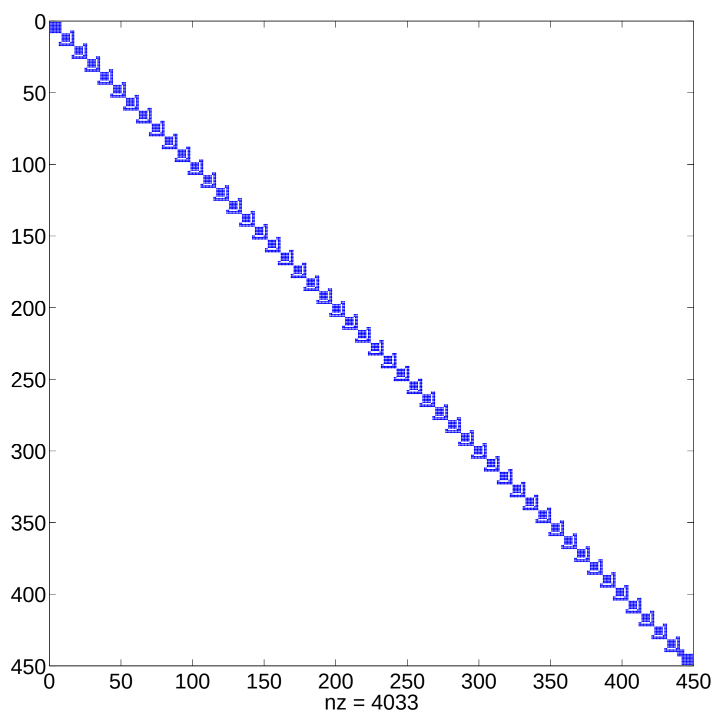

Figure 3: The sparsity pattern of the Schur complement matrix $\boldsymbol{B}$ of the sparse SDP relaxation of minimizing the Chained wood function with $n = 100$

no fill-in. We have confirmed this with the minimization of the Chained wood function in Section 6.

When a given optimization problem does not show enough sparsity to exploit, the method proposed in [14] can be applied to bring out the underlying sparsity. The method is based on a linear transformation of the variables and transforms an optimization problem to one with increased sparsity. Then, the technique of extending the sparsity to the correlative sparsity mentioned in Section 2.1 can be used to improve the computational efficiency of primal-dual interior-point methods.

References

[1] I. Adler, M. G. C. Resende, G. Veiga and N. Karmarkar, "An implementation of Karmarkar's algorithm for linear programming", *Mathematical Programming* **44** (1989) 297–335.

[2] K. D. Anderson, "A modified Schur complement method for handling dense columns in interior-point methods for linear programming", *ACM Transactions on Mathematical Software* **22** (1996) 348–356.

[3] C. Ashcraft, D. Pierce, D. K. Wah and J. Wu, "The reference manual for SPOOLES, release 2.2: An object oriented software library for solving sparse linear systems of equations", Boeing Shared Services Group, P. O. Box 24346, Mail Stop 7L-22, Seattle, WA 98124, January 1999; Available at http://netlib.bell-labs.com/netlib/linalg/spooles/spooles.

[4] J. R. S. Blair and B. Peyton, "An introduction to chordal graphs and clique trees", in *Graph Theory and Sparse Matrix Computation*, A George, J. R. Gilbert and J. W. H. Liu, eds., Springer-Verlag, New York, 1993, pp.1–29.

[5] A. R. Conn, N. I. M. Gould and Ph. L. Toint, "Testing a class of methods for solving minimization problems with simple bounds on the variables", *Mathematics of Computation* **50** (1988) 399–430.

[6] K. Fujisawa, M. Kojima, K. Nakata "Exploiting sparsity in primal-dual Interior-point methods for Semidefinite Programming", *Mathematical Programming* **79** (1997) 235–254.

[7] M. Fukuda, M. Kojima, K. Murota and K. Nakata, "Exploiting sparsity in semidefinite programming via matrix completion I: general framework", *SIAM Journal on Optimization* **11** (2001) 647–674.

[8] M. C. Golumbic, *Algorithmic Graph Theory and Perfect Graphs*, Academic Press, New York, 1980.

[9] D. Goldfarb, K. Scheinberg, "Product-form Cholesky factorization in interior-point methods for second order cone programming", *Mathematical Programming* **103** (2005) 153–179.

[10] A. Griewank and Ph. L. Toint, "On the unconstrained optimization of partially separable functions", in M. J. D. Powell, ed., *Nonlinear Optimization 1981*, Academic Press, New York, 1982, pp.301–312.

[11] C. Helmberg, F. Rendl, R. J. Vanderbei and H. Wolkowicz, "An interior-point method for semidefinite programming", *SIAM Journal on Optimization* **6** (1996) 342-361.

[12] G. Karypis and V. Kumar, "METIS — A software package for partitioning unstructured graphs, partitioning meshes, and computing fill-reducing orderings of sparse matrices, version 4.0 —", Department of Computer Science/Army HPC Research Center, University of Minnesota, Minneapolis, MN 55455, September 1998; Available at http://www-users.cs.umn.edu/~karypis/metis/metis.

[13] S. Kim, M. Kojima and H. Waki, "Generalized Lagrangian duals and sums of squares relaxations of sparse polynomial optimization problems", *SIAM Journal on Optimization* **15** (2005) 697–719.

[14] S. Kim, M. Kojima, and Ph. L. Toint, "Recognizing underlying sparsity in optimization", Research Report B-428, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-Okayama, Meguro, Tokyo 152-8552, Japan, May 2006.

[15] M. Kojima, S. Kim and H. Waki, "Sparsity in sums of squares of polynomials", *Mathematical Programming* **103** (2005) 45–62.

[16] M. Kojima and M. Muramatsu, "A note on sparse SOS and SDP relaxations for polynomial optimization problems over symmetric cones", Research Report B-421, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-Okayama, Meguro-ku, Tokyo 152-8552, Japan, January 2006.

[17] M. Kojima, S. Shindoh and S. Hara, "Interior-point methods for the monotone semidefinite linear complementarity problem in symmetric matrices", *SIAM Journal on Optimization* **7** (1997) 86–125.

[18] J. B. Lasserre, "Convergent semidefinite relaxation in polynomial optimization with sparsity", LAAS-CNRS, 7 Avenue du Colonel Roche, 31077 Toulouse, France (2005).

[19] J. G. Lewis, B. W. Peyton and A. Pothen, "A fast algorithm for reordering sparse matrices for parallel factorization", *SIAM Journal Scientific Computing* **10** (1989) 1146–1173.

[20] R. D. C. Monteiro, "Primal-dual path-following algorithms for semidefinite programming", *SIAM Journal on Optimization* **7** (1997) 663-678.

[21] K. Nakata, K. Fujisawa, and M. Kojima, "Using the conjugate gradient method in interior-point methods for semidefinite programs (in Japanese)", *Proceedings of Institute of Statistical Mathematics* **46** (1998) 297–316.

[22] 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* **95** (2003) 303–327.

[23] Yu. E. Nesterov and M. J. Todd, “Primal-dual interior-point for self-scaled cones”, *SIAM Journal on Optimization* **8** (1988) 324-364.

[24] F. J. Sturm “Implementation of interior point methods for mixed semidefinite and second order cone optimization”, *Optimization Methods and Software* **17** (2002) 1105–1154.

[25] K.C. Toh and M. Kojima, “Solving some large scale semidefinite programs via the conjugate residual method’, *SIAM Journal on Optimization* **12** (2002) 669–691.

[26] K. C. Toh, “Solving large scale semidefinite programs via an iterative solver on the augmented systems”, *SIAM Journal on Optimization* **14** (2004) 670–698.

[27] T. Tsuchiya, “A convergence analysis of the scaling-invariant primal-dual path-following algorithm for second-order cone programming”, *Optimization Methods and Software* **11/12** (1999) 141–182.

[28] R. J. Vanderbei, “Symmetric quasidefinite matrices”, *SIAM Journal on Optimization* **5** (1995) 100–113.

[29] H. Waki, S. Kim, M. Kojima and M. Muramatsu, ”SparsePOP : a Sparse Semidefinite Programming Relaxation of Polynomial Optimization Problems”, Research Report B-414, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-Okayama, Meguro, Tokyo 152-8552, Japan, March 2005.

[30] H. Waki, S. Kim, M. Kojima and M. Muramatsu, “Sums of squares and semidefinite programming relaxations for polynomial optimization problems with structured sparsity” , *SIAM Journal on Optimization* **17** (2006) 218–242.