Linear optimization over homogeneous matrix cones

Levent Tunçel

Department of Combinatorics and Optimization, University of Waterloo,
Waterloo, Ontario N2L 3G1, Canada
E-mail: levent.tuncel@uwaterloo.ca

Lieven Vandenberghe

Department of Electrical and Computer Engineering, UCLA,
Los Angeles, CA 90095-1594, USA
E-mail: vandenbe@ucla.edu

A convex cone is homogeneous if its automorphism group acts transitively on the interior of the cone. Cones that are homogeneous and self-dual are called symmetric. Conic optimization problems over symmetric cones have been extensively studied, particularly in the literature on interior-point algorithms, and as the foundation of modelling tools for convex optimization. In this paper we consider the less well-studied conic optimization problems over cones that are homogeneous but not necessarily self-dual.

We start with cones of positive semidefinite symmetric matrices with a given sparsity pattern. Homogeneous cones in this class are characterized by nested block-arrow sparsity patterns, a subset of the chordal sparsity patterns. Chordal sparsity guarantees that positive define matrices in the cone have zero-fill Cholesky factorizations. The stronger properties that make the cone homogeneous guarantee that the inverse Cholesky factors have the same zero-fill pattern. We describe transitive subsets of the cone automorphism groups, and important properties of the composition of log-det barriers with the automorphisms.

Next, we consider extensions to linear slices of the positive semidefinite cone, and review conditions that make such cones homogeneous. An important example is the matrix norm cone, the epigraph of a quadratic-over-linear matrix function. The properties of homogeneous sparse matrix cones are shown to extend to this more general class of homogeneous matrix cones.

We then give an overview of the algebraic theory of homogeneous cones due to Vinberg and Rothaus. A fundamental consequence of this theory is that every homogeneous cone admits a spectrahedral (linear matrix inequality) representation.

We conclude by discussing the role of homogeneous structure in primal–dual symmetric interior-point methods, contrasting this with the well-developed algorithms for symmetric cones that exploit the strong properties of self-scaled barriers, and with symmetric primal–dual methods for general convex cones.

2020 Mathematics Subject Classification: Primary 90-02, 90C25

Secondary 15B48, 65K05, 90C22, 90C51

CONTENTS

1. Introduction

The conic programming framework has been used extensively in the development of convex optimization theory, applications, algorithms and modelling (Nesterov and Nemirovskii 1994, Ben-Tal and Nemirovski 2001, Boyd and Vandenberghe 2004). As with any optimization problem, a fundamental step in a successful treatment of large-scale conic programs is the identification and efficient exploitation of special structure. In this paper we discuss convex cones represented as slices of the positive semidefinite cone, i.e. as intersections

$$K = \mathcal{V} \cap \mathbb{S}_+^N \tag{1.1}$$

of $\mathbb{S}_+^N$ (the cone of symmetric positive semidefinite $N \times N$ matrices) and a subspace $\mathcal{V}$, and we examine the special structure of $\mathcal{V}$ that makes K a *homogeneous* convex cone. A convex cone is homogeneous if, for every pair of points in its interior, there exists an automorphism of the cone that maps one point to the other.

Inequalities with respect to slices of the positive semidefinite cone arise in non-symmetric formulations of semidefinite programming problems. Consider a semidefinite program (SDP) in inequality form

$$\begin{aligned}
\text{minimize} \quad & c^\top y \\
\text{subject to} \quad & \sum_{i=1}^m y_i A_i + X = B \\
& X \succeq 0
\end{aligned} \tag{1.2a}$$

and its dual problem

$$\begin{aligned}
\text{maximize} \quad & -\langle B, S \rangle & \text{(1.2b)} \\
\text{subject to} \quad & \langle A_i, S \rangle + c_i = 0, \quad i = 1, \ldots, m \\
& S \geq 0.
\end{aligned}$$

The primal variables are $y \in \mathbb{R}^m$, $X \in \mathbb{S}^N$. The dual variable is $S \in \mathbb{S}^N$. The inequalities $X \geq 0$, $S \geq 0$ mean that $X, S \in \mathbb{S}_+^N$. The positive semidefinite matrix cone $\mathbb{S}_+^N$ is a *symmetric cone*, i.e. self-dual and homogeneous, and the special properties of symmetric cones are key to the design and implementation of primal–dual interior-point algorithms for semidefinite optimization.

If the matrices $A_1, \ldots, A_m, B$ all belong to a subspace $\mathcal{V}$ of $\mathbb{S}^N$, the problems (1.2) are equivalent to the pair of conic optimization problems

$$\begin{aligned}
\text{minimize} \quad & c^\top y & \text{(1.3a)} \\
\text{subject to} \quad & \sum_{i=1}^m y_i A_i + X = B \\
& X \in K
\end{aligned}$$

and

$$\begin{aligned}
\text{maximize} \quad & -\langle B, S \rangle & \text{(1.3b)} \\
\text{subject to} \quad & \langle A_i, S \rangle + c_i = 0, \quad i = 1, \ldots, m \\
& S \in K^*,
\end{aligned}$$

where K is defined in (1.1), K^* is the dual cone of K, and the variables X, S are matrices in $\mathcal{V}$. The formulation (1.3) is of interest for large-scale algorithm development because the subspace $\mathcal{V}$ can be of much lower dimension than $\mathbb{S}^N$, possibly as low as the dimension of the span of the coefficient matrices $\bar{\mathcal{V}} = \text{span}\{A_1, \ldots, A_m, B\}$. However, the efficiency of algorithms for handling the conic inequalities with respect to K and K^* depends on more properties of $\mathcal{V}$ than just the dimension, and this may require embedding $\bar{\mathcal{V}}$ in a higher-dimensional subspace. The standard choice in current primal–dual interior-point methods is to embed $\bar{\mathcal{V}}$ in a space of block-diagonal matrices with dense diagonal blocks. For this choice of $\mathcal{V}$, the cone K is symmetric. For almost all other subspaces $\mathcal{V}$, the cone K and its dual K^* are not equal; hence they are not symmetric cones. (The exceptions are semidefinite representations of the small number of symmetric cones, for example direct products of second-order cones.) However, K and K^* may still be homogeneous. Homogeneous convex cones were algebraically classified in the 1960s by Vinberg (1965b) and are the subject of a large literature in algebra and statistics (Letac and Massam 2007, Andersson and Wojnar 2004, Boutouria, Hassairi and Massam 2011, Khare and Rajaratnam 2011). The conditions for a matrix cone of the form (1.1) to be homogeneous have been studied by Letac and Massam (2007) and Ishi (2013, 2015). Homogeneous cones have several important properties in common with symmetric cones. One can note, for example, that their definition contains two fundamental concepts in primal–dual interior-point

Table 1.1. Four classes of sparse positive semidefinite matrix cones, classified by type of sparsity, the linear algebra tools available for their analysis, and fundamental properties of the cones.

Sparsity pattern	Linear algebra	Convex cone
dense	spectral theory	symmetric
homogeneous chordal	zero-fill Cholesky factor and inverse factor	homogeneous
chordal	zero-fill Cholesky factor	slice of PSD cone
general	sparse Cholesky factor	slice of PSD cone

algorithms for optimization over symmetric cones. The automorphisms of a cone (invertible linear transformations that leave the cone invariant) are the *scalings* used in interior-point methods, for example the positive diagonal scalings of the non-negative orthant in algorithms for linear programming. The second property, that the automorphisms act transitively in the interior of the cone, implies that any given pair of primal and dual iterates can be mapped to the same point by a cone automorphism, as we will discuss in Section 4. Hence homogeneous cones are a natural subject of study in conic optimization. However, with some notable exceptions (Güler 1996, Güler and Tunçel 1998, Truong and Tunçel 2004, Chua 2009), work on algorithms for homogeneous conic optimization appears to be quite limited. It is the purpose of this article to describe properties of homogeneous matrix cones that are useful in algorithms for optimization problems of the form (1.3). We also discuss specific examples and structural properties that may be useful for optimization modelling tools.

In Sections 2–4 we first consider matrix subspaces $\mathcal{V}$ defined by sparsity patterns. If the coefficient matrices $A_1, \ldots, A_m, B$ in problem (1.2) have a common (aggregate) sparsity pattern, then the subspace $\mathcal{V}$ in (1.3) can be defined as the set of symmetric $N \times N$ matrices with that pattern, or any extension of the aggregate sparsity pattern. The primal cone K is the cone of positive semidefinite matrices with a given sparsity pattern; the dual cone K^* is the cone of symmetric matrices with the same sparsity pattern that have a positive semidefinite completion. The non-symmetric conic formulation (1.3) has been studied in recent approaches to exploit sparsity in sparse semidefinite optimization (Fukuda, Kojima, Murota and Nakata 2000/01, Benson, Ye and Zhang 2000, Andersen, Dahl and Vandenberghe 2010a, Srijuntongsiri and Vavasis 2004, Burer 2003). Table 1.1 summarizes the definitions that relate this paper to existing literature on semidefinite programming. It distinguishes sparse positive semidefinite matrix cones by type of sparsity. At the top level we have the dense positive semidefinite cones (i.e. without any restriction

on the sparsity pattern). The dense positive semidefinite cone is symmetric (self-dual and homogeneous). Symmetric primal–dual algorithms for them rely heavily on eigenvalue and generalized eigenvalue decompositions of symmetric positive semidefinite matrices (e.g. for computing the matrix geometric mean, or for joint diagonalization of positive definite matrices). At the lowest level of the table we have the positive semidefinite matrix cones with a general, unstructured sparsity pattern. They form lower-dimensional slices of the positive semidefinite cone. Such cones are convex, but not homogeneous or self-dual. Implementations of non-symmetric interior-point algorithms for these cones, for example dual barrier algorithms (Benson *et al.* 2000), benefit from the possibility of computing sparse Cholesky factors, using fill-reducing ordering heuristics. Level three in the table is occupied by the positive semidefinite matrices with chordal sparsity patterns. Chordal sparsity has been studied intensively in sparse semidefinite optimization (see Vandenberghe and Andersen 2014, Zheng, Fantuzzi and Papachristodoulou 2021 for recent surveys). The chordal structure can be exploited to formulate efficient algorithms for key computations needed in semidefinite programming algorithms, such as the evaluation of primal and dual barrier functions and their derivatives, and finding maximum-determinant or minimum-rank positive semi-definite completions (Griewank and Toint 1984, Agler, Helton, McCullough and Rodman 1988, Grone, Johnson, Sá and Wolkowicz 1984). All of these algorithms can be derived from the basic property that positive semidefinite matrices with a chordal sparsity pattern have zero-fill Cholesky factorizations. The second row of the table is the focus of Sections 2–4 of this paper. The sparsity patterns that are referred to here as 'homogeneous chordal' define matrix cones that are homogeneous but not necessarily symmetric. These sparsity patterns have been characterized by Letac and Massam (2007, Theorem 2.2) and Ishi (2013, Theorem A). As we will discuss in Sections 2 and 3, they are block-arrow sparsity patterns and recursive generalizations of block-arrow structures. They form a subset of the chordal patterns, with the additional useful property that the inverse Cholesky factor has the same, zero-fill, sparsity pattern as the Cholesky factor itself.

Note that any class of semidefinite programming problems on a higher level in the table includes the lower ones. Without loss of generality, one can always extend a general sparsity pattern to make it chordal, or a chordal pattern to make it homogeneous chordal, or a homogeneous chordal sparsity pattern to make it dense. However, there is an obvious trade-off. The higher levels come with stronger results and more powerful techniques from linear algebra, and with more efficient primal, dual or primal–dual conic optimization algorithms. They also embed the optimization problem in higher-dimensional spaces and exploit less of the detailed structure in the sparsity pattern.

The three sections on homogeneous sparse matrix cones are organized as follows. Section 2 is a survey of results and algorithms from sparse matrix and graph theory related to chordal and homogeneous chordal sparsity patterns. In Section 3 we show that the positive semidefinite cone with a homogeneous chordal pattern and

the associated dual cone are homogeneous. We establish a transitive subset of the automorphism group constructed from congruences with sparse lower-triangular matrices. In Section 4 we derive implications for the log-det barrier function and its conjugate. We show that the Hessians of the logarithmic barrier functions can be factorized as a composition of a cone automorphism and its adjoint. This leads to a generalization of the Nesterov–Todd scaling point for symmetric cones.

In Section 5 we then turn to more general homogeneous slices of the positive semidefinite matrix cone, with subspaces $\mathcal{V}$ that can be defined by other linear relations than the sparsity pattern. The properties of $\mathcal{V}$ that make the cone (1.1) homogeneous are described by Ishi (2015). The results in this section will parallel the properties of homogeneous sparse matrix cones. In particular, Cholesky factors and inverse Cholesky factors inherit the structure of the subspace $\mathcal{V}$.

Section 6 reviews the general, algebraic classifications of all homogeneous cones and connects these theories to the earlier sections. An important result is that every homogeneous cone has a semidefinite representation, that is, it is linearly isomorphic to a slice of the positive semidefinite cone.

We conclude the paper with a survey of recent work on interior-point methods for non-symmetric conic optimization, and point out the potential benefits of exploiting the special properties of homogeneous cones (Section 7). The two appendices contain background material from graph theory and algorithmic details.

The paper is primarily intended as a survey. Its main contributions are the following.

- We identify a class of conic optimization problems (based on homogeneous sparse matrix cones, called *homogeneous chordal cones*) which lie strictly between SDPs and homogeneous cone programming problems (in the context of the set of convex cones K allowed in the optimization problems (1.3)). In this context, the class of convex optimization problems over homogeneous chordal cones provides a generalization of second-order cone programming that has important computational advantages over semidefinite programming.

- We build on results from convex optimization and analysis, graph theory, data structures and algorithms, sparse matrix computation and theory, and abstract algebra and show how to perform fundamental linear algebra operations in an efficient way for many families of algorithms for our class of conic optimization problems.

- We show how to compute primal and dual scalings that are automorphisms of the underlying cones, and in so doing we solve an open problem about the existence of automorphism-based primal–dual scalings for pairs of interior points in homogeneous cones and in their duals.

- We extend the results from homogeneous sparse matrix cones to homogeneous matrix cones defined by slices of the positive semidefinite cone. Constraints of this type are important in semidefinite representations of the spectral matrix norm and the trace norm.

2. Homogeneous chordal sparsity

We let $\mathbb{S}^N$ denote the space of $N \times N$ symmetric matrices with real entries, let $\mathbb{S}^N_+$ be the convex cone of positive semidefinite matrices in $\mathbb{S}^N$, and let $\mathbb{S}^N_{++} := \mathrm{int}(\mathbb{S}^N_+)$ be the cone of positive definite matrices in $\mathbb{S}^N$. For $X, Y \in \mathbb{S}^N$, the inequalities $X \geq Y$ and $X > Y$ mean that $X - Y \in \mathbb{S}^N_+$ and $X - Y \in \mathbb{S}^N_{++}$, respectively. The standard trace inner product is used for $\mathbb{S}^N$:

$$\langle X, Y \rangle = \mathrm{Tr}(XY) = \sum_{i=1}^{N} \sum_{j=1}^{N} X_{ij} Y_{ij}.$$

The set of $N \times N$ lower-triangular matrices with real entries is denoted by $\mathbb{T}^N$.

2.1. Sparse matrices

An $N \times N$ symmetric sparsity pattern is represented by a simple undirected graph $G = (V, E)$ with vertex set $V = \{1, 2, \ldots, N\}$ and edge set E. An edge connecting vertices i and j is denoted by $\{i, j\}$. A matrix $X \in \mathbb{S}^N$ is said to have sparsity pattern E if $X_{ij} = X_{ji} = 0$ whenever $i \neq j$ and $\{i, j\} \notin E$. The diagonal entries and the entries indexed by E are called the *non-zeros* in the pattern. The other entries (indexed by the complement of E) are the *zeros*. The set of symmetric $N \times N$ matrices with sparsity pattern E is denoted by $\mathbb{S}^N_E$:

$$\mathbb{S}^N_E := \{X \in \mathbb{S}^N : X_{ij} = X_{ji} = 0 \text{ if } i \neq j \text{ and } \{i, j\} \notin E\}.$$

We use Π_E to denote orthogonal projection on $\mathbb{S}^N_E$. For $X \in \mathbb{S}^N$, the matrix $\Pi_E(X)$ is the matrix in $\mathbb{S}^N_E$ with non-zero entries given by $(\Pi_E(X))_{ij} = X_{ij}$ if $i = j$ or if $i \neq j$ and $\{i, j\} \in E$.

The cone of positive semidefinite matrices in $\mathbb{S}^N_E$ is the intersection

$$\mathbb{S}^N_E \cap \mathbb{S}^N_+ = \left\{ X \in \mathbb{S}^N_E : X \geq 0 \right\}. \tag{2.1}$$

This cone is closed, convex and pointed. It also has non-empty interior relative to $\mathbb{S}^N_E$ (it includes the identity matrix I), so it is a *regular* (or *proper*) cone. The cone of matrices in $\mathbb{S}^N_E$ that have a positive semidefinite completion is the projection of $\mathbb{S}^N_+$ on $\mathbb{S}^N_E$. We denote this set by

$$\Pi_E(\mathbb{S}^N_+) = \{\Pi_E(Y) : Y \geq 0\}. \tag{2.2}$$

The cone $\Pi_E(\mathbb{S}^N_+)$ is clearly convex, pointed and has non-empty interior relative to $\mathbb{S}^N_E$. Closedness follows from the fact that if $\Pi_E(Y) = 0$ and $Y \geq 0$ then $Y = 0$. Hence the positive semidefinite completable cone $\Pi_E(\mathbb{S}^N_+)$ is also regular. The two cones $\mathbb{S}^N_E \cap \mathbb{S}^N_+$ and $\Pi_E(\mathbb{S}^N_+)$ are duals of each other under the trace inner product in the space $\mathbb{S}^N_E$.

The graph (V, E) can also be used to describe the sparsity pattern of lower-triangular matrices. We say $L \in \mathbb{T}^N$ has sparsity pattern E if $L + L^\top \in \mathbb{S}^N_E$. The

notation

$$\mathbb{T}_E^N = \left\{ L \in \mathbb{T}^N : L + L^\top \in \mathbb{S}_E^N \right\}$$

will be used for this set.

We define the Cholesky factorization of a positive definite matrix X as a decomposition

$$PXP^\top = LL^\top, \tag{2.3}$$

where P is a permutation matrix and L is lower-triangular with positive diagonal entries. In general, the factorization introduces fill in the sparsity pattern of $PXP^\top$. We say the sparsity pattern of L is an *extension* of the sparsity pattern of $PXP^\top$.

2.2. Chordal sparsity

We now give a short overview of the properties of chordal graphs and chordal sparsity patterns that will be important in the discussion of homogeneous chordal patterns in the next section. The interested reader is referred to the surveys by Vandenberghe and Andersen (2014), Blair and Peyton (1993), Golumbic (2004) and Zheng *et al.* (2021) for more background on chordal graphs and their history.

An undirected graph (V, E) is called *chordal* if it does not contain a cycle C_k of length $k \geq 4$ as a node-induced subgraph (from now on we will simply say *induced graph* to mean *node-induced graph*). A classical result states that a graph is chordal if and only if it has a *perfect elimination ordering* (Fulkerson and Gross 1965). An ordering of the graph is a bijection σ from $\{1, 2, \ldots, |V|\}$ to the vertex set V. An ordering σ is a perfect elimination ordering if

$$\left. \begin{array}{c} \{u, v\} \in E, \ \{u, w\} \in E, \\ \sigma^{-1}(u) < \sigma^{-1}(v) < \sigma^{-1}(w) \end{array} \right\} \quad \Longrightarrow \quad \{v, w\} \in E. \tag{2.4}$$

In other words, the *higher neighbourhood*

$$\mathrm{adj}^+(u) := \{v \in V : \{u, v\} \in E, \ \sigma^{-1}(u) < \sigma^{-1}(v)\}$$

of every vertex induces a complete subgraph of G:

$$v, w \in \mathrm{adj}^+(u) \quad \Longrightarrow \quad \{v, w\} \in E. \tag{2.5}$$

In sparse matrix language, a perfect elimination ordering of a sparsity pattern E defines a permutation matrix that yields a zero-fill Cholesky factorization (2.3), i.e. $P^\top(L + L^\top)P \in \mathbb{S}_E^N$ if $X \in \mathbb{S}_E^N$.

Efficient linear-time algorithms exist for testing chordality of a graph and finding a perfect elimination ordering if one exists (Rose, Tarjan and Lueker 1976, Tarjan and Yannakakis 1984). For non-chordal graphs, the connection with the sparse Cholesky factorization (2.3) suggests a practical heuristic for finding efficient chordal extensions: apply a fill-reducing reordering to the sparsity pattern of X and calculate the sparsity pattern of the Cholesky factor L.
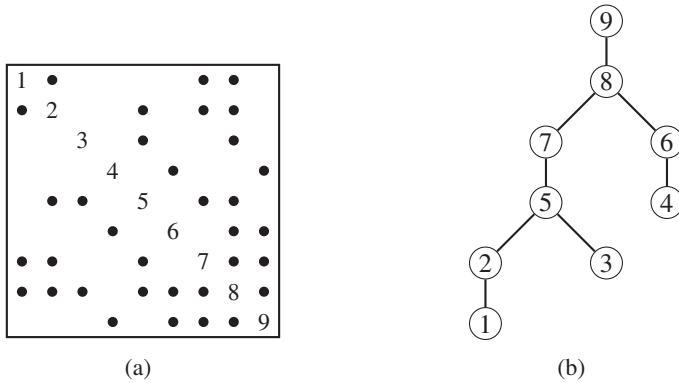
Figure 2.1. (a) A chordal graph with vertices $V = \{1, 2, \ldots, 9\}$ and perfect elimination ordering $1, \ldots, 9$. The dots in the array represent the edges in the graph. (b) Elimination tree.
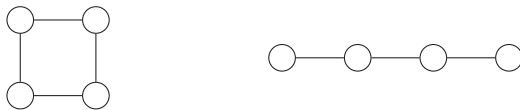

Figure 2.2. C_4 and P_4 are forbidden induced subgraphs in a homogeneous chordal graph.

Elimination trees play an important role in sparse matrix algorithms, such as the multifrontal algorithm for sparse Cholesky factorization (Duff and Reid 1983, Liu 1990). The elimination tree of a chordal graph G with perfect elimination ordering σ is a tree (or a forest if the graph is not connected), with vertex set V. The parent $p(u)$ of a non-root vertex u in the tree is the first element of $\mathrm{adj}^+(u)$. The perfect elimination property (2.5) holds if and only if

$$\mathrm{adj}^+(u) \subseteq \{p(u)\} \cup \mathrm{adj}^+(p(u)) \tag{2.6}$$

for all non-root vertices u. Figure 2.1 shows an example.

It is useful to note that the elimination tree provides a summary of the graph but is not an equivalent representation. For example, from the elimination tree in Figure 2.1 and the property (2.6), we can conclude that vertex 6 is not adjacent to vertex 1; however, the information in the elimination tree does not allow us to decide whether vertex 5 is adjacent to vertex 1.

2.3. Homogeneous chordal sparsity

We define a *homogeneous chordal graph* as an undirected graph that does not contain C_4 (a cycle of length four) or P_4 (a path formed by three edges on four vertices) as induced subgraphs. These forbidden subgraphs are shown in Figure 2.2.

It is clear from the definition that a homogeneous chordal graph does not contain any induced cycle C_k of length $k \geq 5$; so, all homogeneous chordal graphs are chordal.

Homogeneous chordal graphs were first studied by Wolk (1962, 1965), who called them *D-graphs*. Golumbic proposed the more commonly used term *trivially perfect* graphs (Golumbic 1978). They are known as *homogeneous graphs* in the statistics literature on Gaussian graphical models (Letac and Massam 2007, Khare and Rajaratnam 2012). Other names include *quasi-threshold graphs* (Yan, Chen and Chang 1996), *co-chordal graphs* (Khare and Rajaratnam 2012) and *chordal co-graphs*.[1] Our motivation for the name *homogeneous chordal graphs* will become clear in Section 3.

Wolk (1962, 1965) showed that the absence of P_4 and C_4 characterizes the comparability graphs of rooted forests: a graph $G = (V, E)$ is a homogeneous chordal graph if and only if there exists a rooted forest with vertex set V and such that $\{v, w\} \in E$ if and only if v is an ancestor of w or w is an ancestor of v in the forest (in which case we call v and w *comparable* vertices). As a key step in his proof, he also established the important property that every connected component of a homogeneous chordal graph has a *universal vertex*, i.e. a vertex adjacent to all other vertices in the same connected component (Wolk 1962, p. 18). This leads to a useful recursive characterization (Yan *et al.* 1996). Every homogeneous chordal graph can be constructed starting from a single-vertex graph by a repeated application of the following two operations.

- *Disjoint union.* If (V_1, E_1) and (V_2, E_2) are homogeneous chordal graphs and $V_1 \cap V_2 = \emptyset$, then $(V_1 \cup V_2, E_1 \cup E_2)$ is a homogeneous chordal graph.

- *Addition of a universal vertex.* If (V, E) is a homogeneous chordal graph and $w \notin V$, then $(V \cup \{w\}, E \cup \{\{w, v\} : v \in V\})$ is a homogeneous chordal graph.

These two operations have a simple interpretation for graphs that describe sparsity patterns. By making a disjoint union we construct a sparsity pattern of size $N_1 + N_2$ as a block-diagonal pattern with diagonal blocks of size N_1 and N_2 (up to a symmetric reordering). Adding a universal vertex to a sparsity pattern of size $N_1 \times N_1$ corresponds to adding a dense row and column to define a pattern of size $(N_1 + 1) \times (N_1 + 1)$. By repeating the two operations we construct a nested block-arrow pattern (up to a symmetric reordering). Figure 2.3 shows an example.

Chu (2008) presents a linear-time algorithm for recognizing homogeneous chordal graphs. The algorithm, described in detail in Appendix A, is an instance of the *lexicographic breadth-first search* (LBFS) algorithm that was first developed for testing chordality (Rose *et al.* 1976) and later extended for testing a variety of

---

[1] Graphs that do not contain P_4 are also known as *co-graphs* (complement reducible graphs), D^*-graphs or hereditary Dacey graphs (due to a connection to work on orthomodular lattices). So the homogeneous chordal graphs are the chordal co-graphs.
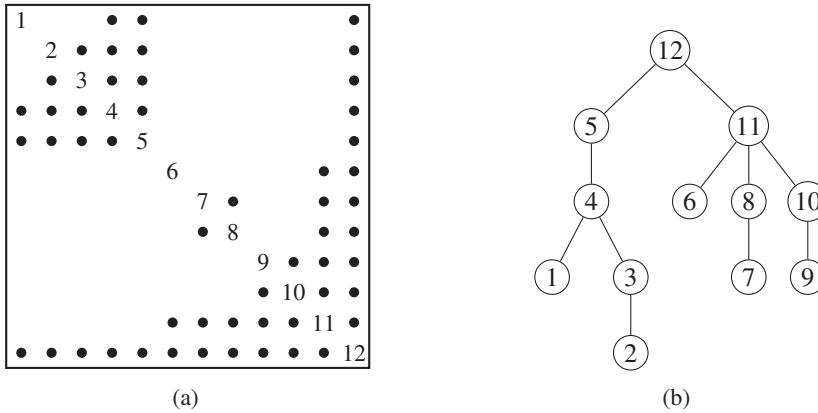
Figure 2.3. The homogeneous chordal graph in (a) is the comparability graph of the tree in (b). This tree is also the elimination tree for the perfect elimination ordering $1, \ldots, 12$.

other graph properties (Corneil 2004, Habib, McConnell, Paul and Viennot 2000). Chu's algorithm also produces a perfect elimination ordering and an elimination tree. The perfect elimination ordering σ produced by the LBFS algorithm has the following property, in addition to (2.4):

$$\left. \begin{array}{l} \{u, v\} \in E, \ \{v, w\} \in E, \\ \sigma^{-1}(u) < \sigma^{-1}(v) < \sigma^{-1}(w) \end{array} \right\} \quad \Longrightarrow \quad \{u, w\} \in E. \tag{2.7}$$

Combined with (2.4), this implies that two vertices are adjacent in the graph if and only if they form an ancestor–descendant pair in the elimination tree: the homogeneous chordal graph is the comparability graph of the elimination tree. We will call a perfect elimination ordering that satisfies (2.7) a *trivially perfect elimination ordering*. For a trivially perfect elimination ordering, property (2.6) can be strengthened to

$$\mathrm{adj}^+(u) = \{p(u)\} \cup \mathrm{adj}^+(p(u)). \tag{2.8}$$

Hence, in contrast to general chordal patterns, a homogeneous chordal graph is completely characterized by an elimination tree. This is illustrated in Figure 2.3. Here the numerical ordering is a trivially perfect elimination ordering of the homogeneous chordal graph in Figure 2.3(a). Each vertex in this graph is adjacent to all its ancestors and descendants in the elimination tree. The ordering in this example is also a *postordering*, that is, if $\sigma^{-1}(v) = j$ and v has k descendants in the elimination tree, then the descendants are numbered $j - 1, \ldots, j - k$. The postordering property holds for all trivially perfect elimination orderings computed by LBFS (see Appendix A).
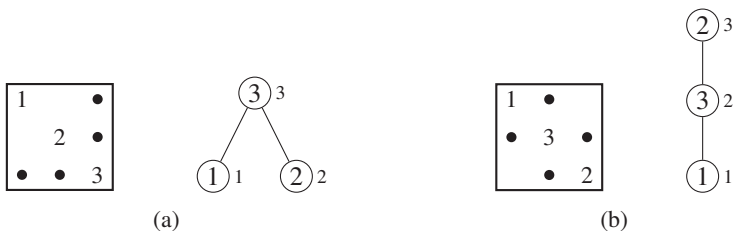
Figure 2.4. Two perfect elimination orderings of a homogeneous chordal graph and the corresponding elimination trees. The number next to node v in the elimination trees is $\sigma^{-1}(v)$, the position of v in the ordering. The ordering in (a) is a trivially perfect elimination ordering. The ordering in (b) is a perfect elimination ordering, but is not trivially perfect.

Note that not every perfect elimination ordering of a homogeneous chordal graph satisfies (2.8). Figure 2.4 shows the smallest non-trivial (not dense and not diagonal) sparsity pattern. The figure shows two perfect elimination orderings and the corresponding elimination trees. The first ordering is trivially perfect. The second ordering is not, because

$$\text{adj}^+(1) = \{3\} \neq \{p(1)\} \cup \text{adj}^+(p(1)) = \{2, 3\}.$$

The elimination tree for a trivially perfect elimination ordering can be compressed into a *supernodal* elimination tree, in which the nodes of the elimination tree are combined into larger supernodes. Each supernode is associated with a *representative* vertex. The representative vertices are the leaf nodes in the elimination tree and all the nodes with more than one child. The supernode with representative vertex v contains the representative vertex v itself plus the nodes in the elimination tree between v and the first ancestor w that is also a representative vertex. In the supernodal elimination tree, the supernode with representative vertex w is the parent of the supernode with representative v. The supernodes therefore form a partition of the vertex set. Each supernode induces a complete subgraph. The vertices in a supernode are adjacent to all vertices in the supernodes that are its ancestors or descendants in the supernodal elimination tree. The definitions are illustrated in Figure 2.5 for the example in Figure 2.3. Note that several other definitions of supernodes exist in the sparse matrix literature. The supernodes as defined here are known as *fundamental supernodes* (Liu, Ng and Peyton 1993).

To conclude, we summarize the properties of the example in Figures 2.3 and 2.5 that generalize to arbitrary homogeneous chordal sparsity patterns in $\mathbb{S}^N$. After applying a symmetric reordering one can assume that the numerical ordering $1, 2, \ldots, N$ is a trivially perfect elimination ordering and a postordering. A matrix
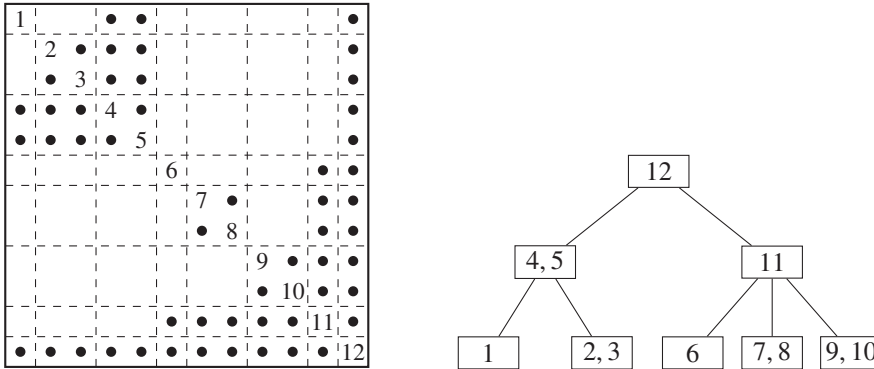
Figure 2.5. Fundamental supernode partition and supernodal elimination tree for the example in Figure 2.3. The representative vertices are 1, 2, 4, 6, 7, 9, 11, 12.

with a homogeneous chordal sparsity pattern will then have the form

$$
X = \begin{bmatrix}
X_{\beta_1\beta_1} & 0 & \cdots & 0 & X_{\beta_1\nu} \\
0 & X_{\beta_2\beta_2} & \cdots & 0 & X_{\beta_2\nu} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & X_{\beta_k\beta_k} & X_{\beta_k\nu} \\
X_{\nu\beta_1} & X_{\nu\beta_2} & \cdots & X_{\nu\beta_k} & X_{\nu\nu}
\end{bmatrix},
\tag{2.9}
$$

where each of the diagonal blocks $X_{\beta_i\beta_i}$, for $i = 1, \ldots, k$, has a block-arrow structure of the same form. If the sparsity pattern is not block-diagonal, ν is the supernode at the root of the supernodal elimination tree. Assume the root ν has k children, denoted by $\nu_1, \ldots, \nu_k$. Then the index set β_i is the union of the supernode ν_i and its descendants in the supernodal elimination tree. The postordering property implies that each of these index sets β_i contains consecutive indices, that precede the indices in ν, so the matrices $X_{\beta_i\beta_i}$ are diagonal blocks. Each of the matrices $X_{\beta_i\beta_i}$ has a homogeneous chordal sparsity pattern, with supernodal elimination tree given by the subtree rooted at ν_i. We have assumed that the entire sparsity pattern is not block-diagonal (ν is not empty). If it is block-diagonal, the associated sparsity graph is not connected, and the supernodal elimination tree is a forest with connected components $\beta_1, \ldots, \beta_k$.

It is easily verified that if the matrix (2.9) is positive definite, then its Cholesky factor in $X = LL^\top$ is structured as

$$
L = \begin{bmatrix}
L_{\beta_1\beta_1} & 0 & \cdots & 0 & 0 \\
0 & L_{\beta_2\beta_2} & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & L_{\beta_k\beta_k} & 0 \\
L_{\nu\beta_1} & L_{\nu\beta_2} & \cdots & L_{\nu\beta_k} & L_{\nu\nu}
\end{bmatrix},
\tag{2.10}
$$

where each block $L_{\beta_i\beta_i}$ is the Cholesky factor of $X_{\beta_i\beta_i}$ and therefore has a similar angular sparsity pattern.

2.4. Homogeneous chordal extension

Homogeneous chordal patterns in the reordered form (2.9) have a long history in many areas, including least-squares fitting (Björck 1996, §6.3, Golub and Plemmons 1980), decomposition methods in optimization (Lasdon 2002) and graphical statistical models (Pearl and Wermuth 1994, Chaudhuri, Drton and Richardson 2007, Drton and Richardson 2008, Letac and Massam 2007). The term *nested block-angularity* is used in Saunders (1972, p. 24).

They also arise naturally as *extensions* of general unstructured sparsity patterns, reordered using a nested dissection ordering (Duff, Erisman and Reid 2017, George and Liu 1981). Here ν is the vertex separator in the first dissection step; the other non-leaf supernodes are the separators in subsequent levels of dissection. Such a pattern is a homogeneous chordal pattern if at each level we treat the last block row and column in (2.9) as dense, and also treat the principal blocks indexed by the leaves of the supernodal elimination tree as dense. In applications to linear equations the matrix will have a large number of additional zeros within these blocks, so the actual sparsity pattern is an unstructured sparsity pattern E' (or a non-homogeneous chordal sparsity pattern if it is the filled pattern of a Cholesky factor) and the homogeneous chordal pattern E is an extension ($E' \subseteq E$).

When used in the non-symmetric formulation (1.3) of a sparse semidefinite program, a homogeneous chordal extension can be obtained by applying nested dissection to the aggregate sparsity pattern of $A_1, \ldots, A_m, B$. If the homogeneous chordal extension is used to define $\mathcal{V}$, then, as we will see in the next section, the cone K is a homogeneous convex cone. The coefficient matrices $A_1, \ldots, A_m, B$ are sparse matrices in $\mathcal{V}$, but their zeros within the homogeneous chordal pattern are not exploited in the definition of the cone K.

Nested dissection ordering provides a heuristic for obtaining homogeneous chordal extensions, with no guarantee of optimality. As proved by Yannakakis (1981), given a sparsity pattern, it is NP-hard to compute the minimum number of edges to add to make the underlying graph chordal. Analogously, El-Mallah and Colbourn (1988) proved that given a sparsity pattern, it is NP-hard to find the smallest number of edges to add to the graph to make it a co-graph (a graph that does not contain P_4 as an induced subgraph). We can show that given a sparsity pattern, it is NP-hard to find the largest induced subgraph which is homogeneous chordal.

Proposition 2.1. Given a graph $G = (V, E)$ describing the sparsity pattern of a symmetric matrix, it is NP-hard to compute the largest principal submatrix with a homogeneous chordal sparsity pattern.

Proof. We use Theorem 3 of Bartholdi (1981/82) (whose proof relies on Yannakakis's related results). This theorem states that given a square matrix A with 0,1

entries, and a positive integer k, it is NP-hard to decide whether A has a $k \times k$ principal submatrix satisfying property $\mathcal{P}$, provided

- property $\mathcal{P}$ is *non-trivial* (meaning that it holds for infinitely many 0,1 matrices and it fails for infinitely many 0,1 matrices);
- property $\mathcal{P}$ holds for identity matrices;
- property $\mathcal{P}$ is hereditary on principal submatrices.

Thus it suffices for us to check that the property of homogeneous chordal sparsity satisfies these required conditions. Using the excluded induced subgraph characterization of homogeneous chordal graphs, we note that identity matrices correspond to empty (no edges) graphs which are homogeneous chordal; sparsity patterns of principal submatrices correspond to induced subgraphs, and if the original graph does not contain a C_4 or P_4 then neither do any of its induced subgraphs. Finally, there are infinitely many graphs which do not contain a C_4 or P_4; moreover, there are infinitely many graphs which do contain either a C_4 or a P_4 (possibly both and many copies). Thus homogeneous chordal sparsity satisfies the assumptions of Theorem 3 of Bartholdi (1981/82) and the underlying problem is NP-hard. □

Therefore one must rely on heuristic algorithms in general (including polynomial-time approximation algorithms for the minimum fill-in problems: Natanzon, Shamir and Sharan 2000), as in the approaches used in applications of chordal extensions of sparsity patterns.

3. Homogeneous sparse matrix cones

We now apply the results of the previous section to derive properties of the two matrix cones

$$K := \mathbb{S}_E^N \cap \mathbb{S}_+^N, \quad K^* = \Pi_E(\mathbb{S}_+^N). \tag{3.1}$$

The cone K is the cone of positive semidefinite matrices with sparsity pattern E. The dual cone K^* is the cone of positive semidefinite completable matrices with sparsity pattern E. Note that $K \subseteq K^*$. We assume that E is a homogeneous chordal sparsity pattern and that the numerical order $1, \ldots, N$ is a trivially perfect elimination ordering, as in the example of Figure 2.3.

The *automorphism group* $\mathrm{Aut}(K)$ of a regular cone K is the set of non-singular linear transformations that map K to itself. A regular cone K is called *homogeneous* if, for every pair of points $x, y \in \mathrm{int}(K)$, there exists an automorphism of K that maps x to y. So, a regular cone K is homogeneous if and only if the automorphism group of K acts transitively in the interior of K. A subset $\mathcal{H} \subseteq \mathrm{Aut}(K)$ is a *transitive subset* of $\mathrm{Aut}(K)$ if, for every pair of points $x, y \in \mathrm{int}(K)$, there exists an automorphism in $\mathcal{H}$ that maps x to y.

Ishi (2013, Theorem A) proves that the sparse matrix cones (3.1) are homogeneous if and only if E is a homogeneous chordal sparsity pattern. In this section we describe transitive subsets of the primal and dual automorphism groups.

3.1. Computations with sparse triangular matrices

The properties of homogeneous chordal sparse matrices that will be needed follow from four facts presented in the next theorem.

Theorem 3.1. Let E be a homogeneous chordal sparsity pattern in $\mathbb{S}^N$, with trivially perfect elimination ordering $1, \ldots, N$, and assume $L \in \mathbb{T}_E^N$.

(i) If $\tilde{L} \in \mathbb{T}_E^N$, then $L\tilde{L} \in \mathbb{T}_E^N$.

(ii) If L is non-singular, then $L^{-1} \in \mathbb{T}_E^N$.

(iii) If $X \in \mathbb{S}_E^N$, then $LXL^\top \in \mathbb{S}_E^N$.

(iv) If $Y \in \mathbb{S}^N$, then $\Pi_E(L^\top Y L) = \Pi_E(L^\top \Pi_E(Y) L)$.

The second property appears in Khare and Rajaratnam (2012). None of the four properties holds for general chordal sparsity patterns, as can be seen by considering the example of a tridiagonal pattern, which is chordal but not homogeneous if $N \geq 4$. We also note the assumption of a trivially perfect elimination ordering. In the example in Figure 2.4(b), the ordering $\sigma(1) = 1, \sigma(2) = 3, \sigma(3) = 2$ is a perfect elimination ordering and results in a zero-fill bidiagonal Cholesky factor. However, the inverse Cholesky factor will generally have a non-zero entry in position 2, 1.

Proof. To simplify the notation, we denote the set $\text{adj}^+(i)$ by α_i. This is the set of row indices of the lower-triangular non-zeros in column i. The set $\{i\} \cup \alpha_i$ is denoted by $\bar{\alpha}_i$. If the order of the elements in α_i and $\bar{\alpha}_i$ matters, it is assumed that they are sorted in increasing order. In this notation, the property (2.8) can be expressed as

$$\alpha_i = \bar{\alpha}_{p(i)} \quad \text{for all } i, \tag{3.2}$$

where we interpret $\bar{\alpha}_{p(i)}$ as the empty set if i is a root of the elimination tree. In the example of Figure 2.3, $\alpha_3 = \{4, 5, 12\}$, $\bar{\alpha}_3 = \{3, 4, 5, 12\}$ and $p(3) = 4$.

To prove property (i), we examine the sparsity pattern of $L\tilde{L}$. The ij element, with $i \geq j$, is

$$(L\tilde{L})_{ij} = \sum_{k=j}^N L_{ik}\tilde{L}_{kj} = \sum_{k \in \bar{\alpha}_j} L_{ik}\tilde{L}_{kj}.$$

The simplification in the second expression follows because $\tilde{L}_{kj} = 0$ for $k \notin \bar{\alpha}_j$. Since L_{ik} is zero if $i \notin \bar{\alpha}_k$, we have $(L\tilde{L})_{ij} = 0$ for $i \notin \bigcup_{k \in \bar{\alpha}_j} \bar{\alpha}_k$. It follows from (3.2) that $\bigcup_{k \in \bar{\alpha}_j} \bar{\alpha}_k = \bar{\alpha}_j$. We conclude that the non-zeros of column j of $L\tilde{L}$ are in the positions indexed by $\bar{\alpha}_j$, i.e. $L\tilde{L} \in \mathbb{T}_E^N$.

For property (ii) we consider the forward substitution method for computing column k of L^{-1}. To solve $Lx = e_k$, where k is the kth unit vector, we set $x = e_k$ and run the iteration

$$\begin{bmatrix} x_j \\ x_{\alpha_j} \end{bmatrix} := \begin{bmatrix} 1/L_{jj} & 0 \\ -L_{\alpha_j j}/L_{jj} & I \end{bmatrix} \begin{bmatrix} x_j \\ x_{\alpha_j} \end{bmatrix}, \quad j = k, k+1, \ldots, N.$$

Since initially $x = e_k$, and α_j is the set of ancestors of vertex j in the elimination tree, the iteration only modifies entries of x on the path between k and the root of the tree. In other words, the iteration can be simplified as

$$\begin{bmatrix} x_j \\ x_{\alpha_j} \end{bmatrix} := \begin{bmatrix} 1/L_{jj} & 0 \\ -L_{\alpha_j j}/L_{jj} & I \end{bmatrix} \begin{bmatrix} x_j \\ x_{\alpha_j} \end{bmatrix}, \quad j = k, p(k), p^2(k), \dots,$$

where $p^2(k) = p(p(k))$, etc., that is, we iterate over $j \in \bar{\alpha}_k$ in ascending order. After completing the iteration, the non-zeros of x are in the positions indexed by $\bar{\alpha}_k$. Therefore $L^{-1} \in \mathbb{T}_E^N$.

Next we prove property (iii). Consider the following expression for the lower-triangular entry of $LXL^\top$ in position ij, with $i > j$:

$$(LXL^\top)_{ij} = \sum_{k=1}^N \left(L_{ik} L_{kj} X_{kk} + \sum_{l \in \alpha_k} X_{lk}(L_{il}L_{jk} + L_{ik}L_{jl}) \right). \tag{3.3}$$

Suppose $i \notin \alpha_j$, that is, i is not an ancestor of j in the elimination tree. We show that $(LXL^\top)_{ij} = 0$. The first term in the sum (3.3) is zero because $L_{ik}L_{kj} \neq 0$ only if $i \in \bar{\alpha}_k$ and $k \in \bar{\alpha}_j$, which implies i is on the path from vertex j to the root. The second term is zero because $L_{il}L_{jk} \neq 0$ implies $i \in \bar{\alpha}_l \subset \bar{\alpha}_k$ and $j \in \bar{\alpha}_k$, so i and j are both on the path from vertex k to the root, and since $i > j$, vertex i is an ancestor of j. Similarly, the last term is zero because $L_{ik}L_{jl} \neq 0$ implies that $i \in \bar{\alpha}_k$ and $j \in \bar{\alpha}_l \subset \alpha_k$, so i and j are both on the path from vertex k to the root and i is an ancestor of j.

The last property in the list follows from property (iii). It is sufficient to show that $\Pi_E(LYL^\top) = 0$ whenever $\Pi_E(Y) = 0$. To see this, we choose any $X \in \mathbb{S}_E^N$ and note that

$$\mathrm{Tr}(X\Pi_E(L^\top YL)) = \mathrm{Tr}(XL^\top YL) = \mathrm{Tr}(LXL^\top Y) = 0$$

because $LXL^\top \in \mathbb{S}_E^N$ by property (iii) and $\Pi_E(Y) = 0$. $\qquad \square$

The properties of Theorem 3.1 are also easily verified by induction for a pattern in the postordered block-matrix form (2.9). To verify property (ii), we note that if L in (2.10) is invertible, its inverse is

$$L^{-1} = \begin{bmatrix} L_{\beta_1\beta_1}^{-1} & 0 & \cdots & 0 & 0 \\ 0 & L_{\beta_2\beta_2}^{-1} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & L_{\beta_k\beta_k}^{-1} & 0 \\ -L_{\nu\nu}^{-1}L_{\nu\beta_1}L_{\beta_1\beta_1}^{-1} & -L_{\nu\nu}^{-1}L_{\nu\beta_2}L_{\beta_2\beta_2}^{-1} & \cdots & -L_{\nu\nu}^{-1}L_{\nu\beta_k}L_{\beta_k\beta_k}^{-1} & L_{\nu\nu}^{-1} \end{bmatrix},$$

and it is clear that L^{-1} has the same sparsity pattern as L.

3.2. Primal cone automorphisms

We now show that the linear transformations of the form

$$\mathcal{L}(X) = LXL^\top, \tag{3.4}$$

with non-singular $L \in \mathbb{T}_E^N$, form a transitive subset of $\mathrm{Aut}(K)$. Property (iii) of Theorem 3.1 shows that $\mathcal{L}(X) \in \mathbb{S}_E^N$ for $X \in \mathbb{S}_E^N$. Since $L^{-1} \in \mathbb{T}_E^N$ (by property (ii)), the same is true for the inverse mapping $\mathcal{L}^{-1}(X) = L^{-1}XL^{-\top}$. The two transformations $\mathcal{L}$ and $\mathcal{L}^{-1}$ preserve positive definiteness, so they are automorphisms for K. To show that the transformations $\mathcal{L}$ form a transitive subset, we show that for every pair of matrices $X_1, X_2 \in \mathrm{int}(K)$ there exists a non-singular $L \in \mathbb{T}_E^N$ such that $LX_1L^\top = X_2$. Let $L_1, L_2 \in \mathbb{T}_E^N$ be the triangular factors in the Cholesky factorizations $X_1 = L_1L_1^\top$ and $X_2 = L_2L_2^\top$. The matrix $L = L_2L_1^{-1}$ is non-singular and in $\mathbb{T}_E^N$ (by the first two properties of Theorem 3.1). The automorphism $\mathcal{L}$ defined by L maps X_1 to X_2:

$$\mathcal{L}(X_1) = LX_1L^\top = LL_1L_1^\top L^\top = L_2L_2^\top = X_2.$$

We will use the notation $\mathrm{Aut}_\triangle(K)$ for the transitive subset of $\mathrm{Aut}(K)$ containing the transformations of the form (3.4) with non-singular $L \in \mathbb{T}_E^N$.

3.3. Dual cone automorphisms

The adjoint of $\mathcal{L}$ is the linear mapping from $\mathbb{S}_E^N$ to $\mathbb{S}_E^N$ that satisfies $\langle \mathcal{L}^*(S), X \rangle = \langle S, \mathcal{L}(X) \rangle$ for all $S, X \in \mathbb{S}_E^N$. Since we use the trace inner product,

$$\langle S, \mathcal{L}(X) \rangle = \mathrm{Tr}(SLXL^\top) = \mathrm{Tr}(L^\top SLX) = \langle \Pi_E(L^\top SL), X \rangle,$$

so the adjoint is given by

$$\mathcal{L}^*(S) = \Pi_E(L^\top SL). \tag{3.5}$$

The projection in the expression $\Pi_E(L^\top SL)$ cannot be omitted because, unlike for the forward mapping $LXL^\top$, the product $L^\top SL$ is not necessarily in $\mathbb{S}_E^N$.

The linear transformations of the form $\mathcal{L}^*$, where $\mathcal{L} \in \mathrm{Aut}_\triangle(K)$, form a transitive subset of $\mathrm{Aut}(K^*)$. The fact that $\mathcal{L}^*$ is an automorphism of K^* follows directly from being the adjoint of an automorphism of K:

$$\begin{aligned}
S \in K^* &\iff \langle S, X \rangle \geq 0 \quad \text{for all } X \in K \\
&\iff \langle S, \mathcal{L}(X) \rangle \geq 0 \quad \text{for all } X \in K \\
&\iff \langle \mathcal{L}^*(S), X \rangle \geq 0 \quad \text{for all } X \in K \\
&\iff \mathcal{L}^*(S) \in K^*.
\end{aligned}$$

On line 2 we use the fact that $\mathcal{L}$ is an automorphism of K. Next we prove that the mappings $\mathcal{L}^*$ form a transitive subset of $\mathrm{Aut}(K^*)$, by showing how for every $S_1, S_2 \in \mathrm{int}(K^*)$ one can find L such that $\mathcal{L}^*(S_1) = S_2$. We use a classical result from the theory of positive definite matrix completions, stating that for every $S \in \mathrm{int}(K^*)$

there exists an $X \in \text{int}(K)$ that satisfies $\Pi_E(X^{-1}) = S$ (Grone *et al.* 1984). The matrix X is the inverse of the maximum-determinant positive definite completion, i.e. the unique solution Y of the convex optimization problem

$$
\begin{array}{ll}
\text{minimize} & -\ln \det(Y) \\
\text{subject to} & \Pi_E(Y) = S
\end{array}
\tag{3.6}
$$

over $Y \in \mathbb{S}_{++}^N$. The optimality conditions for this problem,

$$
Y^{-1} = X > 0, \quad \Pi_E(Y) = S,
$$

where $X \in \mathbb{S}_E^N$ is a multiplier for the equality constraint of (3.6), show that $\Pi_E(X^{-1}) = S$. Now consider two matrices $S_1, S_2 \in \text{int}(K^*)$. To construct an automorphism $\mathcal{L}^*$ (of K^*) that maps S_1 to S_2, we compute the matrices $X_1, X_2 \in \text{int}(K)$ that satisfy $\Pi_E(X_1^{-1}) = S_1, \Pi_E(X_2^{-1}) = S_2$. Let $L_1, L_2 \in \mathbb{T}_E^N$ be the Cholesky factors of X_1 and X_2, and define $L = L_1 L_2^{-1}$. Then

$$
\begin{aligned}
\mathcal{L}^*(S_1) &= \Pi_E(L^\top S_1 L) \\
&= \Pi_E(L^\top \Pi_E(L_1^{-\top} L_1^{-1}) L) \\
&= \Pi_E(L^\top L_1^{-\top} L_1^{-1} L) \\
&= \Pi_E(L_2^{-\top} L_2^{-1}) \\
&= S_2.
\end{aligned}
$$

On line 3 we apply property (iv) of Theorem 3.1.

3.4. Matrix inverse

The inverse of a positive definite matrix $X \in \text{int}(K)$ can be factorized as $X^{-1} = RR^\top$, where the upper-triangular matrix $R = L^{-\top}$ is sparse and satisfies $R^\top \in \mathbb{T}_E^N$. Suppose the pattern is in the postordered block-matrix form (2.9). Then

$$
R = \begin{bmatrix}
R_{\beta_1\beta_1} & 0 & \cdots & 0 & R_{\beta_1 \nu} \\
0 & R_{\beta_2\beta_2} & \cdots & 0 & R_{\beta_2 \nu} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & R_{\beta_k\beta_k} & R_{\beta_k \nu} \\
0 & 0 & \cdots & 0 & R_{\nu\nu}
\end{bmatrix}
$$

$$
= \begin{bmatrix}
L_{\beta_1\beta_1}^{-\top} & 0 & \cdots & 0 & -L_{\beta_1\beta_1}^{-\top} L_{\nu\beta_1}^{\top} L_{\nu\nu}^{-\top} \\
0 & L_{\beta_2\beta_2}^{-\top} & \cdots & 0 & -L_{\beta_2\beta_2}^{-\top} L_{\nu\beta_2}^{\top} L_{\nu\nu}^{-\top} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & L_{\beta_k\beta_k}^{-\top} & -L_{\beta_k\beta_k}^{-\top} L_{\nu\beta_k}^{\top} L_{\nu\nu}^{-\top} \\
0 & 0 & \cdots & 0 & L_{\nu\nu}^{-\top}
\end{bmatrix},
$$

and X^{-1} is the sum of a block-diagonal and a low-rank matrix

$$X^{-1} = \begin{bmatrix} X^{-1}_{\beta_1\beta_1} & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & X^{-1}_{\beta_k\beta_k} & 0 \\ 0 & \cdots & 0 & 0 \end{bmatrix} + \begin{bmatrix} R_{\beta_1\nu} \\ \vdots \\ R_{\beta_k\nu} \\ R_{\nu\nu} \end{bmatrix} \begin{bmatrix} R_{\beta_1\nu} \\ \vdots \\ R_{\beta_k\nu} \\ R_{\nu\nu} \end{bmatrix}^\top .$$

Moreover, each diagonal block $X^{-1}_{\beta_i\beta_i}$ has a similar block-diagonal plus low-rank structure.

Conversely, consider a block-diagonal plus low-rank matrix

$$Y = \begin{bmatrix} Y_{\beta_1\beta_1} & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & Y_{\beta_k\beta_k} & 0 \\ 0 & \cdots & 0 & 0 \end{bmatrix} + \begin{bmatrix} W_{\beta_1\nu} \\ \vdots \\ W_{\beta_k\nu} \\ W_{\nu\nu} \end{bmatrix} \begin{bmatrix} W_{\beta_1\nu} \\ \vdots \\ W_{\beta_k\nu} \\ W_{\nu\nu} \end{bmatrix}^\top ,$$

where the matrices $Y_{\beta_1\beta_1}, \ldots, Y_{\beta_k\beta_k}$ are positive definite, and $W_{\nu\nu}$ is invertible. Then the inverse is a block-arrow matrix

$$Y^{-1} = \begin{bmatrix} Y^{-1}_{\beta_1\beta_1} & \cdots & 0 & -Y^{-1}_{\beta_1\beta_1} W_{\beta_1\nu} W^{-1}_{\nu\nu} \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & Y^{-1}_{\beta_k\beta_k} & -Y^{-1}_{\beta_k\beta_k} W_{\beta_k\nu} W^{-1}_{\nu\nu} \\ -W^{-1}_{\nu\nu} W^\top_{\beta_1\nu} Y^{-1}_{\beta_1\beta_1} & \cdots & -W^{-1}_{\nu\nu} W^\top_{\beta_k\nu} Y^{-1}_{\beta_k\beta_k} & W^{-\top}_{\nu\nu} S W^{-1}_{\nu\nu} \end{bmatrix},$$

where

$$S = I + \sum_{i=1}^{k} W^\top_{\beta_i\nu} Y^{-1}_{\beta_i\beta_i} W_{\beta_i\nu}.$$

4. Logarithmic barriers

The function $-\ln\det(X)$ for symmetric positive definite X has important applications in statistics, machine learning, information theory and semidefinite optimization. Here we restrict the function to the symmetric matrices with a given homogeneous chordal sparsity pattern E. We denote this function by $F \colon \mathbb{S}^N_E \to (-\infty, +\infty]$,

$$F(X) := \begin{cases} -\ln\det(X) & \text{if } X \in \operatorname{int}(K), \\ +\infty & \text{otherwise,} \end{cases} \tag{4.1}$$

where K is the primal cone in (3.1), and refer to F as the *logarithmic barrier* for K.

The gradient and Hessian of F (as a function on $\mathbb{S}^N_E$) at $X \in \operatorname{int}(K)$ are given by

$$F'(X) = -\Pi_E(X^{-1}), \quad F''(X; Y) = \Pi_E(X^{-1} Y X^{-1}). \tag{4.2}$$

Here $F''(X; Y)$ denotes the directional derivative of F' at X in the direction $Y \in \mathbb{S}^N_E$,

that is,

$$F''(X;Y) = \frac{\mathrm{d}}{\mathrm{d}\alpha}F'(X + \alpha Y)\Big|_{\alpha=0}.$$

The *conjugate barrier* of F is defined as

$$F_*(S) = \sup_{X \in \mathrm{int}(K)} \{-\langle S, X \rangle - F(X)\}$$

and has domain $\mathrm{int}(K^*)$. This is the logarithmic barrier for K^*. The maximizer in the optimization problem in the definition is the positive definite solution $\hat{X}$ of the non-linear equation

$$F'(X) = -\Pi_E(X^{-1}) = -S,$$

with variable $X \in \mathbb{S}_E^N$. The inverse $\hat{X}^{-1}$ of the solution is the maximum-determinant positive definite completion of S. From $\hat{X}$ we obtain the function value $F_*(S) = -F(\hat{X}) - N$ and the derivatives

$$F'_*(S) = -\hat{X}, \quad F''_*(S) = F''(\hat{X})^{-1}. \tag{4.3}$$

In this section we derive some interesting properties of compositions of F and F_* with the cone automorphisms (3.4) and (3.5), respectively.

4.1. Composition with primal cone automorphism

As in Section 3, we assume that the numerical order is a trivially perfect elimination ordering for E. Clearly,

$$F(\mathcal{L}(X)) = F(LXL^\top) = F(X) + F(LL^\top) \tag{4.4}$$

for all $X \in \mathrm{int}(K)$ and non-singular $L \in \mathbb{T}_E$. Differentiating the left- and right-hand sides with respect to X shows that

$$F'(\mathcal{L}(X)) = \mathcal{L}^{-*}(F'(X)), \quad F''(\mathcal{L}(X)) = \mathcal{L}^{-*} \circ F''(X) \circ \mathcal{L}^{-1} \tag{4.5}$$

for all $X \in \mathrm{int}(K)$ and non-singular $L \in \mathbb{T}_E$. These properties can also be verified from the definitions (4.2) and Theorem 3.1. For the gradient,

$$
\begin{aligned}
F'(\mathcal{L}(X)) &= -\Pi_E((LXL^\top)^{-1}) \\
&= -\Pi_E(L^{-\top}X^{-1}L^{-1}) \\
&= -\Pi_E(L^{-\top}\Pi_E(X^{-1})L^{-1}) \\
&= \mathcal{L}^{-*}(F'(X)).
\end{aligned}
$$

On line 3 we use property (iv) of Theorem 3.1. The result for the Hessian follows similarly from

$$
\begin{aligned}
F''(\mathcal{L}(X);Y) &= \Pi_E((LXL^\top)^{-1}Y(LXL^\top)^{-1}) \\
&= \Pi_E(L^{-\top}X^{-1}L^{-1}YL^{-\top}X^{-1}L^{-1})
\end{aligned}
$$

$$\begin{aligned}
&= \Pi_E(L^{-\top}\Pi_E(X^{-1}L^{-1}YL^{-\top}X^{-1})L^{-1})\\
&= \mathcal{L}^{-*}(F''(X;\mathcal{L}^{-1}(Y)))
\end{aligned}$$

for every $Y \in \mathbb{S}_E$.

4.2. Composition with dual cone automorphism

Similar properties hold for the dual barrier. Using (4.4) in the definition of the dual barrier, we find that

$$\begin{aligned}
F_*(S) &= \sup_X\{\langle -S, X\rangle - F(X)\}\\
&= \sup_X\{\langle -S, \mathcal{L}(X)\rangle - F(\mathcal{L}(X))\}\\
&= \sup_X\{\langle -\mathcal{L}^*(S), X\rangle - F(X)\} - F(LL^\top)\\
&= F_*(\mathcal{L}^*(S)) - F(LL^\top).
\end{aligned}$$

Hence $F_*(\mathcal{L}^*(S)) = F_*(S) + F(LL^\top)$ for all $S \in \mathrm{int}(K^*)$ and non-singular $L \in \mathbb{T}_E$. Differentiating with respect to S shows that

$$F_*'(\mathcal{L}^*(S)) = \mathcal{L}^{-1}(F_*'(S)), \quad F_*''(\mathcal{L}^*(S)) = \mathcal{L}^{-1}\circ F_*''(S)\circ\mathcal{L}^{-*}. \qquad (4.6)$$

To verify these properties directly, we note that, by definition,

$$\begin{aligned}
\hat{X} = -F_*'(S) &\iff \Pi_E(\hat{X}^{-1}) = S,\\
\hat{Y} = -F_*'(\mathcal{L}^*(S)) &\iff \Pi_E(\hat{Y}^{-1}) = \Pi_E(L^\top SL).
\end{aligned}$$

Combining the two properties, we obtain

$$\Pi_E(\hat{Y}^{-1}) = \Pi_E(L^\top\Pi_E(\hat{X}^{-1})L) = \Pi_E(L^\top\hat{X}^{-1}L).$$

Since the maximum-determinant positive definite completion is unique, we conclude that

$$-F_*'(\mathcal{L}^*(S)) = \hat{Y} = L^{-1}\hat{X}L^{-\top} = -\mathcal{L}^{-1}(F_*'(S)).$$

The Hessian property in (4.6) follows from

$$\begin{aligned}
F_*''(\mathcal{L}^*(S)) &= F''(\hat{Y})^{-1}\\
&= F''(\mathcal{L}^{-1}(\hat{X}))^{-1}\\
&= (\mathcal{L}^*\circ F''(\hat{X})\circ\mathcal{L})^{-1}\\
&= \mathcal{L}^{-1}\circ F''(\hat{X})^{-1}\circ\mathcal{L}^{-*}\\
&= \mathcal{L}^{-1}\circ F_*''(S)\circ\mathcal{L}^{-*}.
\end{aligned}$$

4.3. Hessian factorization

An important consequence of the second relation in (4.5) is that the Hessian of F at any point $X \in \text{int}(K)$ can be factored as

$$F''(X) = \mathcal{L}^{-*} \circ \mathcal{L}^{-1}, \tag{4.7}$$

where $\mathcal{L} \in \text{Aut}_\triangle(K)$, namely the automorphism that maps the identity matrix I to $\mathcal{L}(I) = X$ (and defined by the Cholesky factor of X). Similarly, from (4.6), the Hessian of F_* at any point $S \in \text{int}(K^*)$ admits a factorization

$$F_*''(S) = \mathcal{L} \circ \mathcal{L}^*,$$

where $\mathcal{L}^*$ is the dual cone automorphism that maps S to $\mathcal{L}^*(S) = I$.

Nesterov and Todd (1997) (see also Tunçel 2001, Theorem 3.1) have shown that for every $X \in \text{int}(K)$ and $S \in \text{int}(K^*)$ there exists a unique $W \in \text{int}(K)$ that satisfies

$$F''(W; X) = S,$$

where $F''(W; X)$ denotes the directional derivative of F' at W in the direction X. The matrix W is the solution of the convex optimization problem

$$\text{minimize} \quad -\langle F'(W), X \rangle + \langle S, W \rangle$$

with variable W. By factorizing $F''(W)$ as $F''(W) = \mathcal{L}^{-*} \circ \mathcal{L}^{-1}$, we obtain the following theorem.

Theorem 4.1. For every pair of interior points $X \in \text{int}(K)$ and $S \in \text{int}(K^*)$, there exists a unique $\mathcal{L} \in \text{Aut}_\triangle(K)$ which satisfies

$$\mathcal{L}^{-1}(X) = \mathcal{L}^*(S),$$

that is, there exists a non-singular $L \in \mathbb{T}_E$ such that $L^{-1} X L^{-\top} = \Pi_E(L^\top S L)$.

Theorem 4.1 can be generalized to all homogeneous cones (see the discussion following Theorem 5.3). Efficient computation of the matrix W is a topic of current research.

A closely related result on convex cones is discussed in Tunçel (1998). Theorem 4.2 of Tunçel (1998) states that if there exists a subset $G \subseteq \text{Aut}(K)$ such that for every $x \in \text{int}(K)$ and $s \in \text{int}(K^*)$ there exists a self-adjoint $\mathcal{D} \in G$ which satisfies

$$\mathcal{D}^{-1}(x) = \mathcal{D}(s),$$

then K must be a symmetric cone (homogeneous and self-dual). Theorem 4.1 does not contradict Theorem 4.2 of Tunçel (1998) because the automorphism $\mathcal{L}$ in Theorem 4.1 is not self-adjoint.

5. Homogeneous matrix cones

As an extension of (3.1) we now consider slices of the positive semidefinite cone

$$K := \mathcal{V} \cap \mathbb{S}_+^N, \tag{5.1}$$

where $\mathcal{V}$ is a subspace of $\mathbb{S}^N$. It is clear that K is a closed, pointed and convex cone. We will assume that $\mathcal{V} \cap \mathbb{S}_{++}^N$ is non-empty, so K has non-empty interior (relative to $\mathcal{V}$). The corresponding dual cone (in the subspace $\mathcal{V}$) is given by

$$K^* = \Pi_{\mathcal{V}}(\mathbb{S}_+^N), \tag{5.2}$$

where $\Pi_{\mathcal{V}}$ denotes Euclidean projection on $\mathcal{V}$. To see this, we first note that the cone $\Pi_{\mathcal{V}}(\mathbb{S}_+^N)$ is closed. This follows from Rockafellar (1970, Theorem 9.1) and the fact that if $\Pi_{\mathcal{V}}(Y) = 0$ and $Y \geq 0$ then Y must be zero, because $\Pi_{\mathcal{V}}(Y) = 0$ implies that $\mathrm{Tr}(YX) = 0$ for all $X \in \mathcal{V}$ and, by assumption, $\mathcal{V}$ contains positive definite matrices. Next, it is easily verified that the dual of the cone $\Pi_{\mathcal{V}}(\mathbb{S}_+^N)$ is given by the cone K defined in (5.1):

$$\begin{aligned}
(\Pi_{\mathcal{V}}(\mathbb{S}_+^N))^* &= \{X \in \mathcal{V} \colon \mathrm{Tr}(SX) \geq 0 \text{ for all } S \in \Pi_{\mathcal{V}}(\mathbb{S}_+^N)\} \\
&= \{X \in \mathcal{V} \colon \mathrm{Tr}(YX) \geq 0 \text{ for all } Y \in \mathbb{S}_+^N\} \\
&= \mathcal{V} \cap \mathbb{S}_+^N.
\end{aligned}$$

Hence $K = (\Pi_{\mathcal{V}}(\mathbb{S}_+^N))^*$. Since $\Pi_{\mathcal{V}}(\mathbb{S}_+^N)$ is closed, we have

$$K^* = (\Pi_{\mathcal{V}}(\mathbb{S}_+^N))^{**} = \Pi_{\mathcal{V}}(\mathbb{S}_+^N).$$

We conclude that K and K^* form a dual pair of regular cones. We also note that $K \subseteq K^*$.

Ishi (2015) presents conditions on $\mathcal{V}$ that imply that the cone K defined in (5.1) is homogeneous. Suppose that after a suitable reordering, the matrices $X \in \mathcal{V}$ can be partitioned as $r \times r$ block matrices

$$X = \begin{bmatrix}
X_{11} & X_{21}^\top & X_{31}^\top & \cdots & X_{r1}^\top \\
X_{21} & X_{22} & X_{32}^\top & \cdots & X_{r2}^\top \\
X_{31} & X_{32} & X_{33} & \cdots & X_{r3}^\top \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
X_{r1} & X_{r2} & X_{r3} & \cdots & X_{rr}
\end{bmatrix}, \tag{5.3}$$

with blocks X_{ij} of size $N_i \times N_j$, and that

$$\mathcal{V} := \left\{ X \in \mathbb{S}^N : X_{ij} \in \mathcal{V}_{ij}, \ i \in \{1, \ldots, r\}, j \in \{1, \ldots, i\} \right\}, \tag{5.4}$$

where $\mathcal{V}_{ii}$ is a subspace of $\mathbb{S}^{N_i}$ and, for $i \neq j$, $\mathcal{V}_{ij}$ is a subspace of $\mathbb{R}^{N_i \times N_j}$. For $j > i$ we define $\mathcal{V}_{ij} = \{U^\top : U \in \mathcal{V}_{ji}\}$. Suppose the subspaces $\mathcal{V}_{ij}$ satisfy the following properties.

P1. The diagonal blocks are multiples of the identity: $\mathcal{V}_{ii} = \{\alpha I : \alpha \in \mathbb{R}\}$ for $i = 1, \ldots, r$.

P2. The lower-triangular blocks have orthogonal rows of equal norm: if $i > j$ and $A \in \mathcal{V}_{ij}$, then $AA^\top$ is a multiple of the identity.

P3. If $i > j > k$, then the subspaces $\mathcal{V}_{ij}, \mathcal{V}_{jk}, \mathcal{V}_{ik}$ are related as follows:

$$A \in \mathcal{V}_{ik}, \ B \in \mathcal{V}_{jk} \implies AB^\top \in \mathcal{V}_{ij}.$$

P4. If $i > j > k$, then the subspaces $\mathcal{V}_{ij}, \mathcal{V}_{jk}, \mathcal{V}_{ik}$ are related as follows:

$$A \in \mathcal{V}_{ij}, \ B \in \mathcal{V}_{jk} \implies AB \in \mathcal{V}_{ik}.$$

Ishi (2015, Theorem 3) shows that the cone K is homogeneous. Sections 5.2–5.4 will explain this in more detail.

Property P1 implies that $I \in \mathcal{V}$, so $\mathcal{V} \cap \mathbb{S}_{++}^N \neq \emptyset$, as assumed at the beginning of this section. A useful equivalent form of P2 is the following: if $i > j$ and $B, C \in \mathcal{V}_{ij}$, then $BC^\top + CB^\top$ is a multiple of the identity. This follows from P2 applied to $A = B + C$ and, conversely, clearly implies P2 if we take $B = C = A$.

In the next sections we use the following notation for the set of lower-triangular matrices with $L + L^\top \in \mathcal{V}$:

$$\mathcal{T} := \{ L \in \mathbb{T}^N : L_{ij} \in \mathcal{V}_{ij}, \ i \in \{1, \ldots, r\}, \ j \in \{1, \ldots, i\} \}. \tag{5.5}$$

Here L_{ij} refers to the $N_i \times N_j$ submatrix of L, partitioned as in (5.3).

5.1. Examples

Homogeneous sparse matrix cones. The homogeneous sparse matrix cones of Sections 3–4 are a special case with $\mathcal{V} := \mathbb{S}_E^N$. Suppose E is a homogeneous chordal sparsity pattern and that the numerical order $1, \ldots, N$ is a trivially perfect elimination ordering. Define $r := N$, $N_1 := \cdots := N_r := 1$, and

$$\mathcal{V}_{ij} := \begin{cases} \{0\} & i \neq j \text{ and } \{i, j\} \notin E, \\ \mathbb{R} & \text{otherwise.} \end{cases}$$

Properties P1 and P2 hold trivially, since $N_i = 1$ for all i. Property P3 reduces to

$$i > j > k, \quad \{i, k\} \in E, \quad \{j, k\} \in E \implies \{i, j\} \in E.$$

This is the property (2.4) of a perfect elimination ordering of a chordal graph. Property P4 is

$$i > j > k, \quad \{i, j\} \in E, \quad \{j, k\} \in E \implies \{i, k\} \in E.$$

This is the additional property (2.7) of a trivially perfect elimination ordering.

Block-sparsity. As an extension, we can define a block-sparsity pattern for a matrix partitioned as in (5.3) as an undirected graph with vertex set $V = \{1, 2, \ldots, r\}$ and edge set

$$E = \{\{i, j\} : i \neq j, \mathcal{V}_{ij} \neq \{0\}\}.$$

Properties P3 and P4 imply (among other conditions on the subspaces) that the graph (V, E) represents an $r \times r$ homogeneous chordal sparsity pattern with trivially perfect ordering $1, \ldots, r$. As an example, properties P1–P4 are satisfied by the subspace $\mathcal{V}$ of matrices of the form

$$\begin{bmatrix} \alpha I & 0 & u \\ 0 & \beta I & v \\ u^\top & v^\top & \gamma \end{bmatrix}$$

with $u \in \mathbb{R}^{N_1}$, $v \in \mathbb{R}^{N_2}$ and $\alpha, \beta, \gamma \in \mathbb{R}$. The corresponding homogeneous chordal sparsity pattern is the 3×3 pattern of Figure 2.4.

Rotated quadratic cone. The subspace

$$\mathcal{V} = \left\{ \begin{bmatrix} \alpha I & u \\ u^\top & \beta \end{bmatrix} : \alpha, \beta \in \mathbb{R}, \ u \in \mathbb{R}^{N-1} \right\}$$

is a special case with $r = 2$, $N_1 = N - 1$, $N_2 = 1$ and $\mathcal{V}_{12} = \mathbb{R}^{N-1}$. The cone $K = \mathcal{V} \cap \mathbb{S}_+^N$ is linearly isomorphic to the cone

$$\mathcal{Q}_r = \{(\alpha, \beta, u) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{N-1} : \alpha, \beta \geq 0, \ \alpha\beta \geq u^\top u\} \tag{5.6}$$

$$= \left\{ (\alpha, \beta, u) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{N-1} : \begin{bmatrix} \alpha I & u \\ u^\top & \beta \end{bmatrix} \geq 0 \right\}.$$

The cone $\mathcal{Q}_r$ is known as the *rotated quadratic cone* and is a symmetric cone. It can be used to represent the second-order cone

$$\mathcal{Q} = \{(t, y) \in \mathbb{R} \times \mathbb{R}^N : \|y\|_2 \leq t\}$$

as

$$\mathcal{Q} = \{(t, y) \in \mathbb{R} \times \mathbb{R}^N : (t + y_1, t - y_1, \bar{y}) \in \mathcal{Q}_r\},$$

where $\bar{y} = (y_2, \ldots, y_N)$.

Non-sparse example. Define $\mathcal{V}$ as the set of matrices of the form

$$\begin{bmatrix} \alpha I & 0 & u_1 & -u_2 & v_1 \\ 0 & \alpha I & u_2 & u_1 & v_2 \\ u_1^\top & u_2^\top & \beta & 0 & w_1 \\ -u_2^\top & u_1^\top & 0 & \beta & w_2 \\ v_1^\top & v_2^\top & w_1 & w_2 & \gamma \end{bmatrix}$$

with $\alpha, \beta, \gamma, w_1, w_2 \in \mathbb{R}$ and $u_1, u_2, v_1, v_2 \in \mathbb{R}^M$. This is a special case with $N = 2M + 3$, $r = 3$, $N_1 = 2M$, $N_2 = 2$, $N_3 = 1$ and

$$\mathcal{V}_{12} = \left\{ \begin{bmatrix} u_1 & -u_2 \\ u_2 & u_1 \end{bmatrix} : u_1, u_2 \in \mathbb{R}^M \right\}, \quad \mathcal{V}_{13} = \mathbb{R}^{2M}, \quad \mathcal{V}_{23} = \mathbb{R}^2.$$

Matrix norm cone. Define $\mathcal{V}$ as

$$\mathcal{V} = \left\{ \begin{bmatrix} \alpha I & U^\top \\ U & V \end{bmatrix} : t \in \mathbb{R},\ U \in \mathbb{R}^{K \times L},\ V \in \mathbb{S}^K \right\}.$$

This is a special case of Ishi's general structure, with $r = K + 1$, $N_1 = L$, $N_2 = \cdots = N_r = 1$. The off-diagonal subspaces are defined as

$$\mathcal{V}_{ij} = \begin{cases} \mathbb{R}^L & i = 1,\ j \in \{2, \ldots, r\}, \\ \mathbb{R} & i \in \{2, \ldots, r\},\ j \in \{2, \ldots, i-1\}. \end{cases}$$

The cone $K = \mathcal{V} \cap \mathbb{S}_+^N$ is known as the *matrix norm cone* and is important for trace norm minimization problems (Karimi and Tunçel 2020a,b).

Sparse matrix norm cone. The matrix norm cones and homogeneous sparse matrix cones can be combined in a new class of homogeneous matrix cones. Define $\mathcal{V}$ as

$$\mathcal{V} := \left\{ \begin{bmatrix} \alpha I & U^\top \\ U & V \end{bmatrix} : \alpha \in \mathbb{R},\ U \in \mathcal{U},\ V \in \mathbb{S}_E^K \right\},$$

where $\mathcal{U}$ is a subspace of $\mathbb{R}^{K \times L}$ with the property that for every $U \in \mathcal{U}$, the product $UU^\top \in \mathbb{S}_E^K$. An example is the set of positive semidefinite matrices of the form

$$\left[\begin{array}{ccccc|ccc} \alpha & 0 & 0 & 0 & 0 & u_1 & 0 & u_6 \\ 0 & \alpha & 0 & 0 & 0 & 0 & u_4 & u_7 \\ 0 & 0 & \alpha & 0 & 0 & 0 & u_5 & u_8 \\ 0 & 0 & 0 & \alpha & 0 & u_2 & 0 & u_9 \\ 0 & 0 & 0 & 0 & \alpha & u_3 & 0 & u_{10} \\ \hline u_1 & 0 & 0 & u_2 & u_3 & v_1 & 0 & v_2 \\ 0 & u_4 & u_5 & 0 & 0 & 0 & v_3 & v_4 \\ u_6 & u_7 & u_8 & u_9 & u_{10} & v_2 & v_4 & v_5 \end{array} \right].$$

5.2. Cholesky factorization

In this section we assume that $\mathcal{V}$ satisfies P1, P2, P3 but not necessarily P4. We show that every positive definite matrix $X \in \mathcal{V} \cap \mathbb{S}_{++}^N$ has a Cholesky factorization $X = LL^\top$ where $L \in \mathcal{T}$. This is the counterpart of the zero-fill Cholesky factorization of positive definite matrices with chordal sparsity patterns.

Proof. The proof is by induction on r. For $r = 1$, we have $\mathcal{V} = \{\alpha I : \alpha \in \mathbb{R}\}$ and the result is obvious, with $L = \sqrt{\alpha} I$ if $X = \alpha I$. We show that the result holds for

$r = m$ if it holds for $r = m - 1$. Suppose X is positive definite of the form

$$X = \begin{bmatrix} \alpha_1 I & X_{21}^\top & X_{31}^\top & \cdots & X_{r1}^\top \\ X_{21} & \alpha_2 I & X_{32}^\top & \cdots & X_{r2}^\top \\ X_{31} & X_{32} & \alpha_3 I & \cdots & X_{r3}^\top \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ X_{r1} & X_{r2} & X_{r3} & \cdots & \alpha_r I \end{bmatrix}$$

with $X_{ij} \in \mathcal{V}_{ij}$ for $i \in \{2, \ldots, r\}$, $j \in \{1, \ldots, i-1\}$, and that the subspaces $\mathcal{V}_{ij}$ satisfy P2 and P3. The matrix can be factored as

$$X = \begin{bmatrix} L_{11} & 0 & 0 & \cdots & 0 \\ L_{21} & I & 0 & \cdots & 0 \\ L_{31} & 0 & I & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ L_{r1} & 0 & 0 & \cdots & I \end{bmatrix} \begin{bmatrix} I & 0 & 0 & \cdots & 0 \\ 0 & Y_{22} & Y_{32}^\top & \cdots & Y_{r2}^\top \\ 0 & Y_{32} & Y_{33} & \cdots & Y_{r3}^\top \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & Y_{r2} & Y_{r3} & \cdots & Y_{rr} \end{bmatrix} \begin{bmatrix} L_{11} & 0 & 0 & \cdots & 0 \\ L_{21} & I & 0 & \cdots & 0 \\ L_{31} & 0 & I & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ L_{r1} & 0 & 0 & \cdots & I \end{bmatrix}^\top,$$

$$(5.7)$$

where $L_{11} := \sqrt{\alpha_1} I \in \mathcal{V}_{11}$, and $L_{i1} := X_{i1}/\sqrt{\alpha_1} \in \mathcal{V}_{i1}$ for $i \in \{2, \ldots, r\}$. The matrix Y is the Schur complement

$$\begin{bmatrix} Y_{22} & Y_{32}^\top & \cdots & Y_{r2}^\top \\ Y_{32} & Y_{33} & \cdots & Y_{r3}^\top \\ \vdots & \vdots & \ddots & \vdots \\ Y_{r2} & Y_{r3} & \cdots & Y_{rr} \end{bmatrix} = \begin{bmatrix} \alpha_2 I & X_{32}^\top & \cdots & X_{r2}^\top \\ X_{32} & \alpha_3 I & \cdots & X_{r3}^\top \\ \vdots & \vdots & \ddots & \vdots \\ X_{r2} & X_{r3} & \cdots & \alpha_r I \end{bmatrix} - \frac{1}{\alpha_1} \begin{bmatrix} X_{21} \\ X_{31} \\ \vdots \\ X_{r1} \end{bmatrix} \begin{bmatrix} X_{21} \\ X_{31} \\ \vdots \\ X_{r1} \end{bmatrix}^\top.$$

By property P2, the diagonal blocks $Y_{ii} = \alpha_i I - (1/\alpha_1) X_{i1} X_{i1}^\top$ are multiples of the identity, so $Y_{ii} \in \mathcal{V}_{ii}$. Property P3 implies that for $i > j$,

$$Y_{ij} = X_{ij} - \frac{1}{\alpha_1} X_{i1} X_{j1}^\top \in \mathcal{V}_{ij}.$$

Hence, by the induction hypothesis, the matrix Y can be factored as

$$\begin{bmatrix} Y_{22} & Y_{32}^\top & \cdots & Y_{r2}^\top \\ Y_{32} & Y_{33} & \cdots & Y_{r3}^\top \\ \vdots & \vdots & \ddots & \vdots \\ Y_{r2} & Y_{r3} & \cdots & Y_{rr} \end{bmatrix} = \begin{bmatrix} L_{22} & 0 & \cdots & 0 \\ L_{32} & L_{33} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ L_{r2} & L_{r3} & \cdots & L_{rr} \end{bmatrix} \begin{bmatrix} L_{22}^\top & L_{32}^\top & \cdots & L_{r2}^\top \\ 0 & L_{33}^\top & \cdots & L_{r3}^\top \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & L_{rr}^\top \end{bmatrix}$$

with $L_{ij} \in \mathcal{V}_{ij}$. Substituting the factorization of Y in (5.7) gives a Cholesky factorization $X = LL^\top$ with the desired properties.   □

5.3. Computations with triangular matrices

The following result generalizes Theorem 3.1 to triangular matrices in a subspace with the structure specified in properties P1–P4.

Theorem 5.1. Let $\mathcal{T}$ be the set of triangular matrices (5.5), where the subspaces $\mathcal{V}_{ij}$ satisfy properties P1–P4, and assume $L \in \mathcal{T}$.

(i) If $\tilde{L} \in \mathcal{T}$, then $L\tilde{L} \in \mathcal{T}$.

(ii) If L is non-singular, then $L^{-1} \in \mathcal{T}$.

(iii) If $X \in \mathcal{V}$, then $LXL^{\top} \in \mathcal{V}$.

(iv) If $Y \in \mathcal{V}^{\perp}$, then $L^{\top}YL \in \mathcal{V}^{\perp}$.

Proof. Suppose $L, \tilde{L} \in \mathcal{T}$ are partitioned as

$$
L = \begin{bmatrix} \alpha_1 I & 0 & 0 & \cdots & 0 \\ L_{21} & \alpha_2 I & 0 & \cdots & 0 \\ L_{31} & L_{32} & \alpha_3 I & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ L_{r1} & L_{r2} & L_{r3} & \cdots & \alpha_r I \end{bmatrix}, \quad \tilde{L} = \begin{bmatrix} \tilde{\alpha}_1 I & 0 & 0 & \cdots & 0 \\ \tilde{L}_{21} & \tilde{\alpha}_2 I & 0 & \cdots & 0 \\ \tilde{L}_{31} & \tilde{L}_{32} & \tilde{\alpha}_3 I & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \tilde{L}_{r1} & \tilde{L}_{r2} & \tilde{L}_{r3} & \cdots & \tilde{\alpha}_r I \end{bmatrix}.
$$

The diagonal blocks in the product $L\tilde{L}$ are $(L\tilde{L})_{ii} = \alpha_i \tilde{\alpha}_i I \in \mathcal{V}_{ii}$. For the lower-triangular off-diagonal blocks,

$$
(L\tilde{L})_{ij} = \tilde{\alpha}_j L_{ij} + \sum_{k=j+1}^{i-1} L_{ik} \tilde{L}_{kj} + \alpha_i \tilde{L}_{ij}.
$$

By assumption, the terms L_{ij}, $\tilde{L}_{ij}$ are in $\mathcal{V}_{ij}$. The middle term in the expression on the right-hand side is in $\mathcal{V}_{ij}$ by property P4. Hence P1 and P4 are sufficient to prove statement (i) of the theorem.

Part (ii) is proved by induction on r. For $r = 1$ it is obvious, with $L = \alpha I$ and $L^{-1} = \alpha^{-1} I$. Suppose the result holds for $r = m - 1$ and consider a matrix $L \in \mathcal{T}$, partitioned in $r \times r$ blocks as above. By the induction hypothesis, the blocks in

$$
\begin{bmatrix} L_{22} & 0 & \cdots & 0 \\ L_{32} & L_{33} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ L_{r2} & L_{r3} & \cdots & L_{rr} \end{bmatrix}^{-1} = \begin{bmatrix} \alpha_2^{-1} I & 0 & \cdots & 0 \\ G_{32} & \alpha_3^{-1} I & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ G_{r2} & G_{r3} & \cdots & \alpha_r^{-1} I \end{bmatrix}
$$

satisfy $G_{ij} \in \mathcal{V}_{ij}$ for $i > j > 1$. The inverse of L is

$$L^{-1} = \begin{bmatrix} \alpha_1^{-1}I & 0 & 0 & \cdots & 0 \\ G_{21} & \alpha_2^{-1}I & 0 & \cdots & 0 \\ G_{31} & G_{32} & \alpha_3^{-1}I & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ G_{r1} & G_{r2} & G_{r3} & \cdots & \alpha_r^{-1}I \end{bmatrix}$$

with

$$G_{21} = -(\alpha_1 \alpha_2)^{-1} L_{21}$$
$$G_{31} = -\alpha_1^{-1} G_{32} L_{21} - (\alpha_1 \alpha_3)^{-1} L_{31}$$
$$\vdots$$
$$G_{r1} = -\alpha_1^{-1}(G_{r2} L_{21} + \cdots + G_{r,r-1} L_{r-1,1}) - (\alpha_1 \alpha_r)^{-1} L_{r1}.$$

From the induction hypothesis ($G_{ij} \in \mathcal{V}_{ij}$ for $i > j > 1$) and property P4, we see that $G_{i1} \in \mathcal{V}_{i1}$ for $i \in \{2, \ldots, r\}$. Hence, statement (ii) of the theorem follows from P1 and P4.

Next we show part (iii). First consider the diagonal blocks of $LXL^\top$,

$$(LXL^\top)_{ii} = \sum_{k=1}^{i} L_{ik} X_{kk} L_{ik}^\top + \sum_{k=2}^{i} \sum_{l=1}^{k-1} \left(L_{ik} X_{kl} L_{il}^\top + L_{il} X_{kl}^\top L_{ik}^\top \right).$$

Properties P1, P2 and P3 imply that this is a product of the identity. For the off-diagonal blocks with $i > j$,

$$(LXL^\top)_{ij} = \sum_{k=1}^{j} L_{ik} X_{kk} L_{jk}^\top + \sum_{k=2}^{j} \sum_{l=1}^{k-1} \left(L_{ik} X_{kl} L_{jl}^\top + L_{il} X_{kl}^\top L_{jk}^\top \right).$$

Properties P1, P3 and P4 imply that this is an element in $\mathcal{V}_{ij}$.

Part (iv) is an immediate consequence of part (iii). Suppose $Y \in \mathcal{V}^\perp$. For any $X \in \mathcal{V}$,

$$\mathrm{Tr}(XL^\top YL) = \mathrm{Tr}(LXL^\top Y) = 0$$

because $LXL^\top \in \mathcal{V}$ by part (iii). Therefore $L^\top YL \in \mathcal{V}^\perp$. $\square$

5.4. Primal and dual cone automorphisms

Now consider the cones (5.1) and (5.2), where $\mathcal{V}$ is a subspace that satisfies the four properties P1–P4.

The linear mappings $\mathcal{L}(X) = LXL^\top$ for non-singular $L \in \mathcal{T}$ form a transitive subset of $\mathrm{Aut}(K)$. This is readily shown by extending the arguments in Section 3.2 using Theorem 5.1 and the property that Cholesky factors of matrices in $\mathrm{int}(K)$ are in $\mathcal{T}$ (see Section 5.2). In the remainder of Section 5, $\mathrm{Aut}_\triangle(K)$ will be used to denote this transitive subset of $\mathrm{Aut}(K)$.

The adjoints $\mathcal{L}^*(S) = \Pi_{\mathcal{V}}(L^\top S L)$ of mappings $\mathcal{L} \in \mathrm{Aut}_\triangle(K)$ form a transitive subset of $\mathrm{Aut}(K^*)$. The proof again parallels the arguments in Section 3.3. As noted in Section 3.3, every adjoint of an automorphism of K is an automorphism of K^*. We also note that every $S \in \mathrm{int}(K^*)$ can be expressed as

$$S = \Pi_{\mathcal{V}}(X^{-1}),$$

where $X \in \mathrm{int}(K)$. The matrix X is the solution of the convex optimization problem

$$\text{minimize} \quad \mathrm{Tr}(SX) - \ln \det(X)$$

with variable $X \in \mathcal{V}$. The Cholesky factorization $X = LL^\top$, where $L \in \mathcal{T}$, defines a mapping $\mathcal{L} \in \mathrm{Aut}_\triangle(K)$ that satisfies $\mathcal{L}^{-*}(I) = \Pi_{\mathcal{V}}(L^{-\top}L^{-1}) = S$. Therefore $\mathcal{L}^*(S) = I$.

Now consider any two matrices $S_1, S_2 \in \mathrm{int}(K^*)$. Find $\mathcal{L}_1, \mathcal{L}_2 \in \mathrm{Aut}_\triangle(K)$ that satisfy $\mathcal{L}_1^*(S_1) = I$, $\mathcal{L}_2^*(S_2) = I$. Then $\mathcal{L} = \mathcal{L}_1 \circ \mathcal{L}_2^{-1}$ satisfies $\mathcal{L}^*(S_1) = S_2$ and $\mathcal{L} \in \mathrm{Aut}_\triangle(K)$ by properties (i)–(iii) of Theorem 5.1. By establishing transitive subsets of $\mathrm{Aut}(K)$ and $\mathrm{Aut}(K^*)$, we have shown that these cones are homogeneous.

Using our results, in particular Theorem 5.1, we can establish the following fact.

Theorem 5.2. Let K be a homogeneous cone represented in $\mathbb{S}^N$ as described in (5.1). Then, for every $Z \in \mathcal{V} \cap \mathbb{S}^N$, upon expressing $Z = L + L^\top$ for some $L \in \mathcal{T}$, we have

$$LL^\top + \Pi_{\mathcal{V}}(L^\top L) + L^2 + (L^\top)^2 \in K^*.$$

Decompositions such as the above have potential applications in linear and non-linear complementarity problems over homogeneous cones and in the design of algorithms and theories utilizing Moreau decompositions; see for instance Kong, Tunçel and Xiu (2012).

5.5. Logarithmic barriers

If $K = \mathcal{V} \cap \mathbb{S}_+^N$ and $\mathcal{V}$ satisfies properties P1–P4, then the log-det barrier

$$F(X) := \begin{cases} -\ln \det(X) & \text{if } X \in \mathrm{int}(K), \\ +\infty & \text{otherwise} \end{cases} \tag{5.8}$$

has the same scaling properties (4.5) as the log-det barrier for a homogeneous sparse matrix cone. The proof is exactly the same. From (4.4) and the fact that $\mathcal{L}$ is an automorphism, it follows that (4.5) holds for all $X \in \mathrm{int}(K)$ and all $\mathcal{L} \in \mathrm{Aut}_\triangle(K)$. Similarly, (4.6) holds for all $S \in \mathrm{int}(K^*)$ and all $\mathcal{L} \in \mathrm{Aut}_\triangle(K)$.

With the above definition of the barrier function F, Theorem 4.1 extends to all homogeneous matrix cones K discussed in this section. This is stated in the following theorem.

Theorem 5.3. Let $K = \mathcal{V} \cap \mathbb{S}_+^N$, where $\mathcal{V}$ is a subspace that satisfies properties P1–P4. Then, for every $X \in \mathrm{int}(K)$ and $S \in \mathrm{int}(K^*)$, there exists $\mathcal{L} \in \mathrm{Aut}_\triangle(K)$ such

that

$$\mathcal{L}^{-1}(X) = \mathcal{L}^*(S).$$

Ishi has shown that every homogeneous cone can be represented in the form $K = \mathcal{V} \cap \mathbb{S}_+^N$, where $\mathcal{V}$ satisfies properties P1–P4. Theorem 5.3 therefore shows that problem (2) of Tunçel (1998, p. 711) is solvable for every homogeneous cone, and settles the open problem (i) of Tunçel (1998, p. 714).

Next we relate the above homogeneous matrix cones representation to algebraic classifications of all homogeneous cones and explain why the results of this section apply to all homogeneous cones.

6. Algebraic structure of homogeneous cones

In the previous sections we discussed classes of homogeneous cones defined as linear slices of the positive semidefinite cone. It turns out that every homogeneous cone can be expressed in this form. As mentioned by Faybusovich (2002, p. 214) and Papp and Alizadeh (2013, p. 1406), and worked out in detail by Chua (2003), this result is implicit in Vinberg's T-algebra-based classification of homogeneous cones, because Vinberg's results imply that every homogeneous cone is a 'cone of squares' for a suitable vector product. Rothaus, announcing a similar result first in 1963, proved it using the inductive Siegel-domain-based classification of homogeneous cones and convex cone duality (Rothaus 1963, 1966, 1968). Ishi's approach (Ishi 2013, 2016, 2015), influenced in part by some recent work by Yamasaki and Nomura (2015), brings Rothaus's Siegel-domain-based inductive construction closer to more direct utilization of the T-algebra axioms. In this section we discuss some of the results by Vinberg and Rothaus, and explain their connections to the classes of homogeneous matrix cones described in Sections 3–5.

It is useful to first clarify the meaning of *semidefinite representation* of a convex cone. A convex cone $\mathcal{V} \cap \mathbb{S}_+^N$, where $\mathcal{V} \subseteq \mathbb{S}^N$ is a linear subspace, can be equivalently represented as

$$K = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^n x_i A_i \geq 0 \right\}, \tag{6.1}$$

where $A_1, A_2, \ldots, A_n \in \mathbb{S}^N$. Given the subspace in the representation $\mathcal{V} \cap \mathbb{S}_+^N$, we can pick a basis $A_1, A_2, \ldots, A_n \in \mathbb{S}^N$ for $\mathcal{V}$ to obtain the representation (6.1). Given $A_1, A_2, \ldots, A_n \in \mathbb{S}^N$ in the second representation, we define $\mathcal{V} :=$ span$\{A_1, A_2, \ldots, A_n\}$ to obtain the former representation. The representation (6.1) is called a *linear matrix inequality* (LMI) or *spectrahedral* representation of the cone K. In spectrahedral representations one typically requires that $\mathcal{V} \cap \mathbb{S}_{++}^N \neq \emptyset$ (i.e. there exists $\bar{x} \in \mathbb{R}^n$ such that $\sum_{i=1}^m \bar{x}_i A_i > 0$).

Whenever a regular cone admits a spectrahedral representation with int$(K) = \mathcal{V} \cap \mathbb{S}_{++}^N \neq \emptyset$, the dual cone in the space $\mathbb{S}^N$, under the trace inner product, is given by $\mathcal{V}^\perp + \mathbb{S}_+^N$. (In general, a closure operation is needed on the right-hand side.

However, it can be shown that the cone $\mathcal{V}^\perp + \mathbb{S}_+^N$ is closed if $\mathcal{V} \cap \mathbb{S}_{++}^N \neq \emptyset$, so the closure operation can be omitted.) If we take the dual of $K = \mathcal{V} \cap \mathbb{S}_+^N$ with respect to the smaller space $\mathcal{V}$, the dual cone is

$$K^* = \Pi_{\mathcal{V}}(\mathcal{V}^\perp + \mathbb{S}_+^N) = \Pi_{\mathcal{V}}(\mathbb{S}_+^N).$$

The dual cone can therefore be represented in the form

$$K^* = \left\{ s \in \mathbb{R}^n : \sum_{i=1}^n s_i H_i + \sum_{j=1}^k u_j U_j \succeq 0, \text{ for some } u \in \mathbb{R}^k \right\}, \tag{6.2}$$

where $H_1, H_2, \ldots H_n, U_1, U_2, \ldots, U_k \in \mathbb{S}^N$ are given. This kind of semidefinite representation is called a *lifted-LMI* or *spectrahedral shadow* representation (of K^*); see Helton and Vinnikov (2007), Nemirovski (2007), Chua and Tunçel (2008), Helton and Nie (2010), Gouveia, Parrilo and Thomas (2013), Scheiderer (2018), Averkov (2019), Fawzi (2020) and the references therein. In a spectrahedral shadow representation the dual cone is expressed as the cone of positive semidefinite 'completable' matrices ('completable' by some element of $\mathcal{V}^\perp$). In our context, for the spectrahedral shadow representation (6.2), $\{H_1, H_2, \ldots, H_n\}$ is a basis for $\mathcal{V}$ and $\{U_1, U_2, \ldots, U_k\}$ can be taken as a basis for $\mathcal{V}^\perp$. If so, then $n + k = N(N+1)/2$.

Note that by our choices for these representations of K and K^* (i.e. for this choice of inner product and the space), we always have $K \subseteq K^*$.

6.1. Symmetric bilinear forms

Definition 6.1. Let K be a homogeneous cone in a finite-dimensional real vector space $\mathcal{V}$. A *homogeneous K-bilinear symmetric form* $\mathcal{B}(u, v)$ is a mapping from $\mathbb{R}^p \times \mathbb{R}^p$ to $\mathcal{V}$ that satisfies the following properties.

(i) $\mathcal{B}(\alpha_1 u^{(1)} + \alpha_2 u^{(2)}, v) = \alpha_1 \mathcal{B}(u^{(1)}, v) + \alpha_2 \mathcal{B}(u^{(2)}, v)$ for all $u^{(1)}, u^{(2)}, v \in \mathbb{R}^p$ and $\alpha_1, \alpha_2 \in \mathbb{R}$.

(ii) $\mathcal{B}(u, v) = \mathcal{B}(v, u)$ for all $u, v \in \mathbb{R}^p$.

(iii) $\mathcal{B}(u, u) \in K$ for all $u \in \mathbb{R}^p$.

(iv) $\mathcal{B}(u, u) = 0$ implies $u = 0$.

(v) There exists a transitive subset $G \subseteq \text{Aut}(K)$ such that for every $g \in G$, there exists a linear transformation $\bar{g}$ on $\mathbb{R}^p$ which satisfies

$$g(\mathcal{B}(u, v)) = \mathcal{B}(\bar{g}u, \bar{g}v) \quad \text{for all } u, v \in \mathbb{R}^p. \tag{6.3}$$

In this definition, $p = 0$ is allowed. When $p = 0$, the mapping $\mathcal{B}$ is the trivial bilinear form (a constant zero vector).

We now discuss some implications of the five properties in the definition. We use the standard inner product $u^\top v$ in $\mathbb{R}^p$, an inner product $\langle s, x \rangle$ in $\mathcal{V}$, and let

$$K^* = \{ s \in \mathcal{V} : \langle s, x \rangle \geq 0 \text{ for all } x \in K \}$$

denote the corresponding dual cone. The trace inner product is used for symmetric matrices.

A function $\mathcal{B}$ that satisfies properties (i) and (ii) of Definition 6.1 is called a *symmetric bilinear form*. With every symmetric bilinear form $\mathcal{B}$ one can associate a linear matrix function $\mathcal{H} \colon \mathcal{V} \to \mathbb{S}^p$, defined by the identity

$$u^\top \mathcal{H}(s)v = v^\top \mathcal{H}(s)u = \langle s, \mathcal{B}(u,v) \rangle \quad \text{for all } u, v \in \mathbb{R}^p, s \in \mathcal{V}. \tag{6.4}$$

An explicit formula for the entries of $\mathcal{H}(s)$ is

$$\mathcal{H}(s)_{ij} = \langle s, \mathcal{B}(e_i, e_j) \rangle, \quad i, j = 1, \dots, p, \tag{6.5}$$

where $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0), \dots, e_p = (0, \dots, 0, 1)$ are the standard unit vectors in $\mathbb{R}^p$. This expression follows from (6.4) if we use the bilinearity property (i) of the definition to expand $\mathcal{B}(u,v)$ as

$$\mathcal{B}(u,v) = \mathcal{B}(u_1 e_1 + \cdots + u_p e_p, v_1 e_1 + \cdots + v_p e_p)$$

$$= \sum_{i=1}^{p} \sum_{j=1}^{p} u_i v_j \mathcal{B}(e_i, e_j). \tag{6.6}$$

We will refer to $\mathcal{H}$ as the *dual representation* of the bilinear form $\mathcal{B}$. The adjoint of $\mathcal{H}$ (with respect to the inner product $\langle \cdot, \cdot \rangle$ in $\mathcal{V}$ and the trace inner product in $\mathbb{S}^p$) is the linear mapping $\mathcal{H}^* \colon \mathbb{R}^p \to \mathcal{V}$ that maps a matrix $Y \in \mathbb{S}^p$ to the vector

$$\mathcal{H}^*(Y) = \sum_{i=1}^{p} \sum_{j=1}^{p} Y_{ij} \mathcal{B}(e_i, e_j).$$

Hence, from (6.6), we have the following expression for $\mathcal{B}$:

$$\mathcal{B}(u,v) = \mathcal{H}^*((uv^\top + vu^\top)/2). \tag{6.7}$$

In particular, $\mathcal{B}(u,u) = \mathcal{H}^*(uu^\top)$. If $\mathcal{B}$ is the trivial bilinear form, we define $\mathcal{H}^*$ as the constant zero in $\mathcal{V}$.

The formula $\mathcal{B}(u,u) = \mathcal{H}^*(uu^\top)$ has an important consequence for semidefinite programming applications. It implies that the 'sum of squares' cone

$$C := \left\{ \sum_{i=1}^{k} \mathcal{B}(u^{(i)}, u^{(i)}) \colon \text{for some } k \text{ and } u^{(1)}, \dots, u^{(k)} \in \mathbb{R}^p \right\} \tag{6.8}$$

of any symmetric bilinear form $\mathcal{B}$ has a spectrahedral representation

$$C = \{\mathcal{H}^*(Y) \colon Y \succeq 0\}$$

(Nesterov 2000, Faybusovich 2002, Papp and Alizadeh 2013). This follows from $\mathcal{B}(u,u) = \mathcal{H}^*(uu^\top)$ and linearity of $\mathcal{H}^*$: all elements in C can be expressed as

$$\sum_{i=1}^{k} \mathcal{B}(u^{(i)}, u^{(i)}) = \sum_{i=1}^{k} \mathcal{H}^*(u^{(i)}(u^{(i)})^\top) = \mathcal{H}^*(Y),$$

where $Y = \sum_i u^{(i)}(u^{(i)})^\top \geq 0$, and, conversely, if $x = \mathcal{H}^*(Y)$ with $Y \geq 0$, then any decomposition $Y = \sum_i u^{(i)}(u^{(i)})^\top$ gives an expression $x = \sum_i \mathcal{B}(u^{(i)}, u^{(i)})$ showing that $x \in C$.

Properties (iii)–(v) of Definition 6.1 can be stated in equivalent forms involving the dual representation $\mathcal{H}$ and its adjoint.

Proposition 6.2. Let K be a regular convex cone in a finite-dimensional real vector space $\mathcal{V}$. Let $\mathcal{B} \colon \mathbb{R}^p \times \mathbb{R}^p \to \mathcal{V}$ be a symmetric bilinear form and $\mathcal{H}$ its dual representation defined in (6.4).

(i) Each of the following two statements is equivalent to the property that $\mathcal{B}(u, u) \in K$ for all u:

$$\mathcal{H}^*(Y) \in K \quad \text{for all } Y \geq 0, \tag{6.9a}$$

$$\mathcal{H}(s) \geq 0 \quad \text{for all } s \in K^*. \tag{6.9b}$$

(ii) Each of the following two statements is equivalent to the property that $\mathcal{B}(u, u) \in K \setminus \{0\}$ for all $u \neq 0$:

$$\mathcal{H}^*(Y) \in K \setminus \{0\} \quad \text{for all non-zero } Y \geq 0, \tag{6.10a}$$

$$\mathcal{H}(s) > 0 \quad \text{for all } s \in \text{int}(K^*). \tag{6.10b}$$

(iii) Let $g \colon \mathcal{V} \to \mathcal{V}$ and $\bar{g} \colon \mathbb{R}^p \to \mathbb{R}^p$ be linear transformations. Each of the following two statements is equivalent to the property that $g(\mathcal{B}(u, v)) = \mathcal{B}(\bar{g}u, \bar{g}v)$ for all $u, v \in \mathbb{R}^p$:

$$g(\mathcal{H}^*(Y)) = \mathcal{H}^*(\bar{g}Y\bar{g}^\top) \quad \text{for all } Y \in \mathbb{S}^p, \tag{6.11a}$$

$$\mathcal{H}(g^*(s)) = \bar{g}^\top \mathcal{H}(s)\bar{g} \quad \text{for all } s \in \mathcal{V}. \tag{6.11b}$$

Part (iii) of the proposition follows directly from (6.4) and (6.7). The statements about $\mathcal{H}^*$ in the first two parts follow from $\mathcal{B}(u, u) = \mathcal{H}^*(uu^\top)$ and linearity of $\mathcal{H}^*$. The statements about $\mathcal{H}$ follow from $u^\top \mathcal{H}(s)u = \langle s, \mathcal{B}(u, u) \rangle$, and the equivalences

$$\mathcal{B}(u, u) \in K \iff \langle s, \mathcal{B}(u, u) \rangle \geq 0 \quad \text{for all } s \in K^*$$
$$\iff u^\top \mathcal{H}(s)u \geq 0 \quad \text{for all } s \in K^*$$

and

$$\mathcal{B}(u, u) \in K \setminus \{0\} \iff \langle s, \mathcal{B}(u, u) \rangle > 0 \quad \text{for all } s \in \text{int}(K^*)$$
$$\iff u^\top \mathcal{H}(s)u > 0 \quad \text{for all } s \in \text{int}(K^*).$$

Example 6.3. We take $\mathcal{V} = \mathbb{S}_E^3$, where E is the homogeneous chordal pattern in Figure 2.4, that is, $\mathcal{V}$ is the space of matrices of the form

$$X = \begin{bmatrix} X_{11} & 0 & X_{31} \\ 0 & X_{22} & X_{32} \\ X_{31} & X_{32} & X_{33} \end{bmatrix}.$$

We use the inner product $\langle S, X \rangle = \text{Tr}(SX)$ on $\mathcal{V}$ and define $K = \mathcal{V} \cap \mathbb{S}_+^3$. (This cone

is known as the *Vinberg cone*, the smallest-dimensional homogeneous cone that is not symmetric.) Consider the following symmetric bilinear form $\mathcal{B}\colon \mathbb{R}^6 \times \mathbb{R}^6 \to \mathcal{V}$:

$$
\begin{aligned}
&\mathcal{B}(u, v) \\
&= \frac{1}{2}
\begin{bmatrix} u_1 & u_3 & 0 \\ 0 & 0 & u_5 \\ u_2 & u_4 & u_6 \end{bmatrix}
\begin{bmatrix} v_1 & v_3 & 0 \\ 0 & 0 & v_5 \\ v_2 & v_4 & v_6 \end{bmatrix}^{\top}
+
\begin{bmatrix} v_1 & v_3 & 0 \\ 0 & 0 & v_5 \\ v_2 & v_4 & v_6 \end{bmatrix}
\begin{bmatrix} u_1 & u_3 & 0 \\ 0 & 0 & u_5 \\ u_2 & u_4 & u_6 \end{bmatrix}^{\top} \\
&= \frac{1}{2}
\begin{bmatrix}
2(u_1v_1 + u_3v_3) & 0 & u_1v_2 + u_2v_1 + u_3v_4 + u_4v_3 \\
0 & 2u_5v_5 & u_5v_6 + u_6v_5 \\
u_1v_2 + u_2v_1 + u_3v_4 + u_4v_3 & u_5v_6 + u_6v_5 & 2(u_2v_2 + u_4v_4 + u_6v_6)
\end{bmatrix}.
\end{aligned}
$$

The dual representation $\mathcal{H}\colon \mathcal{V} \to \mathbb{S}^6$ is

$$
\mathcal{H}(S) =
\begin{bmatrix}
S_{11} & S_{31} & 0 & 0 & 0 & 0 \\
S_{31} & S_{33} & 0 & 0 & 0 & 0 \\
0 & 0 & S_{11} & S_{31} & 0 & 0 \\
0 & 0 & S_{31} & S_{33} & 0 & 0 \\
0 & 0 & 0 & 0 & S_{22} & S_{32} \\
0 & 0 & 0 & 0 & S_{32} & S_{33}
\end{bmatrix},
$$

and its adjoint $\mathcal{H}^*\colon \mathbb{S}^6 \to \mathcal{V}$ is

$$
\mathcal{H}^*(Y) =
\begin{bmatrix}
Y_{11} + Y_{33} & 0 & Y_{21} + Y_{43} \\
0 & Y_{55} & Y_{65} \\
Y_{21} + Y_{43} & Y_{65} & Y_{22} + Y_{44} + Y_{66}
\end{bmatrix}.
$$

This bilinear form $\mathcal{B}$ satisfies the five properties of Definition 6.1. It satisfies properties (iii) and (iv), as can be seen from

$$
\begin{aligned}
\mathcal{B}(u, u) &= \frac{1}{2}
\begin{bmatrix} u_1 & u_3 & 0 \\ 0 & 0 & u_5 \\ u_2 & u_4 & u_6 \end{bmatrix}
\begin{bmatrix} u_1 & u_3 & 0 \\ 0 & 0 & u_5 \\ u_2 & u_4 & u_6 \end{bmatrix}^{\top} \\
&=
\begin{bmatrix}
u_1^2 + u_3^2 & 0 & u_1u_2 + u_3u_4 \\
0 & u_5^2 & u_5u_6 \\
u_1u_2 + u_3u_4 & u_5u_6 & u_2^2 + u_4^2 + u_6^2
\end{bmatrix}.
\end{aligned}
$$

The first expression shows that $\mathcal{B}(u, u)$ is positive semidefinite for all u; the second expression shows that $\mathcal{B}(u, u) = 0$ only if $u = 0$. For property (v) we use the transitive subset of $\mathrm{Aut}(K)$ discussed in Section 3.2. The automorphisms in G are the mappings $g = \mathcal{L}\colon \mathcal{V} \to \mathcal{V}$ defined as $\mathcal{L}(X) = LXL^{\top}$, where L is a non-singular triangular matrix

$$
L =
\begin{bmatrix}
L_{11} & 0 & 0 \\
0 & L_{22} & 0 \\
L_{31} & L_{32} & L_{33}
\end{bmatrix}.
$$

Then $\mathcal{L}(\mathcal{B}(u, v)) = \mathcal{B}(\bar{g}u, \bar{g}v)$, where

$$
\bar{g} = \begin{bmatrix}
L_{11} & 0 & 0 & 0 & 0 & 0 \\
L_{31} & L_{33} & 0 & 0 & 0 & 0 \\
0 & 0 & L_{11} & 0 & 0 & 0 \\
0 & 0 & L_{31} & L_{33} & 0 & 0 \\
0 & 0 & 0 & 0 & L_{22} & 0 \\
0 & 0 & 0 & 0 & L_{32} & L_{33}
\end{bmatrix}.
$$

Example 6.4. With the same choice of $\mathcal{V}$ and K, define $\mathcal{B} \colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathcal{V}$ as

$$
\begin{aligned}
\mathcal{B}(u, v) &= \frac{1}{2} \begin{bmatrix} u_1 \\ 0 \\ u_2 \end{bmatrix} \begin{bmatrix} v_1 \\ 0 \\ v_2 \end{bmatrix}^\top + \frac{1}{2} \begin{bmatrix} v_1 \\ 0 \\ v_2 \end{bmatrix} \begin{bmatrix} u_1 \\ 0 \\ u_2 \end{bmatrix}^\top \\
&= \frac{1}{2} \begin{bmatrix}
2u_1 v_1 & 0 & u_1 v_2 + u_2 v_1 \\
0 & 0 & 0 \\
u_1 v_2 + u_2 v_1 & 0 & 2u_2 v_2
\end{bmatrix}.
\end{aligned}
$$

The dual representation $\mathcal{H} \colon \mathcal{V} \to \mathbb{S}^2$ and its adjoint $\mathcal{H}^* \colon \mathbb{S}^2 \to \mathcal{V}$ are

$$
\mathcal{H}(S) = \begin{bmatrix} S_{11} & S_{31} \\ S_{31} & S_{33} \end{bmatrix}, \quad \mathcal{H}^*(Y) = \begin{bmatrix} Y_{11} & 0 & Y_{21} \\ 0 & 0 & 0 \\ Y_{21} & 0 & Y_{22} \end{bmatrix}.
$$

Here,

$$
\mathcal{B}(u, u) = \begin{bmatrix} u_1 \\ 0 \\ u_2 \end{bmatrix} \begin{bmatrix} u_1 \\ 0 \\ u_2 \end{bmatrix}^\top = \begin{bmatrix} u_1^2 & 0 & u_1 u_2 \\ 0 & 0 & 0 \\ u_1 u_2 & 0 & u_2^2 \end{bmatrix},
$$

which satisfies properties (iii) and (iv) of Definition 6.1. Property (v) holds for the same transitive subset G as in the previous example and

$$
\bar{g} = \begin{bmatrix} L_{11} & 0 \\ L_{31} & L_{33} \end{bmatrix}.
$$

Therefore $\mathcal{B}$ is another homogeneous K-bilinear symmetric form for the same cone K. Note that, in contrast to the previous example, K is not equal to the sum-of-squares cone (6.8). Here, strict inclusions $K \supset \{\mathcal{H}^*(Y) \colon Y \succeq 0\}$ and $K^* \subset \{S \colon \mathcal{H}(S) \succeq 0\}$ hold.

6.2. Siegel cone

Let K be a homogeneous cone and $\mathcal{B}$ a homogeneous K-bilinear symmetric form as defined in Definition 6.1. We define the *Siegel cone* associated with K and $\mathcal{B}$ as

$$\mathrm{SC}(K, \mathcal{B})$$
$$:= \{(x, u, \alpha) \in \mathcal{V} \times \mathbb{R}^p \times \mathbb{R} \colon x \in K, \ \alpha \geq 0, \ \alpha x - \mathcal{B}(u, u) \in K\}$$
$$= \left\{(x, u, \alpha) \colon \alpha > 0, \ x - \frac{1}{\alpha}\mathcal{B}(u, u) \in K\right\} \cup \{(x, 0, 0) \colon x \in K\}. \qquad (6.12)$$

If $\mathcal{B}$ is the trivial bilinear form ($p = 0$), the Siegel cone is $\mathrm{SC}(K, \mathcal{B}) = K \times \mathbb{R}_+$.

The following equivalent definition follows from the results in the previous section and makes it clear that $\mathrm{SC}(K, \mathcal{B})$ is convex: if $p \geq 1$,

$$\mathrm{SC}(K, \mathcal{B})$$
$$= \left\{(x, u, \alpha) \colon x - \mathcal{H}^*(Y) \in K, \begin{bmatrix} \alpha & u^\top \\ u & Y \end{bmatrix} \succeq 0 \text{ for some } Y \in \mathbb{S}^p\right\}. \qquad (6.13)$$

This definition also shows that $\mathrm{SC}(K, \mathcal{B})$ has a spectrahedral shadow representation if the cone K has a spectrahedral shadow representation. The equivalence of (6.12) and (6.13) can be seen as follows. We first note that in both definitions the only elements (x, u, α) with $\alpha = 0$ are the vectors $(x, 0, 0)$, $x \in K$. If $\alpha = 0$, the matrix inequality in (6.13) requires $u = 0$ and $Y \succeq 0$. Since $\mathcal{H}^*(Y) \in K$ for all $Y \succeq 0$, the condition on x then reduces to $x \in K$. Next, suppose $\alpha > 0$ and (x, u, α) is in the cone (6.12). Then $Y = (1/\alpha)uu^\top$ satisfies the conditions in (6.13), so (x, u, α) is in the cone (6.13). Conversely, suppose $\alpha > 0$ and (x, u, α) satisfies the conditions in (6.13) for some $Y \succeq 0$. Then $Y \succeq (1/\alpha)uu^\top$ and therefore $\mathcal{H}^*(Y) - (1/\alpha)\mathcal{H}^*(uu^\top) \in K$. Hence $x - \mathcal{B}(u, u)/\alpha = x - \mathcal{H}^*(uu^\top)/\alpha \in K$, so (x, u, α) is an element of the cone (6.12).

To establish the equivalence between (6.13) and (6.12) we only used properties (i)–(iii) of Definition 6.1. Clearly, $\mathrm{SC}(K, \mathcal{B})$ has non-empty interior in $\mathcal{V} \times \mathbb{R}^p \times \mathbb{R}$, since K has non-empty interior in $\mathcal{V}$. Property (iv) further implies that $\mathrm{SC}(K, \mathcal{B})$ is closed and pointed, so it is a regular cone. It is closed because $\mathrm{SC}(K, \mathcal{B})$ can be expressed as the image of a closed convex cone $\mathbb{S}_+^{p+1} \times K$ under the linear transformation

$$\mathcal{A}(W, w) = (\mathcal{H}^*(W_{22}) + w, W_{21}, W_{11}),$$

where W_{11} is scalar, $W_{21} \in \mathbb{R}^p$ and W_{22} is the trailing $p \times p$ submatrix in

$$W = \begin{bmatrix} W_{11} & W_{21}^\top \\ W_{21} & W_{22} \end{bmatrix}.$$

Property (iv) in its form (6.10a) implies that $\mathcal{A}(W, w) = 0$, $W \succeq 0$, $w \in K$ only holds for $W = 0$, $w = 0$. Hence, by Theorem 9.1 of Rockafellar (1970), the set $\mathrm{SC}(K, \mathcal{B})$ is closed. By a similar argument, $\mathrm{SC}(K, \mathcal{B})$ is pointed. Suppose

$(x, u, \alpha) \in SC(K, \mathcal{B})$ and $-(x, u, \alpha) \in SC(K, \mathcal{B})$, so

$$(x, u, \alpha) = \mathcal{A}(W, w), \quad -(x, u, \alpha) = \mathcal{A}(\tilde{W}, \tilde{w}),$$

for some $W, \tilde{W} \geq 0$, $w, \tilde{w} \in K$. Therefore $0 = \mathcal{A}(W + \tilde{W}, w + \tilde{w})$, and by property (iv), $W = \tilde{W} = 0$ and $w = \tilde{w} = 0$. Hence $(x, u, \alpha) = 0$.

The dual cone of $SC(K, \mathcal{B})$, if we use the inner product $\langle s, x \rangle + 2v^\top u + \beta\alpha$ between (s, v, β) and (x, u, α), is given by

$$SC(K, \mathcal{B})^* = \left\{ (s, v, \beta) \in \mathcal{V} \times \mathbb{R}^p \times \mathbb{R} : s \in K^*, \begin{bmatrix} \beta & v^\top \\ v & \mathcal{H}(s) \end{bmatrix} \succeq 0 \right\} \qquad (6.14)$$

if $p \geq 1$. If $\mathcal{B}$ is the trivial bilinear form ($p = 0$), the dual Siegel cone is $SC(K, \mathcal{B})^* = K^* \times \mathbb{R}_+$. The dual Siegel cone is closed, convex and pointed, and property (iv) of Definition 6.1 implies that it has non-empty interior.

Note that the expression (6.14) shows that $SC(K, \mathcal{B})^*$ has a spectrahedral representation if K^* has a spectrahedral representation.

So far we have only used properties (i)–(iv) of Definition 6.1. Property (v) further implies that the Siegel cone is a homogeneous cone. This result is due to Vinberg (1965a). To see this, it is sufficient to verify that the group generated by the following linear transformations on $\mathcal{V} \times \mathbb{R}^p \times \mathbb{R}$ forms a transitive subset of $\mathrm{Aut}(SC(K, \mathcal{B}))$:

$$T_1(\gamma) \colon (x, u, \alpha) \mapsto (x, \sqrt{\gamma}u, \gamma\alpha), \qquad (6.15)$$

$$T_2(w) \colon (x, u, \alpha) \mapsto (x + 2\mathcal{B}(w, u) + \alpha\mathcal{B}(w, w), u + \alpha w, \alpha), \qquad (6.16)$$

$$T_3(g) \colon (x, u, \alpha) \mapsto (g(x), \bar{g}u, \alpha). \qquad (6.17)$$

Here, T_1 is parametrized by a scalar $\gamma > 0$, T_2 by a vector $w \in \mathbb{R}^p$, and T_3 by an automorphism $g \in G$, where G is the transitive subset of $\mathrm{Aut}(K)$ mentioned in property (v) of Definition 6.1. The mapping $\bar{g}$ is the corresponding linear transformation in $\mathbb{R}^p$ and satisfies (6.3). It is easy to check, using (6.12) or (6.13), that these transformations are automorphisms of $SC(K, \mathcal{B})$. To verify that they form a transitive subset, consider an arbitrary pair of points (x, u, α) and $(\hat{x}, \hat{u}, \hat{\alpha})$ in the interior of $SC(K, \mathcal{B})$. Let $g \in G$ be an automorphism that maps $x - \mathcal{B}(u, u)/\alpha$ to $\hat{x} - \mathcal{B}(\hat{u}, \hat{u})/\hat{\alpha}$. Then the mapping

$$T_1(\hat{\alpha}) \circ T_2(\hat{u}/\hat{\alpha}^{1/2}) \circ T_3(g) \circ T_2(-u/\alpha^{1/2}) \circ T_1(1/\alpha)$$

is an automorphism of $SC(K, \mathcal{B})$ that maps (x, u, α) to $(\hat{x}, \hat{u}, \hat{\alpha})$.

By duality, the adjoints of the mappings $T_1(\gamma)$, $T_2(w)$, $T_3(g)$ form a transitive subset of the automorphism group of $SC(K, \mathcal{B})^*$. The adjoints are given by

$$T_1(\gamma)^* \colon (s, v, \beta) \mapsto (s, \sqrt{\gamma}v, \gamma\beta), \qquad (6.18)$$

$$T_2(w)^* \colon (s, v, \beta) \mapsto (s, \mathcal{H}(s)w + v, w^\top\mathcal{H}(s)w + 2w^\top v + \beta), \qquad (6.19)$$

$$T_3(g)^* \colon (s, v, \beta) \mapsto (g^*(s), \bar{g}^\top v, \beta) \qquad (6.20)$$

and are exploited in the work of Rothaus (1966). These mappings correspond to

congruence operations

$$\begin{bmatrix} \sqrt{\gamma} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \beta & v^\top \\ v & \mathcal{H}(s) \end{bmatrix} \begin{bmatrix} \sqrt{\gamma} & 0 \\ 0 & I \end{bmatrix}, \quad \begin{bmatrix} 1 & w^\top \\ 0 & I \end{bmatrix} \begin{bmatrix} \beta & v^\top \\ v & \mathcal{H}(s) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ w & I \end{bmatrix}$$

and

$$\begin{bmatrix} 1 & 0 \\ 0 & \bar{g}^\top \end{bmatrix} \begin{bmatrix} \beta & v^\top \\ v & \mathcal{H}(s) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \bar{g} \end{bmatrix},$$

respectively, where on the last line we use the identity $\bar{g}^\top \mathcal{H}(s)\bar{g} = \mathcal{H}(g^*(s))$.

Rothaus calls the mapping $\mathcal{H}$, associated with a homogeneous K-bilinear symmetric form $\mathcal{B}$ via the definition (6.4), a *representation* of K^*, and he calls the dual Siegel cone $\mathrm{SC}(K, \mathcal{B})^*$ an *extension* of K^* from the representation $\mathcal{H}$ (Rothaus 1966). We have used the term *dual representation* for $\mathcal{H}$, to avoid confusion with general semidefinite representations of convex cones (i.e. spectrahedral representations or spectrahedral shadow representations).

Example 6.5. We take $K = \mathbb{R}_+$ and $\mathcal{V} = \mathbb{R}$, and the trivial bilinear form ($p = 0$ and $\mathcal{B} = 0$). The Siegel cone is

$$\mathrm{SC}(K, \mathcal{B}) = \{(x, \alpha) \in \mathbb{R} \times \mathbb{R} : x \geq 0, \alpha \geq 0, \alpha x \geq 0\}$$
$$= \left\{ (x, \alpha) \in \mathbb{R} \times \mathbb{R} : \begin{bmatrix} \alpha & 0 \\ 0 & x \end{bmatrix} \succeq 0 \right\}.$$

This cone is linearly isomorphic to $\mathbb{R}_+^2$.

For the same K and $\mathcal{V}$, consider $\mathcal{B} \colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defined as $\mathcal{B}(u, v) = uv$. The symmetric form clearly satisfies properties (i)–(iv) of Definition 6.1, and property (v) with automorphisms $g(x) = \gamma x$ for $\gamma > 0$, and linear transformations $\bar{g} = \sqrt{\gamma}$. The Siegel cone

$$\mathrm{SC}(K, \mathcal{B}) = \{(x, u, \alpha) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R} : x \geq 0, \alpha \geq 0, \alpha x \geq u^2\}$$
$$= \left\{ (x, u, \alpha) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R} : \begin{bmatrix} \alpha & u \\ u & x \end{bmatrix} \succeq 0 \right\}$$

is linearly isomorphic to $\mathbb{S}_+^2$.

Finally, consider $\mathcal{B} \colon \mathbb{R}^p \times \mathbb{R}^p$ with $p > 1$, defined as $\mathcal{B}(u, v) = u^\top v$. This is another K-bilinear homogeneous form. In property (v), we take automorphisms $g(x) = \gamma x$ for $\gamma > 0$ and $\bar{g} = \sqrt{\gamma} I$. The dual representation $\mathcal{H} \colon \mathbb{R} \to \mathbb{S}^p$ and its adjoint $\mathcal{H}^* \colon \mathbb{S}^p \to \mathbb{R}$ are $\mathcal{H}(s) = sI$ and $\mathcal{H}^*(Y) = \mathrm{Tr}(Y)$. With this choice of $\mathcal{B}$, we obtain

$$\mathrm{SC}(K, \mathcal{B}) = \{(x, u, \alpha) \in \mathbb{R} \times \mathbb{R}^p \times \mathbb{R} : x \geq 0, \alpha \geq 0, \alpha x \geq u^\top u\}$$
$$= \left\{ (x, u, \alpha) \in \mathbb{R} \times \mathbb{R}^p \times \mathbb{R} : \begin{bmatrix} \alpha I & u \\ u^\top & x \end{bmatrix} \succeq 0 \right\}.$$

This cone is linearly isomorphic to the rotated quadratic cone (5.6).

The most general form of a homogeneous K-bilinear symmetric form for $K = \mathbb{R}^+$ is $\mathcal{B}(u, v) = u^\top Q^{-1} v$, with $Q \in \mathbb{S}_{++}^p$. With this choice,

$$\text{SC}(K, \mathcal{B}) = \{(x, u, \alpha) \in \mathbb{R} \times \mathbb{R}^p \times \mathbb{R} : x \geq 0, \ \alpha \geq 0, \ \alpha x \geq u^\top Q^{-1} u\}$$

$$= \left\{ (x, u, \alpha) \in \mathbb{R} \times \mathbb{R}^p \times \mathbb{R} : \begin{bmatrix} \alpha Q & u \\ u^\top & x \end{bmatrix} \geq 0 \right\}.$$

This cone is a rotated quadratic cone after a linear transformation.

Hence $\mathbb{R}_+^2$, $\mathbb{S}_+^2$ and the rotated quadratic cones are essentially the only types of cones that can be constructed as Siegel cones of $\mathbb{R}_+$.

Example 6.6. We continue Examples 6.3 and 6.4. The Siegel cone for the bilinear form in Example 6.3 is

$$\text{SC}(K, \mathcal{B}) = \left\{ (x, u, \alpha) \in \mathbb{S}_E^3 \times \mathbb{R}^6 \times \mathbb{R} : \begin{bmatrix} \alpha & 0 & 0 & u_1 & 0 & u_2 \\ 0 & \alpha & 0 & u_3 & 0 & u_4 \\ 0 & 0 & \alpha & 0 & u_5 & u_6 \\ u_1 & u_3 & 0 & X_{11} & 0 & X_{31} \\ 0 & 0 & u_5 & 0 & X_{22} & X_{32} \\ u_2 & u_4 & u_6 & X_{31} & X_{32} & X_{33} \end{bmatrix} \geq 0 \right\}.$$

This is an example of a *sparse matrix norm cone* discussed in Section 5.1. For the bilinear form of Example 6.4, we obtain

$$\text{SC}(K, \mathcal{B}) = \left\{ (x, u, \alpha) \in \mathbb{S}_E^3 \times \mathbb{R}^2 \times \mathbb{R} : \begin{bmatrix} \alpha & u_1 & 0 & u_2 \\ u_1 & X_{11} & 0 & X_{31} \\ 0 & 0 & X_{22} & X_{32} \\ u_2 & X_{31} & X_{32} & X_{33} \end{bmatrix} \geq 0 \right\}.$$

This is a homogeneous sparse matrix cone (ordered using a trivially perfect elimination ordering).

6.3. Siegel-domain construction of homogeneous cones

The Siegel cone is the key tool in Vinberg's recursive construction of all homogeneous cones. A homogeneous cone K and a homogeneous K-bilinear symmetric form $\mathcal{B}$ together yield a Siegel cone $\text{SC}(K, \mathcal{B})$, which is a homogeneous cone in a higher-dimensional space. The converse is also true. For every homogeneous cone $\widehat{K}$ of dimension at least two, there exists a lower-dimensional homogeneous cone K and a homogeneous K-bilinear symmetric form $\mathcal{B}$ such that $\widehat{K}$ is linearly isomorphic to the Siegel cone $\text{SC}(K, \mathcal{B})$; see for example Rothaus (1966) or Gindikin (1992). This provides an inductive characterization (called *Siegel-domain construction*) of all homogeneous cones, starting with the ray $\mathbb{R}_+$ in $\mathbb{R}$ as the first homogeneous cone. The construction may be viewed as an abstraction of the commonly used concept of *Schur complement*.

The minimum number of steps required to construct K in this recursive way is called the *Siegel-rank of K*. We denote this integer-valued function of homogeneous

cones by $r(K)$ and define $r(\mathbb{R}_+) := 1$. Since K is homogeneous if and only if K^* is, Vinberg's classification theory described above also applies to K^*. Furthermore, the Siegel-ranks of K and K^* are always the same.

Example 6.7. The Vinberg cone

$$K := \left\{ x \in \mathbb{R}^5 : \begin{bmatrix} x_1 & 0 & x_2 \\ 0 & x_3 & x_4 \\ x_2 & x_4 & x_5 \end{bmatrix} \geq 0 \right\} \tag{6.21}$$

has Siegel-rank three. Let us construct it via the recursive procedure.

- Let $K_1 = \mathbb{R}_+$. Define $\mathcal{B}_1 \colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ by $\mathcal{B}(u, v) := uv$. Then K_1 and $\mathcal{B}_1$ satisfy the conditions of Definition 6.1 and their Siegel cone is linearly isomorphic to $\mathbb{S}_2^+$; see Example 6.5. This shows that $r(\mathbb{S}_+^2) = 2$.

- Let $K_2 := \mathbb{S}_+^2$. Define $\mathcal{B}_2 \colon \mathbb{R} \times \mathbb{R} \to \mathbb{S}^2$ by

$$\mathcal{B}_2(u, v) := \frac{1}{2} \begin{bmatrix} 0 \\ u \end{bmatrix} \begin{bmatrix} 0 & v \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 0 \\ v \end{bmatrix} \begin{bmatrix} 0 & u \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & uv \end{bmatrix}.$$

Again, K_2 and $\mathcal{B}_2$ satisfy the conditions of Definition 6.1. To check the fifth condition, we can take as the transitive subset G the set of linear maps $g(X) = LXL^\top$, where L is a non-singular 2×2 lower-triangular matrix

$$L = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix},$$

and define $\bar{g} = L_{22}$, so that $L(\mathcal{B}(u, v)L^\top = \mathcal{B}(L_{22}u, L_{22}v)$ as desired. The Siegel cone for K_2 and the bilinear form $\mathcal{B}_2$ is

$$\mathrm{SC}(K_2, \mathcal{B}_2) = \left\{ (X, u, \alpha) \in \mathbb{S}^2 \times \mathbb{R} \times \mathbb{R} : \begin{bmatrix} \alpha & 0 & u \\ 0 & X_{11} & X_{21} \\ u & X_{21} & X_{21} \end{bmatrix} \geq 0 \right\},$$

which is linearly isomorphic to the Vinberg cone (6.21). Thus we have derived the Vinberg cone as a homogeneous cone with $r(K) = 3$.

6.4. Semidefinite representations of homogeneous cones

Since every homogeneous cone of Siegel-rank $r \geq 2$ arises from a homogeneous cone of Siegel-rank $r - 1$, via the above construction, we can establish many properties of homogeneous cones by induction on the Siegel-rank. For example, from (6.13) we see that $\mathrm{SC}(K, \mathcal{B})$ has a spectrahedral shadow representation if K has a spectrahedral shadow representation. By induction, starting with $\mathbb{R}_+$, it follows that every homogeneous cone has a spectrahedral shadow or lifted-LMI representation.

From (6.14), we also see that if the dual of a homogeneous cone K has a spectrahedral representation then so does $SC(K, \mathcal{B})^*$. For example, if $\mathcal{V} = \mathbb{R}^n$ and $K^* = \{s \in \mathbb{R}^n : \mathcal{A}(s) \geq 0\}$ is an LMI representation of K^*, then

$$SC(K, \mathcal{B})^* = \left\{ (s, v, \beta) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R} : \begin{bmatrix} \beta & v^\top & 0 \\ v & \mathcal{H}(s) & 0 \\ 0 & 0 & \mathcal{A}(s) \end{bmatrix} \geq 0 \right\}$$

is an LMI representation of $SC(K, \mathcal{B})^*$. Since the set of homogeneous cones is closed under duality, every homogeneous cone of Siegel-rank $r \geq 2$ must arise as $SC(K, \mathcal{B})^*$ for some homogeneous cone K of Siegel-rank $r - 1$ and some homogeneous K-bilinear symmetric form $\mathcal{B}$. Therefore, by induction on r, we can establish that every homogeneous cone has a spectrahedral representation of the form $\mathcal{V} \cap \mathbb{S}^N_+$ for some $N \geq 1$ and a linear subspace $\mathcal{V}$ of $\mathbb{S}^N$. Moreover, it can be assumed that there exists a transitive subset of $\text{Aut}(\mathcal{V} \cap \mathbb{S}^N_+)$, consisting of congruences $\mathcal{R}(U) = R^\top U R$. This again follows by induction from (6.14) and the fact that the group generated by the mappings (6.18)–(6.20) forms a transitive subset of $\text{Aut}(SC(K, \mathcal{B})^*)$.

This high-level description of a recursive construction of a spectrahedral representation does not necessarily lead to an efficient representation. In Section 5 we saw a more structured and potentially more efficient canonical form for the spectrahedral representation of a homogeneous cone, due to Ishi (2015).

In Example 6.5 we enumerated the three types of cones (up to linear isomorphisms) that can be constructed as Siegel cones of $K = \mathbb{R}_+$ by choosing different bilinear forms $\mathcal{B}$. The possible homogeneous K-bilinear symmetric forms are the trivial symmetric form and the inner products $\mathcal{B}(u, v) = v^\top u$ in $\mathbb{R}^p$, for $p \geq 1$. The corresponding dual representations are $\mathcal{H}(s) = sI$, where I is a $p \times p$ identity matrix. To continue the inductive construction, one needs to characterize all homogeneous $\widehat{K}$-bilinear symmetric forms $\widehat{\mathcal{B}}$ for a Siegel cone $\widehat{K} = SC(K, \mathcal{B})$, or, equivalently, the dual representations $\widehat{\mathcal{H}}$ of $\widehat{\mathcal{B}}$, from the dual representations $\mathcal{H}$ of $\mathcal{B}$. By definition,

$$\widehat{\mathcal{H}}(s, v, \beta) = \begin{bmatrix} \beta & v^\top \\ v & \mathcal{H}(s) \end{bmatrix}$$

is a dual representation of one possible $\widehat{\mathcal{B}}$. Rothaus (1963, 1966, 1968) has characterized the possible dual representations of homogeneous $\widehat{K}$-bilinear symmetric forms. Using properties (i)–(v) of Definition 6.1 (Rothaus 1966, Lemma 3.5 and Theorem 3.7), he proves that they are all of the form

$$\widehat{\mathcal{H}}(s, v, \beta) = \begin{bmatrix} \beta I & \mathcal{C}(v)^\top \\ \mathcal{C}(v) & \widetilde{\mathcal{H}}(s) \end{bmatrix},$$

where $\widetilde{\mathcal{H}} : \mathcal{V} \to \mathbb{S}^q$ is the dual representation of a homogeneous K-bilinear form $\widetilde{\mathcal{B}} : \mathbb{R}^q \times \mathbb{R}^q \to \mathcal{V}$, and $\mathcal{C} : \mathbb{R}^p \to \mathbb{R}^{q \times m}$ is a linear matrix function. Property (v) of Definition 6.1 imposes additional constraints on $\mathcal{C}$ and $\widetilde{H}$; specifically, the matrix

equations

$$\mathcal{C}(u)^\top \mathcal{C}(v) + \mathcal{C}(v)^\top \mathcal{C}(u) = 2(u^\top v)I, \quad \mathcal{C}(\mathcal{H}(s)v) = \widetilde{\mathcal{H}}(s)\mathcal{C}(v)$$

are satisfied for all s, u, v.

7. Primal–dual interior-point methods

Theorems 4.1 and 5.3 provide a technique for 'scaling' primal and dual conic optimization problems over homogeneous matrix cones. These types of primal–dual scalings can be used to design and analyse *scale-invariant* and *primal–dual symmetric* algorithms in the sense of Tunçel (1998).

Consider a linear conic optimization problem in standard form (P) and its dual (D), with cones K and K^* in a finite-dimensional real vector space $\mathcal{W}$:

$$\begin{array}{llll}
\text{(P)} & \text{minimize} & \langle c, x \rangle & \qquad \text{(D)} \quad \text{maximize} \quad b^\top y \qquad\qquad (7.1)\\
& \text{subject to} & \mathcal{A}(x) = b & \qquad\qquad\quad\ \text{subject to} \quad \mathcal{A}^*(y) + s = c\\
& & x \in K & \qquad\qquad\qquad\qquad\qquad\qquad s \in K^*.
\end{array}$$

The linear mapping $\mathcal{A}\colon \mathcal{W} \to \mathbb{R}^m$ and the vectors $b \in \mathbb{R}^m$ and $c \in \mathcal{W}$ are given. The primal optimization variable is $x \in \mathcal{W}$; the dual variables are $y \in \mathbb{R}^m$ and $s \in \mathcal{W}$. The linear conic optimization problem is a natural extension of linear programming (the special case with $\mathcal{W} = \mathbb{R}^n$, $K = K^* = \mathbb{R}^n_+$), and has been widely used in the development of interior-point methods; see the surveys by Renegar (2001), Nemirovski and Todd (2008) and Ben-Tal and Nemirovski (2001). Advances in conic optimization algorithms and software have also enabled the creation of highly influential modelling software for convex optimization (Lofberg 2004, Grant and Boyd 2014, Diamond and Boyd 2016, Fu, Narasimhan and Boyd 2020, Sagnol and Stahlberg 2022). These modelling tools take advantage of the fact that a few different types of convex cones (the positive semidefinite cone, the second-order cone and the exponential cone) are sufficient to reformulate most convex optimization problems encountered in practice as conic optimization problems.

In this section we first review some recent literature on interior-point algorithms for conic optimization, and then comment on the special case of homogeneous cones K and K^*.

Until 2001, the research and the literature on primal–dual interior-point methods with polynomial iteration complexity were dominated by a very high level of activity concentrated on semidefinite programming, i.e. the special case $\mathcal{W} = \mathbb{S}^N$, $K = K^* = \mathbb{S}^N_+$. We must note, however, that the monograph by Nesterov and Nemirovskii (1994) already contained primal–dual interior-point algorithms, with polynomial iteration complexity, for general conic programming. Their results were based on the primal–dual symmetric, generalized Tanabe–Todd–Ye potential function: generalized because the original Tanabe–Todd–Ye potential function was proposed for linear programming and linear complementarity problems, whereas Nesterov and Nemirovski's generalization replaced the logarithmic barrier functions for the

non-negative orthant with *self-concordant* barrier functions for K, and used the Legendre–Fenchel conjugates for the dual cone K^*. Even though the algorithm and analysis were based on a primal–dual symmetric potential function, the underlying algorithm was not primal–dual symmetric (the algorithm chose different kinds of search directions and steps in the primal and dual spaces).

In a breakthrough work, Nesterov and Todd (1997, 1998) identified properties of self-concordant barriers (they called such special self-concordant barriers *self-scaled*) allowing the design and analysis of primal–dual symmetric interior-point algorithms with an outstanding number of desired properties and matching the best iteration complexity bounds. Only *symmetric cones* (those that are homogeneous and self-dual) admit self-scaled barriers. Therefore the Nesterov–Todd algorithms only apply to symmetric cones, that is, they are limited to second-order cone programming and semidefinite programming (over symmetric matrices with real entries, Hermitian matrices with complex entries or quaternion entries, and Hermitian 3×3 matrices over the octonions). It quickly became clear that generalizing all of the desired properties of Nesterov–Todd algorithms beyond symmetric cones was impossible; see a result of Nesterov in Theorem 7.2 of Güler (1997), Tunçel (1998) and Lemma 6.4 of Nesterov and Tunçel (2016). However, as explained below, Theorem 5.3 does allow some possibilities for homogeneous cones that are not self-dual.

Until recently, the literature on primal–dual symmetric interior-point algorithms for general conic optimization was quite sparse (beyond the special case of symmetric cones), but papers on this subject have been increasing in number and in their depth. A general framework for primal–dual symmetric interior-point algorithms with polynomial iteration complexity was proposed in Tunçel (2001). The same paper also showed how the theory of quasi-Newton methods and quasi-Newton-like updates can be applied to the computation of a primal–dual scaling in interior-point methods. Chares (2009) considered p-norm and power cone optimization problems, and recently Roy and Xiao (2022) proved Chares' conjecture on self-concordance of a very efficient barrier function for generalized power cones. Nesterov (2012) proposed primal–dual interior-point algorithms for general conic optimization which are based on a primal–dual scaling that approximately satisfies the conditions used in Nesterov and Todd (1997, 1998) and Tunçel (2001). Myklebust and Tunçel (2014) streamlined the computation of the primal–dual scaling in Tunçel (2001), which involved a rank-four update to a symmetric positive definite matrix, by expressing it as a composition of two rank-two quasi-Newton updates. Further, they proved that short-step path-following algorithms based on the framework of Tunçel (2001) achieve the same worst-case polynomial iteration complexity as the current best interior-point algorithms for symmetric cone programming. Skajaa and Ye (2015) also used the idea of quasi-Newton methods to design and analyse a primal–dual interior-point algorithm for general conic optimization. Their algorithm with some necessary modifications has been implemented by Papp and Yıldız (2022). Dahl and Andersen (2022) followed the framework

from Tunçel (2001) and Myklebust and Tunçel (2014), made a connection to the work of Schnabel (1983), and designed and implemented primal–dual symmetric interior-point algorithms for exponential cone programming problems.

There are interior-point algorithms for hyperbolicity cones (Renegar and Sondjaja 2014, Nesterov and Tunçel 2016), and there is further potential for interesting primal–dual algorithms utilizing self concordant barriers for sophisticated matrix cones, such as those arising from the quantum relative entropy (Faybusovich and Zhou 2022, Karimi and Tunçel 2020*a*, Fawzi and Saunderson 2022).

Chua (2009) proposed a primal–dual interior-point algorithm for conic optimization with general homogeneous cones, based on Vinberg's axioms and exploiting the underlying structure, including the transitive subset of the cone automorphism group. Chua's algorithm achieves the current best iteration complexity bound for symmetric cone programming. However, it is not primal–dual symmetric (in each iteration, the scaling is computed based on the automorphism which maps the current dual iterate to identity). Moreover, the search direction is well-defined only in a narrow neighbourhood of the central path.

By means of computational experiments, several advantages of primal–dual algorithms have been observed. In theoretical contexts there are additional justifications. See Nesterov and Todd (2002) for a justification of the usage of primal–dual central path setting, and see Todd (2009) for a geometric justification of the primal–dual scaling in the case of symmetric cones.

Next, based on the results and insights from the earlier sections, we outline a new way of computing the primal–dual scaling in new primal–dual symmetric interior-point algorithms for homogeneous cones.

We assume the cone K in (7.1) is a homogeneous matrix cone of the types discussed in Sections 3–5, i.e. a homogeneous sparse matrix cone (3.1) or the more general homogeneous matrix cone (5.1). In the first case we take $\mathcal{W} = \mathbb{S}_E^N$ in (7.1), and in the second case $\mathcal{W} = \mathcal{V}$. We let F denote the logarithmic barrier function (4.1) and (5.8). This is a ϑ-*logarithmically homogeneous self-concordant barrier* (or ϑ-*normal barrier*), with parameter $\vartheta = N$. To simplify the notation and the application to homogeneous cones in other representations (e.g. with $\mathcal{W} = \mathbb{R}^p$), we continue to use lower-case symbols x, s for the variables.

Let $x_k \in \text{int}(K)$, $s_k \in \text{int}(K^*)$ be current iterates in a primal–dual algorithm for (7.1). Theorems 4.1 and 5.3 state that there exists an automorphism $\mathcal{L}$ of K that satisfies

$$\mathcal{L}^{-1}(x_k) = \mathcal{L}^*(s_k). \tag{7.2}$$

Moreover, $\mathcal{L}^{-*} \circ \mathcal{L}^{-1} = F''(w)$, where w is the *primal–dual scaling point* defined by

$$F''(w; x_k) = s_k. \tag{7.3}$$

(If $\mathcal{W} = \mathbb{R}^p$, this equation is written more simply as $F''(w)x_k = s_k$.) If we make a change of variables $\bar{x} = \mathcal{L}^{-1}(x)$, $\bar{s} = \mathcal{L}^*(s)$, problems (P) and (D) in (7.1) are

transformed to $(\bar{\text{P}})$ and $(\bar{\text{D}})$ given by

$$
\begin{array}{llll}
(\bar{\text{P}}) & \text{minimize} & \langle \bar{c}, \bar{x} \rangle & \qquad (\bar{\text{D}}) \quad \text{maximize} \quad b^\top y \\
& \text{subject to} & \bar{\mathcal{A}}(\bar{x}) = b & \qquad \qquad \quad \text{subject to} \quad \bar{\mathcal{A}}^*(y) + \bar{s} = \bar{c} \\
& & \bar{x} \in K & \qquad \qquad \qquad \qquad \qquad \quad \bar{s} \in K^*.
\end{array}
$$

Here $\bar{\mathcal{A}} := A \circ \mathcal{L}$ and $\bar{c} := \mathcal{L}^*(c)$. In the scaled problem the current iterates x_k, s_k are mapped to the same point

$$
v_k := \mathcal{L}^{-1}(x_k) = \mathcal{L}^*(s_k). \tag{7.4}
$$

The scaled problem generalizes the *v-space formulation* from the interior-point literature for linear complementarity problems and linear optimization problems over symmetric cones (Kojima, Megiddo, Noma and Yoshise 1991, Jansen, Roos and Terlaky 1996, Sturm and Zhang 1999). Extending the definition in Tunçel (2001, Section 3), we can define a *primal–dual affine scaling direction* at x_k, s_k as the solution (d_x, d_y, d_s) of the linear system

$$
\mathcal{A}(d_x) = 0, \quad \mathcal{A}^*(d_y) + d_s = 0, \quad \mathcal{L}^{-1}(d_x) + \mathcal{L}^*(d_s) = -v_k.
$$

If F is a self-scaled barrier of a symmetric cone, equation (7.3) defines the Nesterov–Todd scaling point w (Nesterov and Todd 1997, 1998), and automatically implies

$$
F''(w; \tilde{x}_k) = \tilde{s}_k, \tag{7.5}
$$

where $\tilde{x}_k := -F'_*(s_k)$ and $\tilde{s}_k := -F'(x_k)$. For general convex cones, equation (7.3) can still be used to define primal–dual scaling points, and algorithms based on such scalings have been studied in Tunçel (2001) and Nesterov (2012). An important difference is that for general cones (and for the non-self-dual homogeneous cones discussed in this paper), equation (7.3) does not imply (7.5). In the algorithms of Tunçel (2001) and Myklebust and Tunçel (2014) this difficulty is addressed by making a rank-four update to $F''(w)$, or to an approximation of $F''(w)$, to define a positive definite self-adjoint mapping $\mathcal{H}$ that satisfies both $\mathcal{H}(x_k) = s_k$ and $\mathcal{H}(\tilde{x}_k) = \tilde{s}_k$. In these algorithms, a primal–dual search direction (d_x, d_y, d_s) is computed from $\mathcal{H}$ by solving the equation

$$
\mathcal{A}(d_x) = 0, \quad \mathcal{A}^*(d_y) + d_s = 0, \quad \mathcal{H}^{-1/2}(d_x) + \mathcal{H}^{1/2}(d_s) = -v_k + \gamma \mu \tilde{v}_k, \tag{7.6}
$$

where $\gamma \in [0, 1]$ is a centering parameter, $\mu = \langle s_k, x_k \rangle / \vartheta$, and

$$
v_k := \mathcal{H}^{1/2}(x_k) = \mathcal{H}^{-1/2}(s_k), \quad \tilde{v}_k := \mathcal{H}^{1/2}(\tilde{x}_k) = \mathcal{H}^{-1/2}(\tilde{s}_k).
$$

For homogeneous cones, different and simpler updates are possible, because the factor $\mathcal{L}$ of the primal–dual scaling matrix $F''(w) = \mathcal{L}^{-*} \circ \mathcal{L}^{-1}$ can be modified by a rank-one update to obtain a scaling $\mathcal{L}_+$ that satisfies the two conditions

$$
\mathcal{L}_+^{-1}(x_k) = \mathcal{L}_+^*(s_k), \quad \mathcal{L}_+^{-1}(\tilde{x}_k) = \mathcal{L}_+^*(\tilde{s}_k). \tag{7.7}
$$

The construction of $\mathcal{L}_+$ is similar to a quasi-Newton update, with the difference that $\mathcal{L}_+$ must satisfy the two equations (7.7), as opposed to one (secant) equation in standard quasi-Newton updates. A simplification of the updates in Myklebust and Tunçel (2014) that achieves this goal proceeds as follows. Define

$$\delta_{\mathrm{p}} := x_k - \mu \tilde{x}_k, \quad \delta_{\mathrm{d}} := s_k - \mu \tilde{s}_k,$$

where $\mu = \langle s_k, x_k \rangle / \vartheta$. General properties of ϑ-logarithmically homogeneous barriers ($\langle F'(x), x \rangle = \langle s, F_*'(s) \rangle = -\vartheta$) imply that

$$\langle s_k, \delta_{\mathrm{p}} \rangle = \langle \delta_{\mathrm{d}}, x_k \rangle = 0.$$

One can also show that

$$\langle \delta_{\mathrm{d}}, \delta_{\mathrm{p}} \rangle \geq 0$$

with equality only if δ_{p} and δ_{d} are both zero (Tunçel 2001, Corollary 4.1). We note the simple v-space expressions

$$\mathcal{L}^{-1}(\delta_{\mathrm{p}}) = v_k + \mu F_*'(v_k), \quad \mathcal{L}^*(\delta_{\mathrm{d}}) = v_k + \mu F'(v_k),$$

which follow from (7.4) and the composition properties (4.5), (4.6). The second terms on the two right-hand sides are not equal because $F_*' \neq F'$, unless the cone is symmetric. The purpose of the update of $\mathcal{L}$ is to achieve $\mathcal{L}_+^{-1}(\delta_{\mathrm{p}}) = \mathcal{L}_+^*(\delta_{\mathrm{d}})$ while preserving $\mathcal{L}_+^{-1}(x_k) = \mathcal{L}_+^*(s_k)$.

If $\delta_{\mathrm{p}} = \delta_{\mathrm{d}} = 0$, no update is needed and we take $\mathcal{L}_+ = \mathcal{L}$. Otherwise, $\langle \delta_{\mathrm{d}}, \delta_{\mathrm{p}} \rangle > 0$ and we use the (Broyden) rank-one update

$$\mathcal{L}_+(\cdot) = \mathcal{L}(\cdot) + \frac{\langle \hat{v}_k, \cdot \rangle}{\|\hat{v}_k\|^2} (\delta_{\mathrm{p}} - \mathcal{L}(\hat{v}_k)), \tag{7.8}$$

where $\hat{v}_k$ is a multiple of $\mathcal{L}^*(\delta_{\mathrm{d}})$, scaled to have norm $\|\hat{v}_k\| = \langle \delta_{\mathrm{p}}, \delta_{\mathrm{d}} \rangle^{1/2}$, that is,

$$\hat{v}_k = \frac{1}{\alpha} \mathcal{L}^*(\delta_{\mathrm{d}}), \quad \alpha = \frac{\|\mathcal{L}^*(\delta_{\mathrm{d}})\|}{\langle \delta_{\mathrm{d}}, \delta_{\mathrm{p}} \rangle^{1/2}}. \tag{7.9}$$

The mapping $\mathcal{L}_+$ is invertible with inverse

$$\mathcal{L}_+^{-1}(\cdot) = \mathcal{L}^{-1}(\cdot) + \alpha \frac{\langle \mathcal{L}^{-*}(\hat{v}_k), \cdot \rangle}{\|\hat{v}_k\|^2} (\hat{v}_k - \mathcal{L}^{-1}(\delta_{\mathrm{p}})).$$

We verify that the update $\mathcal{L}_+$ satisfies

$$\mathcal{L}_+^{-1}(x_k) = \mathcal{L}_+^*(s_k) = v_k, \quad \mathcal{L}_+^{-1}(\delta_{\mathrm{p}}) = \mathcal{L}_+^*(\delta_{\mathrm{d}}) = \hat{v}_k. \tag{7.10}$$

This is equivalent to (7.7) because δ_{p} and δ_{d} are linear combinations of $x_k, \tilde{x}_k$ and $s_k, \tilde{s}_k$, respectively. To show (7.10) we first note that

$$\langle \hat{v}_k, v_k \rangle = \frac{1}{\alpha} \langle \mathcal{L}^*(\delta_{\mathrm{d}}), v_k \rangle = \frac{1}{\alpha} \langle \delta_{\mathrm{d}}, \mathcal{L}(v_k) \rangle = \frac{1}{\alpha} \langle \delta_{\mathrm{d}}, x_k \rangle = 0.$$

Therefore, applying (7.8) to v_k gives

$$\mathcal{L}_+(v_k) = \mathcal{L}(v_k) + \frac{\langle \hat{v}_k, v_k \rangle}{\|\hat{v}_k\|^2}(\delta_p - \mathcal{L}(\hat{v}_k)) = \mathcal{L}(v_k) = x_k.$$

Applying the adjoint to s_k gives

$$\begin{aligned}
\mathcal{L}_+^*(s_k) &= \mathcal{L}^*(s_k) + \frac{\langle s_k, \delta_p - \mathcal{L}(\hat{v}_k) \rangle}{\|\hat{v}_k\|^2}\hat{v}_k \\
&= v_k - \frac{\langle \mathcal{L}^*(s_k), \hat{v}_k \rangle}{\|\hat{v}_k\|^2}\hat{v}_k \\
&= v_k - \frac{\langle v_k, \hat{v}_k \rangle}{\|\hat{v}_k\|^2}\hat{v}_k \\
&= v_k.
\end{aligned}$$

This proves the first two equations in (7.10). The equation $\mathcal{L}_+(\hat{v}_k) = \delta_p$ is immediate from (7.8). The last equation $\mathcal{L}_+^*(\delta_d) = \hat{v}_k$ follows from

$$\begin{aligned}
\mathcal{L}_+^*(\delta_d) &= \mathcal{L}^*(\delta_d) + \frac{\langle \delta_d, \delta_p \rangle - \langle \mathcal{L}^*(\delta_d), \hat{v}_k \rangle}{\|\hat{v}_k\|^2}\hat{v}_k \\
&= \alpha\hat{v}_k + \frac{\|\hat{v}_k\|^2 - \alpha\|\hat{v}_k\|^2}{\|\hat{v}_k\|^2}\hat{v}_k \\
&= \hat{v}_k.
\end{aligned}$$

The rank-one update (7.8) is the square-root form of a Broyden–Fletcher–Goldfarb–Shanno (BFGS) update (i.e. $\mathcal{H}_+ = \mathcal{L}_+ \circ \mathcal{L}_+^*$ is the BFGS update of $\mathcal{H} = \mathcal{L} \circ \mathcal{L}^*$ for the secant condition $\mathcal{H}_+(\delta_d) = \delta_p$). Many other rank-one updates will serve the same purpose. For example, in Sorensen (1982, p. 140) a closely related family of quasi-Newton updates is defined. Sorensen's updates are parametrized by the vector $\hat{v}_k$ (in our notation). Instead of (7.9) one can choose for $\hat{v}_k$ any vector that satisfies

$$\|\hat{v}_k\|^2 = \langle \delta_d, \delta_p \rangle, \quad \langle \hat{v}_k, v \rangle = 0, \tag{7.11}$$

and $\langle \mathcal{L}^*(\delta_d), \hat{v}_k \rangle \neq \langle \delta_d, \delta_p \rangle$, $\langle \hat{v}_k, \mathcal{L}^{-1}(\delta_p) \rangle \neq \langle \delta_d, \delta_p \rangle$. Then the mapping $\mathcal{L}_+$ defined by

$$\mathcal{L}_+(\cdot) = \mathcal{L}(\cdot) + \frac{\langle \mathcal{L}^*(\delta_d) - \hat{v}, \cdot \rangle}{\langle \mathcal{L}^*(\delta_d) - \hat{v}, \hat{v} \rangle}(\delta_p - \mathcal{L}(\hat{v}))$$

is invertible and satisfies (7.10). When $\alpha \neq 1$, the update (7.8) is a special case if we choose the $\hat{v}_k$ given in (7.9).

Corresponding to the updated primal–dual scaling $\mathcal{L}_+$ that satisfies (7.10), a primal–dual search direction at x_k, s_k can be defined as the solution of the equation

$$\mathcal{A}(d_x) = 0, \quad \mathcal{A}^*(d_y) + d_s = 0, \quad \mathcal{L}_+^{-1}(d_x) + \mathcal{L}_+^*(d_s) = -v_k + \gamma\mu\tilde{v}_k,$$

where $\gamma \in [0, 1]$ is a centering parameter, $\mu = \langle s_k, x_k \rangle / \vartheta$, and

$$\tilde{v}_k := \mathcal{L}_+^{-1}(\tilde{x}_k) = \mathcal{L}_+^*(\tilde{s}_k) = \frac{1}{\mu}(v_k - \hat{v}_k).$$

This primal–dual search direction simplifies the algorithms developed and analysed in the framework of Tunçel (2001), and are similar to the algorithms in Myklebust and Tunçel (2014), based on the search direction defined in (7.6). However, the primal–dual scalings for homogeneous cones described above have stronger properties than the primal–dual scalings for general convex cones discussed in Myklebust and Tunçel (2014). As mentioned earlier, instead of a rank-four update of the scaling matrix $F''(w)$, we perform a rank-one update to the factors (each of which is an automorphism of the underlying cone) in the decomposition $F''(w) = \mathcal{L}^{-*} \circ \mathcal{L}^{-1}$ to satisfy the second of the key equations (7.7). Moreover, if we apply a short-step strategy satisfying the assumptions in Myklebust and Tunçel (2014), these new algorithms achieve the same polynomial-time iteration complexity as the current best primal–dual symmetric interior-point algorithms for symmetric cone programming.

Our approach above offers more possibilities for the design and analysis of algorithms that have a significant part operating in the v-space. In addition to the references for v-space-based algorithms mentioned above, another example is the algorithm for linear programming proposed by Nesterov (2008).

8. Conclusion

Special cases of homogeneous matrix cones have been studied in the conic optimization literature. Sparse SDPs with arrow patterns are quite common, and arise, for example, in robust least-squares and robust quadratic programming (Andersen, Vandenberghe and Dahl 2010*b*, El Ghaoui and Lebret 1997, Ben-Tal, El Ghaoui and Nemirovski 2009) and in structural optimization (Kočvara 2021). They also appear in semidefinite relaxations of optimization problem with quadratic equality constraints, when the constraints involve only squares x_i^2 of variables but not cross-products $x_i x_j$ with $i \neq j$ (e.g. Boolean constraints expressed as $x_i(x_i - 1) = 0$). Sparse matrix cones with block-arrow structure are often highlighted as an important example of chordal structure (Vandenberghe and Andersen 2014, Zheng *et al.* 2021). As we have seen, their homogeneous cone property actually distinguishes them from general chordal sparse matrix cones. The matrix norm cones described in Section 5.1 have also been studied separately, for their important role in optimization problems involving the matrix trace norm and spectral norm. Except for these special cases, homogeneous matrix cones have been largely unexplored in modelling convex optimization problems and in the development of scalable algorithms.

Our approach in this paper builds on fundamental results from various disciplines: abstract algebra, graph theory, sparse matrix computation and theory, convex

conic optimization. An exciting next step in research is the development of special-ized algorithms and software which exploit the special structures we exposed here. There are many other interesting directions to be explored. Hyperbolicity cones are the next class of well-known convex cones which contain homogeneous cones as a strict subset. It may be fruitful to find a class of convex cones strictly between homogeneous cones and hyperbolicity cones providing a common generalization of homogeneous cones and cones of symmetric positive semidefinite matrices with chordal sparsity.

A. Background on homogeneous chordal graphs

This appendix contains additional details for Section 2. We first review some results by Wolk (1962, 1965) and Golumbic (1978), and then discuss the LBFS algorithm for recognizing and reordering homogeneous chordal graphs (Chu 2008). We will use the term *D-graph* when discussing Wolk's results in the next section, and the term *trivially perfect graph* when discussing Golumbic's results. After that we use the term *homogeneous chordal graph* as in the rest of the paper.

A.1. D-graphs

Wolk (1962, 1965) defines a *D-graph* or *graph with the diagonal property* as an undirected graph that does not contain P_4 or C_4 as an induced subgraph. He shows that this property characterizes the comparability graphs of rooted forests.

It is easy to show by contradiction that the absence of induced subgraphs C_4 or P_4 is a necessary condition for a graph $G = (V, E)$ to be the comparability graph of a rooted forest. Suppose the vertices u, v, w, x induce C_4 or P_4 (Figure A.1) and that there exists a rooted forest with G as its comparability graph. We use the notation $a < b$ to denote that b is an ancestor of a in the forest (b is on the unique path from a to a root of the forest). This defines a partial ordering: if $a < b$ and $b < c$, then $a < c$. There are two possible orientations for the edge $\{u, v\}$ in Figure A.1, and for each orientation there is only one possible orientation of the edges $\{v, w\}$ and $\{w, x\}$ that is compatible with the fact that $\{u, w\} \notin E$ and $\{v, x\} \notin E$. For example, if $u < v$ as in the graph in Figure A.1(a), then necessarily $v > w$, because $v < w$ would imply that $u < w$ and therefore $\{u, w\} \in E$. Now the two orientations in the figure are incompatible with a tree structure because in
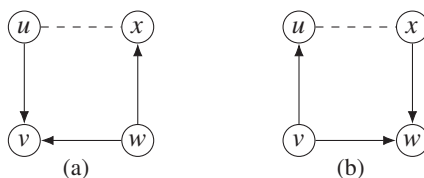
Figure A.1. The vertices u, v, w, x induce C_4 (if u and x are adjacent) or P_4 (otherwise). In (a) we assume $v > u$. Transitivity of the partial ordering and the fact that $\{u, w\} \notin E$ and $\{v, x\} \notin E$ imply that $w < v$ and $w < x$. This ordering is incompatible with a tree structure because the vertex w has two ancestors v and x that do not form an ancestor–descendant pair in the tree. In (b) we assume $u > v$. Here, transitivity of the partial ordering implies that $v < w$ and $x < w$. Now v has two ancestors u and w that do not form an ancestor–descendant pair.

each case we find a vertex (w in (a) and v in (b)) with two ancestors that are not mutually comparable (do not form an ancestor–descendant pair in the tree).

For the second part of Wolk's result (every D-graph is the comparability graph of a rooted forest), we refer to Section A.4, where we discuss how to construct a rooted forest with comparability graph G.

Wolk also established the important property that every connected component of a D-graph has a *universal vertex*, i.e. a vertex adjacent to all other vertices (Wolk 1962, lemma). This can be seen as follows. Without loss of generality we assume that G is connected. Let v be the vertex with highest degree, and denote its neighbourhood by $\mathrm{adj}(v) = \{u_1, \dots, u_k\}$, where k is the degree of v. We need to show that v is a universal vertex, i.e. $k = |V| - 1$. Assume that $k < |V| - 1$. Since the graph is connected, there exists a vertex w adjacent to one of the vertices u_i and not adjacent to v. Thus $\{w, u_i\} \in E$, $\{u_i, v\} \in E$ and $\{w, v\} \notin E$. Consider any vertex u_j, $j \neq i$. Since $\{v, u_j\} \in E$, the vertices v, u_i, w, u_j induce a P_4 or C_4 unless u_i and u_j are adjacent. Therefore, if the graph is a D-graph, u_i must be adjacent to all u_j, $j \neq i$. However, it is also adjacent to v and to w, so its degree is higher than the degree of v. This contradicts our assumption that v is a vertex with maximum degree.

It was mentioned on page 684 that this property leads to useful recursive characterization of D-graphs. One consequence of this characterization is that D-graphs are *interval graphs* (Yan *et al.* 1996). (In an interval graph the vertices represent intervals in $\mathbb{R}$; two vertices are adjacent if and only if the corresponding intervals intersect.) This follows from the construction method above, since clearly a disjoint union of interval graphs is an interval graph, and the addition of a universal vertex to an interval graph results in an interval graph. The interval graphs are a subclass of the chordal graphs (Golumbic 2004, Chapter 8).

A.2. Trivially perfect graphs

Golumbic (1978) defines a graph $G = (V, E)$ to be *trivially perfect* if $\alpha(G_W) = m(G_W)$ holds for all $W \subseteq V$, where G_W denotes the subgraph induced by W, $\alpha(G_W)$ is the stability number and $m(G_W)$ is the number of maximal cliques. To motivate the name, recall that a graph is *perfect* if $\alpha(G_W) = \bar{\chi}(G_W)$ for all W, where $\bar{\chi}(G_W)$ is the clique cover number of G_W. Clearly, $\bar{\chi}(G_W) \leq m(G_W)$, so $\alpha(G_W) = m(G_W)$ immediately implies that $\alpha(G_W) = \bar{\chi}(G_W)$.

Golumbic gives the following simple proof to show that the trivially perfect graphs are exactly the graphs that do not contain C_4 or P_4 as induced subgraphs. First, we note that $\alpha(C_4) = 2 < m(C_4) = 4$ and $\alpha(P_4) = 2 < m(P_4) = 3$, so a trivially perfect graph cannot contain C_4 or P_4. To show that the condition is sufficient, assume that G does not contain C_4 or P_4 as induced subgraphs. Suppose $\alpha(G_W) < m(G_W)$ for some $W \subseteq V$. Let S be a maximum stable set of G_W. Since $|S| = \alpha(G_W) < m(G_W)$, there exists a vertex $s \in S$ that belongs to two different maximal cliques of G_W, so we can find $x, y \in W$ with $\{s, x\} \in E$, $\{s, y\} \in E$, $\{x, y\} \notin E$. Let u be any element of $S \setminus \{s\}$ (note that $|S| = \alpha(G_W) \geq 2$ since $\{x, y\} \notin E$). Therefore $\{s, u\} \notin E$. If $\{x, u\} \in E$ and $\{u, y\} \in E$, then the vertices s, x, u, y induce a subgraph C_4. If $\{x, u\} \in E$ and $\{u, y\} \notin E$, or $\{x, u\} \notin E$ and $\{u, y\} \in E$, then they induce a subgraph P_4. We conclude that $\{x, u\} \notin E$ and $\{u, y\} \notin E$ for all $u \in S \setminus \{s\}$. However, this means that the set $(S \setminus \{s\}) \cup \{x, y\}$ is a stable set larger than S, contradicting the assumption that S is a maximum stable set.

A.3. Lexicographic breadth-first search

We now discuss Chu's algorithm for recognizing homogeneous chordal graphs and constructing a trivially perfect elimination ordering $\sigma \colon \{1, 2, \ldots, |V|\} \to V$ (Chu 2008). The algorithm can be interpreted as reversing the recursive construction of a homogeneous chordal graph via the operations of disjoint union and addition of a universal vertex. We number the vertices in the order $|V|, \ldots, 1$, that is, select $\sigma(|V|), \ldots, \sigma(1)$ in that order. At each step we find a universal vertex, give it the next available number and remove it from the graph. Note that a universal vertex in a homogeneous chordal graph is easily found as a vertex with highest degree.

Chu's algorithm maintains a list $L = (V_1, \ldots, V_K)$ of non-empty disjoint subsets of V. The vertices in each set V_i are ordered by non-decreasing degree (in G).

- Define $K = 1$ and $L = (V_1)$, with V_1 containing the elements of V sorted in order of non-decreasing degree.

- For $i = |V|, \ldots, 1$:

  1. Let v be the last vertex in V_K. Define $\sigma(i) = v$.

  2. If $\mathrm{adj}(v) \cap V_j \neq \emptyset$ for some $j < K$, terminate. The graph is not a homogeneous chordal graph.
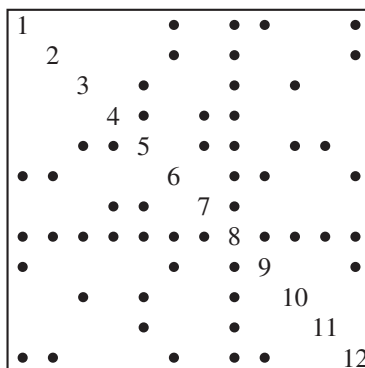
Vertex	1	2	3	4	5	6	7	8	9	10	11	12
Degree	4	3	3	3	6	5	3	11	4	3	2	5

Figure A.2. Undirected graph with vertex set $V = \{1, \ldots, 12\}$ and edges indicated by dots. The table lists the degrees of the 12 vertices.

3. Otherwise, partition $V_K \setminus \{v\}$ into two sets

$$W' = V_K \cap \text{adj}(v), \quad W = (V_K \setminus \{v\}) \setminus W'.$$

The vertices in W and W' are kept in the order of non-decreasing degree (in G). Replace the list L with

$$L := (V_1, \ldots, V_{K-1}, W, W'). \tag{A.1}$$

If W or W' is empty, remove the empty sets from L. Set K equal to the length of the new list L.

The complexity of the algorithm is $O(|E| + |V|)$.

As an example we apply the algorithm to the graph in Figure A.2. The sequence of partitions L is shown in Table A.1. The ordering found by the algorithm is

$$(\sigma(1), \ldots, \sigma(12)) = (2, 1, 9, 6, 12, 11, 4, 7, 3, 10, 5, 8), \tag{A.2a}$$
$$(\sigma^{-1}(1), \ldots, \sigma^{-1}(12)) = (2, 1, 9, 7, 11, 4, 8, 12, 3, 10, 6, 5). \tag{A.2b}$$

We now verify that the algorithm recognizes homogeneous chordal graphs (Chu 2008, Theorem 3). First, assume that G is a homogeneous chordal graph. Let $L = (V_1, \ldots, V_K)$ be the partition at the start of a cycle in the for-loop. Assume that each set V_j induces a homogeneous chordal subgraph G_{V_j}, disconnected from the other induced subgraphs G_{V_k}, $k \neq j$. If G is a homogeneous chordal graph, this assumption holds at the start of the algorithm. Since $v \in V_K$, we have $\text{adj}(v) \cap V_j = \emptyset$ for $j < K$, so the algorithm does not terminate in step 2. Since G_{V_K} is a homogeneous chordal graph, the sets W and W', which are subsets of V_K,

Table A.1. The partition L at the start of each cycle in the LBFS algorithm. We start from the vertices V sorted by degree.

i	L
12	$\boxed{11, 2, 3, 4, 7, 10, 1, 9, 6, 12, 5, 8}$
11	$\boxed{11, 2, 3, 4, 7, 10, 1, 9, 6, 12, 5}$
10	$\boxed{2, 1, 9, 6, 12}$ $\boxed{11, 3, 4, 7, 10}$
9	$\boxed{2, 1, 9, 6, 12}$ $\boxed{11, 4, 7}$ $\boxed{3}$
8	$\boxed{2, 1, 9, 6, 12}$ $\boxed{11, 4, 7}$
7	$\boxed{2, 1, 9, 6, 12}$ $\boxed{11}$ $\boxed{4}$
6	$\boxed{2, 1, 9, 6, 12}$ $\boxed{11}$
5	$\boxed{2, 1, 9, 6, 12}$
4	$\boxed{2, 1, 9, 6}$
3	$\boxed{2, 1, 9}$
2	$\boxed{2}$ $\boxed{1}$
1	$\boxed{2}$

also induce homogeneous chordal graphs. Moreover, v is a vertex with maximum degree in V_K, and therefore a universal vertex in the connected component of G_{V_K} to which it belongs. This implies that G_W is disconnected from $G_{W'}$. We conclude that the sets in the new partition computed in step 3 of the algorithm define homogeneous chordal subgraphs that are mutually disconnected. Therefore the algorithm completes the for-loop and does not terminate early.

Chu also shows that when the algorithm terminates early in step 2, a subgraph P_4 or C_4 certifying that the graph is not a homogeneous chordal graph is easily obtained (Chu 2008, Lemma 4).

Next we show that if the algorithm terminates successfully, the graph G is a homogeneous chordal graph. Let $(V_1, \ldots, V_{K-1}, W, W')$ be the partition (A.1) at the end of cycle i in the for-loop (with W and W' possibly empty). Assume that each set in this partition defines a homogeneous chordal graph, disconnected from the graphs induced by the other sets. This is certainly true for $i = 2$, since $L = (V_1, W, W')$ with $V_1 = \{\sigma(1)\}$ and $W = W' = \emptyset$. The set V_K at the beginning of cycle i can be constructed by first adding a universal vertex v to W' and then making the disjoint union with the graph induced by W. Therefore
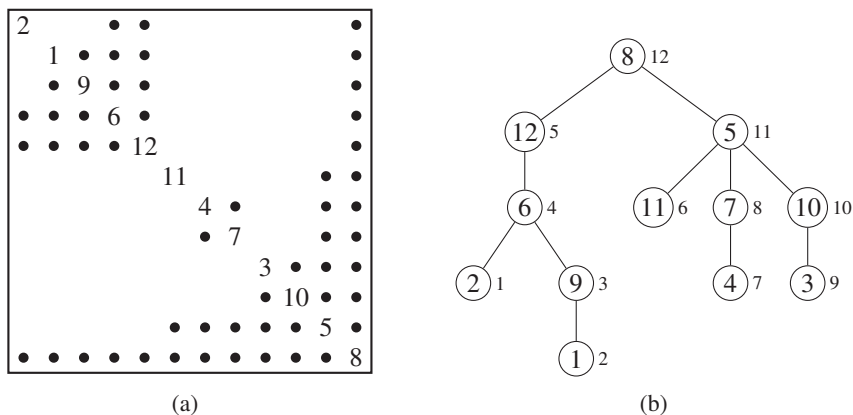
Figure A.3. (a) The graph of Figure A.2 ordered using the ordering (A.2) and (b) the corresponding elimination tree. The number next to node v in the elimination tree is $\sigma^{-1}(v)$.

V_K induces a homogeneous chordal graph, disconnected from the graphs induced by $V_1, \ldots, V_{K-1}$. We conclude that the sets in the partition $L = (V_1, \ldots, V_K)$ at the beginning of cycle i induce mutually disconnected homogeneous chordal subgraphs. Therefore, if the algorithm terminates the for-loop, the initial graph $G = (V, E)$ is a homogeneous chordal graph.

A.4. Elimination tree

We now discuss the ordering σ produced by LBFS. We use the notation

$$\mathrm{adj}^+(v) = \{w \in \mathrm{adj}(v) \colon \sigma^{-1}(w) > \sigma^{-1}(v)\},$$
$$\mathrm{adj}^-(v) = \{w \in \mathrm{adj}(v) \colon \sigma^{-1}(w) < \sigma^{-1}(v)\}$$

for the *higher* and *lower neighbourhoods* of v. We also define

$$p(v) = \arg\min \{\sigma^{-1}(w) \colon w \in \mathrm{adj}^+(v)\}$$

with the convention that $p(v) = v$ if $\mathrm{adj}^+(v)$ is empty. The graph with vertex set V and edges $\{v, p(v)\}$ for $p(v) \neq v$ is acyclic, since, by definition, $\sigma^{-1}(p(v)) > \sigma^{-1}(v)$. It is a rooted forest if we take the vertices with $p(v) = v$ as its roots. The vertex $p(v)$ is the *parent* of v in the rooted forest.

Figure A.3 illustrates these definitions for the example. In the array representation of the ordered graph, vertex v appears on the diagonal of the array in position $\sigma^{-1}(v)$. The elements of $\mathrm{adj}^+(v)$ are found as the non-zeros below the diagonal in column $\sigma^{-1}(v)$. The elements of $\mathrm{adj}^-(v)$ are the elements to the left of the diagonal in row $\sigma^{-1}(v)$. The parent $p(v)$ of v is the first non-zero below the diagonal.

The parent function can be computed by modifying the LBFS algorithm as follows (Chu 2008, p. 11). Let $v = \sigma(i)$ be the vertex selected in step 1 of cycle i of the algorithm. This is called the *pivot* (Chu 2008). The set W' in step 3 is the

Table A.2. Parent function $p(w)$ at the end of cycle $i = 12, \ldots, 1$.

	Vertex											
i	1	2	3	4	5	6	7	8	9	10	11	12
12	8	8	8	8	8	8	8	8	8	8	8	8
11	8	8	5	5	8	8	5	8	8	5	5	8
10	8	8	10	5	8	8	5	8	8	5	5	8
9	8	8	10	5	8	8	5	8	8	5	5	8
8	8	8	10	7	8	8	5	8	8	5	5	8
7	8	8	10	7	8	8	5	8	8	5	5	8
6	8	8	10	7	8	8	5	8	8	5	5	8
5	12	12	10	7	8	12	5	8	12	5	5	8
4	6	6	10	7	8	12	5	8	6	5	5	8
3	9	6	10	7	8	12	5	8	6	5	5	8
2	9	6	10	7	8	12	5	8	6	5	5	8
1	9	6	10	7	8	12	5	8	6	5	5	8

lower neighbourhood adj$^-(v)$, since it contains the vertices adjacent to v that will be numbered after v. Since $w \in$ adj$^-(v)$ if and only if $v \in$ adj$^+(w)$, we find the parent $p(w)$ as the last pivot v before w is numbered for which $w \in$ adj$^-(v)$. To construct the parent function, we initialize $p(v) = v$ for all $v \in V$ at the start of the algorithm. In step 3 of the algorithm we set $p(w) = v$ for all $w \in W'$. Table A.2 shows the value of $p(v)$ at the end of each LBFS cycle in the example.

Assume now, without loss of generality, that the graph G is connected, so the rooted forest defined by the parent function $p(v)$ is a tree, called the *elimination tree*. Consider the partition (A.1) in cycle $i = \sigma^{-1}(v)$, when v is the pivot. A vertex $w \in W' =$ adj$^-(v)$ receives $p(w) = v$. This vertex is not adjacent to any of the elements of $V_1, \ldots, V_{K-1}, W$. If, in subsequent cycles, the value of $p(w)$ is updated, the new value can only be another element in adj$^-(v)$. It follows that the vertices in adj$^-(v)$ form the subtree in the elimination tree with root v. Moreover, by definition of $W' =$ adj$^-(v)$, the vertex v is adjacent to every element in adj$^-(v)$, i.e. all the descendants of v in the elimination tree. Equivalently, every vertex w is adjacent to all its ancestors in the elimination tree (all vertices on the unique path between w and the root). Finally, if two vertices w, z do not form an ancestor–descendant pair, they were placed in different sets of the partition when their least common ancestor was the pivot. Therefore w and z are not adjacent. In summary, two vertices in G are adjacent if and only if they are comparable (form an ancestor–descendant pair) in the elimination tree. In other words, G is the comparability graph of the elimination tree. It also follows that σ is a *trivially perfect elimination ordering*,

that is, adj$^+(v)$ induces a complete subgraph of G and adj$^+(v)$ contains the vertices on the path from v to the root.

Finally, we note that placing W' last in the list (A.1) ensures that the computed ordering is a *postordering*, that is, if $\sigma^{-1}(v) = i$ and v has k descendants in the elimination tree, then the descendants w will have consecutive positions $\sigma^{-1}(w) = i - k, \ldots, i - 1$ in the ordering.

B. Matrix algorithms for homogeneous chordal sparsity

In this appendix we outline algorithms for the basic matrix operations discussed in Sections 3 and 4. The algorithms are similar to the multifrontal algorithms for matrices with chordal sparsity patterns described in Andersen, Dahl and Vandenberghe (2013) and Vandenberghe and Andersen (2014), with additional simplifications to exploit homogeneous chordal sparsity.

We consider a sparsity pattern described by a homogeneous chordal graph $G = (V, E)$ with $V = \{1, 2, \ldots, N\}$, and assume the numerical order $1, 2, \ldots, N$ is a trivially perfect elimination ordering of V. We let α_i denote the set of row indices of the lower-triangular non-zeros in column i, and let $\bar{\alpha}_i$ be the set $\{i\} \cup \alpha_i$. The parent of a non-root vertex i in the elimination tree is denoted by $p(i)$. By definition, this is the first element of α_i.

If $1, \ldots, N$ is a perfect elimination ordering of a chordal pattern, we have the important property

$$\alpha_i \subseteq \bar{\alpha}_{p(i)} \tag{B.1}$$

for all non-root vertices i. By applying this recursively, we see that the vertices indexed by α_i are on the path from vertex i to the root, i.e. $\alpha_i \subseteq \{p(i), p^2(i), \ldots, p^k(i)\}$, if k is the depth (distance to the root) of vertex i. If $1, \ldots, N$ is a trivially perfect elimination ordering of a homogeneous chordal pattern, we have equality:

$$\alpha_i = \bar{\alpha}_{p(i)}. \tag{B.2}$$

Therefore $\alpha_i = \{p(i), p^2(i), \ldots, p^k(i)\}$, the set of ancestors of vertex i in the elimination tree.

The algorithms presented in the rest of this section use a recursion on the elimination tree. A recursion in *topological order* visits each node of the elimination tree before its parent. A recursion in *inverse topological order* visits each node before its children. We also use the notation ch(i) for the set of children of node i in the elimination tree. Supernodal elimination trees can be used to formulate faster supernodal or blocked versions, but this extension will not be discussed in detail.

B.1. Cone automorphisms

Our main interest in this section is the evaluation of the linear mappings $\mathcal{L}$ and $\mathcal{L}^*$ defined in (3.4) and (3.5). We first consider two simpler operations, matrix–matrix products $L\tilde{L}$ and $L\tilde{L}^\top$, where $L, \tilde{L} \in \mathbb{T}_E^N$.

B.1.1. Products of lower-triangular matrices

Let $L, \tilde{L} \in \mathbb{T}_E^N$. Column k of the product $Y = L\tilde{L}$ can be computed by initializing the column as zero, and running the iteration

$$\begin{bmatrix} Y_{jk} \\ Y_{\alpha_j k} \end{bmatrix} := \begin{bmatrix} Y_{jk} \\ Y_{\alpha_j k} \end{bmatrix} + \tilde{L}_{jk} \begin{bmatrix} L_{jj} \\ L_{\alpha_j j} \end{bmatrix}, \quad j \in \bar{\alpha}_k.$$

Here we rely on the fact that the non-zero elements in column k of $\tilde{L}$ are in the rows indexed by $\bar{\alpha}_k$, and the non-zeros in column j of L are in the rows indexed by $\bar{\alpha}_j$. Now, the property (3.2) implies that for a trivially perfect elimination ordering, $\alpha_j \subset \alpha_k$ for $j \in \alpha_k$. Therefore the non-zeros in the kth column of Y are in the rows indexed by $\bar{\alpha}_k$. This again shows that $Y = L\tilde{L} \in \mathbb{T}_E^N$, as already noted in Theorem 3.1.

Next we consider products $Y = L\tilde{L}^\top$, where $L, \tilde{L} \in \mathbb{T}_E^N$. The matrix Y is not symmetric but has a symmetric sparsity pattern, and if E is chordal and $1, \ldots, N$ is a perfect elimination order, then the sparsity pattern of Y is E. To see this, consider the formula for the ij element of Y:

$$Y_{ij} = \sum_{k=1}^{\min\{i,j\}} L_{ik}\tilde{L}_{jk}.$$

For a perfect elimination ordering of a chordal graph, $\{i, k\} \in E$, $\{j, k\} \in E$ implies that $\{i, j\} \in E$. So if $\sum_k L_{ik}\tilde{L}_{jk}$ is non-zero, then $\{i, j\} \in E$.

An efficient method for computing Y can be formulated as a recursion on the elimination tree, using ideas from multifrontal Cholesky factorization (see Section B.3). As in the multifrontal method, we start from the equation for the $\bar{\alpha}_i \times \bar{\alpha}_i$ block:

$$\begin{bmatrix} Y_{ii} & Y_{i\alpha_i} \\ Y_{\alpha_i i} & Y_{\alpha_i \alpha_i} \end{bmatrix} = \begin{bmatrix} L_{ii} \\ L_{\alpha_i i} \end{bmatrix}\begin{bmatrix} \tilde{L}_{ii} \\ \tilde{L}_{\alpha_i i} \end{bmatrix}^\top + \sum_{k<i} \begin{bmatrix} L_{ik} \\ L_{\alpha_i k} \end{bmatrix}\begin{bmatrix} \tilde{L}_{ik} \\ \tilde{L}_{\alpha_i k} \end{bmatrix}^\top + \sum_{k>i} \begin{bmatrix} 0 \\ L_{\alpha_i k} \end{bmatrix}\begin{bmatrix} 0 \\ \tilde{L}_{\alpha_i k} \end{bmatrix}^\top. \tag{B.3}$$

We define for each vertex j a non-symmetric *update matrix*

$$U_j = -\sum_{k \in T_j} L_{\alpha_j k}\tilde{L}_{\alpha_j k}^\top, \tag{B.4}$$

where T_j is the subtree of the elimination tree rooted at node j. With this notation, the first column and row of equation (B.3), and the definition of the update matrix U_i using (B.4), can be combined in the equation

$$\begin{bmatrix} Y_{ii} & Y_{i\alpha_i} \\ Y_{\alpha_i i} & -U_i \end{bmatrix} = \begin{bmatrix} L_{ii} \\ L_{\alpha_i i} \end{bmatrix}\begin{bmatrix} \tilde{L}_{ii} \\ \tilde{L}_{\alpha_i i} \end{bmatrix}^\top - \sum_{j \in \text{ch}(i)} U_j,$$

where $\text{ch}(i)$ is the set of children of node i in the elimination tree. This recursion allows us to compute Y. We enumerate the vertices i of the elimination tree

in topological order (i.e. visiting each node before its parent, e.g. in the order $1, \ldots, N$). For each i we compute Y_{ii}, $Y_{i\alpha_i}$, $Y_{\alpha_i i}$ and U_i from column i of L and $\tilde{L}$, and from the update matrices of the children of i. After the update at vertex i the matrices U_j for $j \in \text{ch}(i)$ can be discarded.

B.1.2. Triangular scaling of symmetric matrix

We now turn to the computation of $\mathcal{L}(X) = LXL^\top$, where $X \in \mathbb{S}_E^N$ and $L \in \mathbb{T}_E^N$. The operation can be reduced to a combination of the previous cases by splitting X as $X = \tilde{L} + \tilde{L}^\top$, where $\tilde{L}$ is lower-triangular with non-zero elements $\tilde{L}_{ii} = X_{ii}/2$, $\tilde{L}_{\alpha_i i} = X_{\alpha_i i}$ for $i = 1, \ldots, n$. The other entries of $\tilde{L}$ are zero. Then $\mathcal{L}(X)$ can be written as

$$\mathcal{L}(X) = L(\tilde{L} + \tilde{L}^\top)L^\top = (L\tilde{L})L^\top + L(L\tilde{L})^\top.$$

We first compute $\hat{L} = L\tilde{L}$ column by column using

$$\begin{bmatrix} \hat{L}_{ii} \\ \hat{L}_{\alpha_i i} \end{bmatrix} = \begin{bmatrix} L_{ii} & 0 \\ L_{\alpha_i i} & L_{\alpha_i \alpha_i} \end{bmatrix} \begin{bmatrix} X_{ii}/2 \\ X_{\alpha_i i} \end{bmatrix}, \quad i = 1, \ldots, N.$$

Then $Y = \mathcal{L}(X)$ can be computed via

$$\begin{bmatrix} Y_{ii} & Y_{i\alpha_i} \\ Y_{\alpha_i i} & -U_i \end{bmatrix} = \begin{bmatrix} \hat{L}_{ii} \\ \hat{L}_{\alpha_i i} \end{bmatrix} \begin{bmatrix} L_{ii} \\ L_{\alpha_i i} \end{bmatrix}^\top + \begin{bmatrix} L_{ii} \\ L_{\alpha_i i} \end{bmatrix} \begin{bmatrix} \hat{L}_{ii} \\ \hat{L}_{\alpha_i i} \end{bmatrix}^\top - \sum_{j \in \text{ch}(i)} U_j$$

in topological order. Combining the two steps gives the formula

$$\begin{bmatrix} Y_{ii} & Y_{i\alpha_i} \\ Y_{\alpha_i i} & -U_i \end{bmatrix} = \begin{bmatrix} L_{ii} & 0 \\ L_{\alpha_i i} & L_{\alpha_i \alpha_i} \end{bmatrix} \begin{bmatrix} X_{ii}/2 \\ X_{\alpha_i i} \end{bmatrix} \begin{bmatrix} L_{ii} \\ L_{\alpha_i i} \end{bmatrix}^\top$$
$$+ \begin{bmatrix} L_{ii} \\ L_{\alpha_i i} \end{bmatrix} \begin{bmatrix} X_{ii}/2 \\ X_{\alpha_i i} \end{bmatrix}^\top \begin{bmatrix} L_{ii} & L_{\alpha_i i}^\top \\ 0 & L_{\alpha_i \alpha_i}^\top \end{bmatrix} - \sum_{j \in \text{ch}(i)} U_j$$
$$= \begin{bmatrix} L_{ii} & 0 \\ L_{\alpha_i i} & L_{\alpha_i \alpha_i} \end{bmatrix} \begin{bmatrix} X_{ii} & X_{\alpha_i i}^\top \\ X_{\alpha_i i} & 0 \end{bmatrix} \begin{bmatrix} L_{ii} & L_{\alpha_i i}^\top \\ 0 & L_{\alpha_i \alpha_i}^\top \end{bmatrix} - \sum_{j \in \text{ch}(i)} U_j, \quad \text{(B.5)}$$

which can be evaluated by a recursion in topological order. The algorithm is summarized in Algorithm B.1. The intermediate variable W and the computation in step 1 are introduced to clarify the adjoint relation with the algorithm for $\mathcal{L}^*$ below.

B.1.3. Adjoint triangular scaling of symmetric matrix

The next operation is $\mathcal{L}^*(S) = \Pi_E(L^\top S L)$. The $\bar{\alpha}_i \times \bar{\alpha}_i$ block of $Y = L^\top S L$ is

$$\begin{bmatrix} Y_{ii} & Y_{i\alpha_i} \\ Y_{\alpha_i i} & Y_{\alpha_i \alpha_i} \end{bmatrix} = \begin{bmatrix} L_{ii} & L_{\alpha_i i}^\top \\ 0 & L_{\alpha_i \alpha_i}^\top \end{bmatrix} \begin{bmatrix} S_{ii} & S_{\alpha_i i}^\top \\ S_{\alpha_i i} & S_{\alpha_i \alpha_i} \end{bmatrix} \begin{bmatrix} L_{ii} & 0 \\ L_{\alpha_i i} & L_{\alpha_i \alpha_i} \end{bmatrix}. \quad \text{(B.6)}$$

This follows from the fact that the block column of L indexed by α_i has no non-zeros outside the rows α_i. This is not true for a general chordal pattern. For a

---

Algorithm B.1. *Forward mapping $\mathcal{L}$.*

Input. A matrix $X \in \mathbb{S}_E^N$ with homogeneous chordal sparsity pattern and trivially perfect elimination ordering $\sigma = (1, 2, \ldots, N)$, a lower-triangular matrix $L \in \mathbb{T}_E^N$ and the elimination tree for σ.
Output. The matrix $Y = LXL^\top$.

1. Define a lower-triangular matrix $W \in \mathbb{T}_E^N$ with
$$W_{ii} = X_{ii}, \quad W_{\alpha_i i} = L_{\alpha_i \alpha_i} X_{\alpha_i i}, \quad i = 1, \ldots, N.$$

2. Enumerate the vertices $i = 1, 2, \ldots, N$ of the elimination tree in topological order. For each i, compute $U_i, Y_{ii}, Y_{\alpha_i i}$ using the formula
$$\begin{bmatrix} Y_{ii} & Y_{i\alpha_i} \\ Y_{\alpha_i i} & -U_i \end{bmatrix} = \begin{bmatrix} L_{ii} & 0 \\ L_{\alpha_i i} & I \end{bmatrix} \begin{bmatrix} W_{ii} & W_{\alpha_i i}^\top \\ W_{\alpha_i i} & 0 \end{bmatrix} \begin{bmatrix} L_{ii} & L_{\alpha_i i}^\top \\ 0 & I \end{bmatrix} - \sum_{j \in \mathrm{ch}(i)} U_j.$$

Algorithm B.2. *Adjoint mapping $\mathcal{L}^*$.*

Input. A matrix $S \in \mathbb{S}_E^N$ with a homogeneous chordal sparsity pattern and trivially perfect elimination ordering $\sigma = (1, 2, \ldots, N)$, a lower-triangular matrix $L \in \mathbb{T}_E^N$, and the elimination tree for σ.
Output. The matrix $Y = \Pi_E(L^\top S L)$.

1. Compute a lower-triangular matrix $W \in \mathbb{T}_E^N$ by running the following recursion in reverse topological order. For each i, compute W_{ii} and $W_{i\alpha_i}$ from
$$\begin{bmatrix} W_{ii} & W_{i\alpha_i}^\top \\ W_{i\alpha_i} & \times \end{bmatrix} = \begin{bmatrix} L_{ii} & L_{\alpha_i i}^\top \\ 0 & I \end{bmatrix} \begin{bmatrix} S_{ii} & S_{\alpha_i i}^\top \\ S_{\alpha_i i} & V_i \end{bmatrix} \begin{bmatrix} L_{ii} & 0 \\ L_{\alpha_i i} & I \end{bmatrix},$$
and define
$$V_j = \begin{bmatrix} S_{ii} & S_{\alpha_i i}^\top \\ S_{\alpha_i i} & V_i \end{bmatrix}, \quad j \in \mathrm{ch}(i).$$

2. For $i = 1, \ldots, N$, set
$$Y_{ii} = W_{ii}, \quad Y_{\alpha_i i} = L_{\alpha_i \alpha_i}^\top W_{\alpha_i i}.$$

---

general chordal pattern the expression (B.6) gives the wrong value for the 22 block $Y_{\alpha_i \alpha_i}$, although the expressions for Y_{ii} and $Y_{\alpha_i i}$ are correct. Even for a homogeneous chordal pattern, we do not actually use the 22 block, since these elements are part of other columns and we need to compute them only once. A possible implementation is shown in Algorithm B.2. The intermediate variable V_i in this algorithm is simply $S_{\alpha_i \alpha_i}$. Passing this dense matrix from nodes to their children is more efficient than retrieving $S_{\alpha_i \alpha_i}$ from a sparse matrix data structure (Andersen *et al.* 2013).

B.2. Inverse cone automorphisms

Next we consider the inverses of the mappings $\mathcal{L}$ and $\mathcal{L}^*$. Again we start with some observations about simpler operations with the inverse of a sparse lower-triangular matrix.

B.2.1. Products with inverse of lower-triangular matrix

To solve $Lx = b$, we set $x := b$ and run the iteration

$$\begin{bmatrix} x_j \\ x_{\alpha_j} \end{bmatrix} := \begin{bmatrix} 1/L_{jj} & 0 \\ -L_{\alpha_j j}/L_{jj} & I \end{bmatrix} \begin{bmatrix} x_j \\ x_{\alpha_j} \end{bmatrix}, \quad j = 1, \ldots, N. \tag{B.7}$$

The algorithm does not require chordality or homogeneous chordality, but the order of the recursion matters. If the pattern is chordal and the right-hand side b is sparse, we can simplify the iteration and iterate over a 'pruned' elimination tree, defined by the vertices k with $b_k \neq 0$ and their ancestors. This follows from (B.1): all the elements of $\bar{\alpha}_j$ are on the path from vertex j to the root, so the iteration (B.7) does not change entries outside this pruned elimination tree. In particular, if b has only one non-zero entry b_k, then in (B.7) we can iterate over the vertices $j = k, p(k), p^2(k), \ldots$, on the path from k to the root of the elimination tree.

The product $X = L^{-1}\tilde{L}$ can be computed column by column, by forward substitution. Set $X = \tilde{L}$. For each $k = 1, \ldots, n$, run the iteration

$$\begin{bmatrix} X_{jk} \\ X_{\alpha_j k} \end{bmatrix} := \begin{bmatrix} 1/L_{jj} & 0 \\ -L_{\alpha_j j}/L_{jj} & I \end{bmatrix} \begin{bmatrix} X_{jk} \\ X_{\alpha_j k} \end{bmatrix}, \quad j = k, p(k), p^2(k), \ldots.$$

This works for any chordal sparsity pattern. In general, however, the sets α_j for $j = p(k), p^2(k), \ldots$ are not subsets of α_k, so the final sparsity pattern of X_k can include non-zeros outside α_k. For a homogeneous chordal pattern and trivially perfect elimination ordering, the property (3.2) implies that the indices of all lower-triangular non-zeros of X_k are in α_k. Therefore $X = L^{-1}\tilde{L}$ has the same sparsity pattern as L and $\tilde{L}$.

Applying this with $\tilde{L} = I$, we see that the inverse L^{-1} has the same sparsity pattern as L: $L^{-1} \in \mathbb{T}_E^N$; see Theorem 5.3. This property does not hold for general chordal sparsity patterns. As a consequence, the identities

$$(L^{-1})_{\bar{\alpha}_j \bar{\alpha}_j} = L^{-1}_{\bar{\alpha}_j \bar{\alpha}_j} = \begin{bmatrix} 1/L_{jj} & 0 \\ -(1/L_{jj})L^{-1}_{\alpha_j \alpha_j}L_{\alpha_j j} & L^{-1}_{\alpha_j \alpha_j} \end{bmatrix}, \quad j = 1, \ldots, N \tag{B.8}$$

(which hold for any non-singular triangular matrix and any index set $\bar{\alpha}_j$) characterize all the non-zero elements in L^{-1}.

B.2.2. Inverse of triangular scaling

The inverse of the mapping $\mathcal{L}(X) = LXL^\top$ is $L^{-1}XL^{-\top}$. It can be evaluated via the formula (B.5) applied to the inverse of L:

$$\begin{bmatrix} Y_{ii} & Y_{i\alpha_i} \\ Y_{\alpha_i i} & -U_i \end{bmatrix} = L^{-1}_{\bar{\alpha}_i \bar{\alpha}_i} \begin{bmatrix} X_{ii} & X^\top_{\alpha_i i} \\ X_{\alpha_i i} & 0 \end{bmatrix} L^{-\top}_{\bar{\alpha}_i \bar{\alpha}_i} - \sum_{j \in \text{ch}(i)} U_j.$$

Algorithm B.3. *Inverse forward mapping $\mathcal{L}^{-1}$.*

Input. A matrix $X \in \mathbb{S}_E^N$ with homogeneous chordal sparsity pattern and trivially perfect elimination ordering $\sigma = (1, 2, \ldots, N)$, a non-singular lower-triangular matrix $L \in \mathbb{T}_E^N$, and the elimination tree for σ.
Output. The matrix $Y = L^{-1}XL^{-\top}$.

1. Enumerate the vertices $i = 1, 2, \ldots, N$ of the elimination tree in topological order. For each i, compute V_i, W_{ii}, $W_{\alpha_i i}$ using the formula

$$\begin{bmatrix} W_{ii} & W_{i\alpha_i} \\ W_{\alpha_i i} & -V_i \end{bmatrix}$$
$$= \begin{bmatrix} 1/L_{ii} & 0 \\ -L_{\alpha_i i}/L_{ii} & I \end{bmatrix} \left(\begin{bmatrix} X_{ii} & X_{\alpha_i i}^\top \\ X_{\alpha_i i} & 0 \end{bmatrix} - \sum_{j \in \text{ch}(i)} V_j \right) \begin{bmatrix} 1/L_{ii} & -L_{\alpha_i i}^\top/L_{ii} \\ 0 & I \end{bmatrix}.$$

2. For $i = 1, \ldots, N$, compute

$$Y_{ii} = W_{ii}, \quad Y_{\alpha_i i} = L_{\alpha_i \alpha_i}^{-1} W_{\alpha_i i}.$$

This can be simplified if we define update matrices $V_i = L_{\alpha_i \alpha_i} U_i L_{\alpha_i \alpha_i}^\top$ instead of U_i:

$$\begin{bmatrix} Y_{ii} & Y_{i\alpha_i} \\ Y_{\alpha_i i} & -L_{\alpha_i \alpha_i}^{-1} V_i L_{\alpha_i \alpha_i}^{-\top} \end{bmatrix} = L_{\bar{\alpha}_i \bar{\alpha}_i}^{-1} \left(\begin{bmatrix} X_{ii} & X_{\alpha_i i}^\top \\ X_{\alpha_i i} & 0 \end{bmatrix} - \sum_{j \in \text{ch}(i)} V_j \right) L_{\bar{\alpha}_i \bar{\alpha}_i}^{-\top},$$

and, using (B.8),

$$\begin{bmatrix} Y_{ii} & Y_{i\alpha_i} L_{\alpha_i \alpha_i}^\top \\ L_{\alpha_i \alpha_i} Y_{\alpha_i i} & -V_i \end{bmatrix}$$
$$= \begin{bmatrix} 1/L_{ii} & 0 \\ -L_{\alpha_i i}/L_{ii} & I \end{bmatrix} \left(\begin{bmatrix} X_{ii} & X_{\alpha_i i}^\top \\ X_{\alpha_i i} & 0 \end{bmatrix} - \sum_{j \in \text{ch}(i)} V_j \right) \begin{bmatrix} 1/L_{ii} & -L_{\alpha_i i}^\top/L_{ii} \\ 0 & I \end{bmatrix}.$$

This is summarized in Algorithm B.3.

B.2.3. Inverse of adjoint triangular scaling

Applying (B.6) with L^{-1} shows that $\mathcal{L}^{-*}(S) = \Pi_E(L^{-\top}SL^{-1})$ and that the $\bar{\alpha}_i \times \bar{\alpha}_i$ block of $Y = (\mathcal{L}^*)^{-1}(S)$ is given by

$$\begin{bmatrix} Y_{ii} & Y_{i\alpha_i} \\ Y_{\alpha_i i} & Y_{\alpha_i \alpha_i} \end{bmatrix}$$
$$= \begin{bmatrix} 1/L_{ii} & -L_{\alpha_i i}^\top L_{\alpha_i \alpha_i}^{-\top}/L_{ii} \\ 0 & L_{\alpha_i \alpha_i}^{-\top} \end{bmatrix} \begin{bmatrix} S_{ii} & S_{\alpha_i i}^\top \\ S_{\alpha_i i} & S_{\alpha_i \alpha_i} \end{bmatrix} \begin{bmatrix} 1/L_{ii} & 0 \\ -L_{\alpha_i \alpha_i}^{-1} L_{\alpha_i i}/L_{ii} & L_{\alpha_i \alpha_i}^{-1} \end{bmatrix}.$$

Algorithm B.4. *Inverse adjoint mapping* $(\mathcal{L}^*)^{-1}$.

Input. A matrix $S \in \mathbb{S}_E^N$ with homogeneous chordal sparsity pattern and trivially perfect elimination ordering $\sigma = (1, 2, \ldots, N)$, a non-singular lower-triangular matrix $L \in \mathbb{T}_E^N$, and the elimination tree for σ.
Output. The matrix $Y = \Pi_E(L^{-\top}SL^{-1})$.

1. For $i = 1, \ldots, N$, set

$$W_{ii} = S_{ii}, \quad W_{\alpha_i i} = L_{\alpha_i \alpha_i}^{-\top} S_{\alpha_i i}.$$

2. Enumerate the vertices $i = 1, \ldots, N$ in reverse topological order. For each i, compute Y_{ii} and $Y_{i\alpha_i}$ from

$$\begin{bmatrix} Y_{ii} & Y_{i\alpha_i}^{\top} \\ Y_{\alpha_i i} & \times \end{bmatrix} = \begin{bmatrix} 1/L_{ii} & -L_{\alpha_i i}^{\top}/L_{ii} \\ 0 & I \end{bmatrix} \begin{bmatrix} W_{ii} & W_{\alpha_i i}^{\top} \\ W_{\alpha_i i} & V_i \end{bmatrix} \begin{bmatrix} 1/L_{ii} & 0 \\ -L_{\alpha_i i}/L_{ii} & I \end{bmatrix},$$

and define

$$V_j = \begin{bmatrix} Y_{ii} & Y_{\alpha_i i}^{\top} \\ Y_{\alpha_i i} & V_i \end{bmatrix}, \quad j \in \mathrm{ch}(i).$$

It can be computed as shown in Algorithm B.4. Here the 'update matrices' V_j are defined as

$$V_i = Y_{\alpha_i \alpha_i} = L_{\alpha_i \alpha_i}^{-\top} S_{\alpha_i \alpha_i} L_{\alpha_i \alpha_i}^{-1}.$$

B.3. *Cholesky factorization*

Assume X is positive definite with sparsity pattern E. We define the Cholesky factorization as a factorization $X = LL^{\top}$ with L lower-triangular with positive diagonal elements. If $\sigma = (1, 2, \ldots, N)$ is a perfect elimination order, then L has the same sparsity pattern as X, i.e. $L \in \mathbb{T}_E^N$. In this section we specialize the multifrontal Cholesky factorization algorithm (Duff and Reid 1983, Liu 1990, 1992) to homogeneous chordal sparsity patterns.

Consider the $\bar{\alpha}_i \times \bar{\alpha}_i$ block of the factorization:

$$\begin{bmatrix} X_{ii} & X_{\alpha_i i}^{\top} \\ X_{\alpha_i i} & X_{\alpha_i \alpha_i} \end{bmatrix} = \begin{bmatrix} L_{ii} \\ L_{\alpha_i i} \end{bmatrix} \begin{bmatrix} L_{ii} \\ L_{\alpha_i i} \end{bmatrix}^{\top} + \sum_{k<i} \begin{bmatrix} L_{ik} \\ L_{\alpha_i k} \end{bmatrix} \begin{bmatrix} L_{ik} \\ L_{\alpha_i k} \end{bmatrix}^{\top} + \sum_{k>i} \begin{bmatrix} 0 \\ L_{\alpha_i k} \end{bmatrix} \begin{bmatrix} 0 \\ L_{\alpha_i k} \end{bmatrix}^{\top}.$$

If we consider only the first row and column in this equation, we can drop the last term on the right-hand side. In the second term, we can limit the sum to the vertices k that are proper descendants of i in the elimination tree:

$$\begin{bmatrix} X_{ii} & X_{\alpha_i i}^{\top} \\ X_{\alpha_i i} & \times \end{bmatrix} = \begin{bmatrix} L_{ii} \\ L_{\alpha_i i} \end{bmatrix} \begin{bmatrix} L_{ii} \\ L_{\alpha_i i} \end{bmatrix}^{\top} + \sum_{j \in \mathrm{ch}(i)} \sum_{k \in T_j} \begin{bmatrix} L_{ik} \\ L_{\alpha_i k} \end{bmatrix} \begin{bmatrix} L_{ik} \\ L_{\alpha_i k} \end{bmatrix}^{\top}. \tag{B.9}$$

Here T_j denotes the subtree of the elimination tree rooted at j. In the multifrontal

Algorithm B.5. *Cholesky factorization.*

Input. A matrix $X \in \mathbb{S}_E^N \cap \mathbb{S}_{++}^N$, with homogeneous chordal sparsity pattern and trivially perfect elimination ordering $\sigma = (1, 2, \ldots, N)$, and the elimination tree for σ.

Output. The Cholesky factorization $X = LL^\top$.

Enumerate the vertices $i = 1, 2, \ldots, N$ of the elimination tree in topological order. For each i, form the *frontal matrix*

$$\begin{bmatrix} F_{11} & F_{21}^\top \\ F_{21} & F_{22} \end{bmatrix} = \begin{bmatrix} X_{ii} & X_{\alpha_i i}^\top \\ X_{\alpha_i i} & 0 \end{bmatrix} + \sum_{j \in \text{ch}(i)} U_j,$$

and calculate L_{ii}, $L_{\alpha_i i}$, and the update matrix U_i from

$$L_{ii} = \sqrt{F_{11}}, \quad L_{\alpha_i i} = \frac{1}{L_{ii}} F_{21}, \quad U_i = F_{22} - L_{\alpha_i i} L_{\alpha_i i}^\top.$$

method we define for each node j in the elimination tree an *update matrix*

$$U_j = -\sum_{k \in T_j} L_{\alpha_j k} L_{\alpha_j k}^\top.$$

For a trivially perfect elimination ordering, $\alpha_j = \bar{\alpha}_i$ if $j \in \text{ch}(i)$. The last term in (B.9) is therefore equal to $-\sum_{j \in \text{ch}(i)} U_j$, and the 22 block of the entire right-hand side is $-U_i$. Therefore

$$\begin{bmatrix} X_{ii} & X_{\alpha_i i}^\top \\ X_{\alpha_i i} & -U_i \end{bmatrix} = \begin{bmatrix} L_{ii} \\ L_{\alpha_i i} \end{bmatrix} \begin{bmatrix} L_{ii} \\ L_{\alpha_i i} \end{bmatrix}^\top - \sum_{j \in \text{ch}(i)} U_j.$$

Rearranging this as

$$\begin{bmatrix} X_{ii} & X_{\alpha_i i}^\top \\ X_{\alpha_i i} & 0 \end{bmatrix} + \sum_{j \in \text{ch}(i)} U_j = \begin{bmatrix} L_{ii} \\ L_{\alpha_i i} \end{bmatrix} \begin{bmatrix} L_{ii} \\ L_{\alpha_i i} \end{bmatrix}^\top + \begin{bmatrix} 0 & 0 \\ 0 & U_i \end{bmatrix} \qquad \text{(B.10)}$$

suggests a recursive algorithm for computing the factorization, as shown in Algorithm B.5.

B.4. Maximum-determinant positive definite completion.

A matrix with a chordal sparsity pattern has a positive definite completion if and only if all completely specified principal submatrices are positive definite (Grone *et al.* 1984). In our notation, $S \in \Pi_E(\mathbb{S}_{++}^N)$ if and only if $S_{\bar{\alpha}_i \bar{\alpha}_i} > 0$ for all i. The positive definite completion with maximum determinant is the inverse of a matrix $X \in \mathbb{S}_E^N \cap \mathbb{S}_{++}^N$. If we parametrize $X = LL^\top$ by its Cholesky factor L, then $L \in \mathbb{T}_E^N$ is the solution of the non-linear equation

$$\Pi_E(L^{-\top} L^{-1}) = S.$$

Algorithm B.6. *Maximum-determinant positive definite completion.*

Input. A matrix $S \in \Pi_E(\mathbb{S}_{++}^N)$ with homogeneous chordal sparsity pattern and trivially perfect elimination ordering $\sigma = (1, 2, \dots, N)$, and the elimination tree for σ.

Output. The non-singular matrix $L \in \mathbb{T}_E^N$ that satisfies $\Pi(L^{-\top}L^{-1}) = S$.

Enumerate the vertices $i = 1, 2, \dots, N$ of the elimination tree in inverse topological order. For each i, compute

$$u = V_i^\top S_{\alpha_i i}, \quad L_{ii} = \frac{1}{(S_{ii} - u^\top u)^{1/2}}, \quad L_{\alpha_i i} = -L_{ii} V_i u.$$

Then set

$$V_j = \begin{bmatrix} L_{ii} & 0 \\ L_{\alpha_i i} & V_i \end{bmatrix}, \quad j \in \text{ch}(i).$$

The solution can be computed as follows (Andersen *et al.* 2013). Consider the $\bar{\alpha}_i \times \bar{\alpha}_i$ block of the equation $X^{-1}L = L^{-\top}$,

$$S_{\bar{\alpha}_i \bar{\alpha}_i} L_{\bar{\alpha}_i \bar{\alpha}_i} = L_{\bar{\alpha}_i \bar{\alpha}_i}^{-\top}. \tag{B.11}$$

On the right-hand side we use $(L^{-1})_{\bar{\alpha}_i \bar{\alpha}_i} = L_{\bar{\alpha}_i \bar{\alpha}_i}^{-1}$, which holds for any non-singular lower-triangular matrix and any index set $\bar{\alpha}_i$. On the left-hand side we use the fact that the block column $\bar{\alpha}_i$ of L has no zeros outside the rows indexed by $\bar{\alpha}_i$, since $1, \dots, N$ is a trivially perfect elimination ordering. An algorithm for computing the Cholesky factor L follows from the first column of the equation (B.11):

$$\begin{bmatrix} S_{ii} & S_{\alpha_i i}^\top \\ S_{\alpha_i i} & S_{\alpha_i \alpha_i} \end{bmatrix} \begin{bmatrix} L_{ii} \\ L_{\alpha_i i} \end{bmatrix} = \begin{bmatrix} 1/L_{ii} \\ 0 \end{bmatrix}.$$

The subvector $L_{\alpha_i i}/L_{ii}$ satisfies

$$\frac{1}{L_{ii}} L_{\alpha_i i} = -S_{\alpha_i \alpha_i}^{-1} S_{\alpha_i i} = -L_{\alpha_i \alpha_i} L_{\alpha_i \alpha_i}^\top S_{\alpha_i i}.$$

Substituting this in the first equation gives an expression for L_{ii}:

$$L_{ii} = (S_{ii} + S_{\alpha_i i}^\top (L_{\alpha_i i}/L_{ii}))^{-1/2} = (S_{ii} - S_{\alpha_i i}^\top L_{\alpha_i \alpha_i} L_{\alpha_i \alpha_i}^\top S_{\alpha_i i})^{-1/2}.$$

In other words, if we define $u = L_{\alpha_i \alpha_i}^\top S_{\alpha_i i}$, then

$$L_{ii} = (S_{ii} - \|u\|^2)^{-1/2}, \quad L_{\alpha_i i} = -L_{ii} L_{\alpha_i \alpha_i} u.$$

In Algorithm B.6 we define $V_i = L_{\alpha_i \alpha_i}$.

B.5. Gradient and Hessian of primal barrier

In Section 4 we introduced the function $F(X) = -\ln \det X$ as the logarithmic barrier function for the cone $K = \mathbb{S}_+^N \cap \mathbb{S}_E^N$. Define $\mathcal{L}(Y) = LYL^\top$, where L is the Cholesky

Algorithm B.7. *Projected inverse.*

Input. The Cholesky factor L of a positive definite matrix $X \in \mathbb{S}_{++}^N \cap \mathbb{S}_E^N$, with a homogeneous chordal sparsity pattern and trivially perfect elimination ordering $1, \ldots, N$, and the elimination tree for σ.

Output. The projected inverse $Y = \Pi_E(L^{-\top}L^{-1})$.

Enumerate the vertices $i = 1, 2, \ldots, N$ in inverse topological order. For each i, calculate

$$Y_{\alpha_i i} = -\frac{1}{L_{ii}}V_i L_{\alpha_i i}, \quad Y_{ii} = \frac{1}{L_{ii}}\left(\frac{1}{L_{ii}} - L_{\alpha_i i}^\top Y_{\alpha_i i}\right)$$

and define the update matrices

$$V_j = \begin{bmatrix} Y_{ii} & Y_{\alpha_i i}^\top \\ Y_{\alpha_i i} & V_i \end{bmatrix}, \quad j \in \mathrm{ch}(i).$$

---

factor of X. Then the gradient of F at X, which is given by $F'(X) = -\Pi_E(X^{-1})$, can be computed as

$$F'(X) = -\Pi_E(L^{-\top}L^{-1}) = -(\mathcal{L}^*)^{-1}(I).$$

Algorithm B.7 is Algorithm B.4 with $X = I$. It is also easily derived directly by considering the $\bar{\alpha}_i \times i$ block of the equation $X^{-1}L = L^{-\top}$, that is,

$$\begin{bmatrix} Y_{ii} & Y_{\alpha_i i}^\top \\ Y_{\alpha_i i} & Y_{\alpha_i \alpha_i} \end{bmatrix} \begin{bmatrix} L_{ii} \\ L_{\alpha_i i} \end{bmatrix} = \begin{bmatrix} 1/L_{ii} \\ 0 \end{bmatrix}. \tag{B.12}$$

We define $V_i = (X^{-1})_{\alpha_i \alpha_i}$. The Hessian of F at X is the linear mapping

$$F''(X;Y) = (\mathcal{L} \circ \mathcal{L}^*)^{-1}(Y)$$

(see (4.7)), and can be evaluated by calling Algorithms B.3 and B.4.

B.6. Gradient and Hessian of dual barrier

The barrier for the cone $\Pi_E(\mathbb{S}_+^N)$ is

$$F_*(S) = \sup_X \left(-\mathrm{Tr}(SX) - F(X)\right) = N - F(\hat{X}),$$

where $\hat{X}$ is the maximizer in the definition, i.e. the solution of the equation $\Pi_E(X^{-1}) = S$. Define $\mathcal{L}(Y) = LYL^\top$, where L is the Cholesky factor of $\hat{X}$, which can be computed by Algorithm B.6. The gradient of F_* at S is

$$F_*'(S) = -\hat{X} = -\mathcal{L}(I)$$

and can be computed by applying Algorithm B.1 with $X = I$, as shown in Algorithm B.8. The Hessian of F_* is given by $F_*''(S) = F''(\hat{X})^{-1} = \mathcal{L} \circ \mathcal{L}^*$ and can be evaluated via Algorithms B.1 and B.2.

Algorithm B.8. *Dual gradient.*

Input. The Cholesky factor L of the inverse of the maximum-determinant positive definite completion of a matrix $S \in \Pi_E(\mathbb{S}_{++}^N)$, with a homogeneous chordal sparsity pattern and trivially perfect elimination ordering $\sigma = (1, 2, \ldots, N)$, and the elimination tree for σ.
Output. The matrix $Y = LL^\top$.

Enumerate the vertices $i = 1, 2, \ldots, N$ of the elimination tree in topological order. For each i, compute $U_i, Y_{ii}, Y_{\alpha_i i}$ using the formula

$$\begin{bmatrix} Y_{ii} & Y_{i\alpha_i} \\ Y_{\alpha_i i} & -U_i \end{bmatrix} = \begin{bmatrix} L_{ii} \\ L_{\alpha_i i} \end{bmatrix} \begin{bmatrix} L_{ii} \\ L_{\alpha_i i} \end{bmatrix}^\top - \sum_{j \in \mathrm{ch}(i)} U_j.$$

References

J. Agler, J. W. Helton, S. McCullough and L. Rodman (1988), Positive semidefinite matrices with a given sparsity pattern, *Linear Algebra Appl.* **107**, 101–149.

M. S. Andersen, J. Dahl and L. Vandenberghe (2010*a*), Implementation of nonsymmetric interior-point methods for linear optimization over sparse matrix cones, *Math. Program. Comput.* **2**, 167–201.

M. S. Andersen, J. Dahl and L. Vandenberghe (2013), Logarithmic barriers for sparse matrix cones, *Optim. Methods Softw.* **28**, 396–423.

M. S. Andersen, L. Vandenberghe and J. Dahl (2010*b*), Linear matrix inequalities with chordal sparsity patterns and applications to robust quadratic optimization, in *Proceedings of the 2010 IEEE International Symposium on Computer-Aided Control System Design (CACSD)*, pp. 7–12.

S. A. Andersson and G. G. Wojnar (2004), Wishart distributions on homogeneous cones, *J. Theoret. Probab.* **17**, 781–818.

G. Averkov (2019), Optimal size of linear matrix inequalities in semidefinite approaches to polynomial optimization, *SIAM J. Appl. Algebra Geom.* **3**, 128–151.

J. J. Bartholdi, III (1981/82), A good submatrix is hard to find, *Oper. Res. Lett.* **1**, 190–193.

A. Ben-Tal and A. Nemirovski (2001), *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*, SIAM.

A. Ben-Tal, L. El Ghaoui and A. Nemirovski (2009), *Robust Optimization*, Princeton University Press.

S. J. Benson, Y. Ye and X. Zhang (2000), Solving large-scale sparse semidefinite programs for combinatorial optimization, *SIAM J. Optim.* **10**, 443–461.

Å. Björck (1996), *Numerical Methods for Least Squares Problems*, SIAM.

J. R. S. Blair and B. Peyton (1993), An introduction to chordal graphs and clique trees, in *Graph Theory and Sparse Matrix Computation* (A. George *et al.*, eds), Springer, pp. 1–29.

I. Boutouria, A. Hassairi and H. Massam (2011), Extension of the Olkin and Rubin characterization to the Wishart distribution on homogeneous cones, *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **14**, 591–611.

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

S. Burer (2003), Semidefinite programming in the space of partial positive semidefinite matrices, *SIAM J. Optim.* **14**, 139–172.

P. R. Chares (2009), Cones and interior-point algorithms for structured convex optimization involving powers and exponentials. PhD thesis, Université Catholique de Louvain.

S. Chaudhuri, M. Drton and T. S. Richardson (2007), Estimation of a covariance matrix with zeros, *Biometrika* **94**, 199–216.

F. P. M. Chu (2008), A simple linear time certifying LBFS-based algorithm for recognizing trivially perfect graphs and their complement, *Inform. Process. Lett.* **107**, 7–12.

C. B. Chua (2003), Relating homogeneous cones and positive definite cones via T-algebras, *SIAM J. Optim.* **14**, 500–506.

C. B. Chua (2009), A T-algebraic approach to primal–dual interior-point algorithms, *SIAM J. Optim.* **20**, 503–523.

C. B. Chua and L. Tunçel (2008), Invariance and efficiency of convex representations, *Math. Program.* **111**, 113–140.

D. G. Corneil (2004), Lexicographic breadth first search: A survey, in *Graph-Theoretic Concepts in Computer Science*, Vol. 3353 of Lecture Notes in Computer Science, Springer, pp. 1–19.

J. Dahl and E. D. Andersen (2022), A primal–dual interior-point algorithm for nonsymmetric exponential-cone optimization, *Math. Program.* **194**, 341–370.

S. Diamond and S. Boyd (2016), CVXPY: A Python-embedded modeling language for convex optimization, *J. Mach. Learn. Res.* **17**, 1–5.

M. Drton and T. S. Richardson (2008), Graphical methods for efficient likelihood inference in Gaussian covariance models, *J. Mach. Learn. Res.* **9**, 893–914.

I. S. Duff and J. K. Reid (1983), The multifrontal solution of indefinite sparse symmetric linear equations, *ACM Trans. Math. Softw.* **9**, 302–325.

I. S. Duff, A. M. Erisman and J. K. Reid (2017), *Direct Methods for Sparse Matrices*, Oxford University Press.

L. El Ghaoui and H. Lebret (1997), Robust solutions to least-squares problems with uncertain data, *SIAM J. Matrix Anal. Appl.* **18**, 1035–1064.

E. S. El-Mallah and C. J. Colbourn (1988), The complexity of some edge deletion problems, *IEEE Trans. Circuits Systems* **35**, 354–362.

H. Fawzi (2020), Lifts of convex sets, in *Sum of Squares: Theory and Applications*, Vol. 77 of Proceedings of Symposia in Applied Mathematics, American Mathematical Society, pp. 37–57.

H. Fawzi and J. Saunderson (2022), Optimal self-concordant barriers for quantum relative entropies. Available at arXiv:2205.04581.

L. Faybusovich (2002), On Nesterov's approach to semi-infinite programming, *Acta Appl. Math.* **74**, 195–215.

L. Faybusovich and C. Zhou (2022), Long-step path-following algorithm for quantum information theory: Some numerical aspects and applications, *Numer. Algebra Control Optim.* **12**, 445–467.

A. Fu, B. Narasimhan and S. Boyd (2020), CVXR: An R package for disciplined convex optimization, *J. Statist. Softw.* **94**, 1–34.

M. Fukuda, M. Kojima, K. Murota and K. Nakata (2000/01), Exploiting sparsity in semidefinite programming via matrix completion, I: General framework, *SIAM J. Optim.* **11**, 647–674.

D. R. Fulkerson and O. Gross (1965), Incidence matrices and interval graphs, *Pacific J. Math.* **15**, 835–855.

A. George and J. W. H. Liu (1981), *Computer Solution of Large Sparse Positive Definite Systems*, Prentice-Hall.

S. Gindikin (1992), *Tube Domains and the Cauchy Problem*, Vol. 111 of Translations of Mathematical Monographs, American Mathematical Society.

G. H. Golub and R. J. Plemmons (1980), Large-scale geodetic least-squares adjustment by dissection and orthogonal decomposition, *Linear Algebra Appl.* **34**, 3–27.

M. C. Golumbic (1978), Trivially perfect graphs, *Discrete Math.* **24**, 105–107.

M. C. Golumbic (2004), *Algorithmic Graph Theory and Perfect Graphs*, second edition, Elsevier.

J. Gouveia, P. A. Parrilo and R. R. Thomas (2013), Lifts of convex sets and cone factorizations, *Math. Oper. Res.* **38**, 248–264.

M. Grant and S. Boyd (2014), CVX: Matlab software for disciplined convex programming, version 2.1. Available at http://cvxr.com/cvx.

A. Griewank and P. L. Toint (1984), On the existence of convex decompositions of partially separable functions, *Math. Program.* **28**, 25–49.

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

O. Güler (1996), Barrier functions in interior point methods, *Math. Oper. Res.* **21**, 860–885.

O. Güler (1997), Hyperbolic polynomials and interior point methods for convex programming, *Math. Oper. Res.* **22**, 350–377.

O. Güler and L. Tunçel (1998), Characterization of the barrier parameter of homogeneous convex cones, *Math. Program.* **81**, 55–76.

M. Habib, R. McConnell, C. Paul and L. Viennot (2000), Lex-BFS and partition refinement, with applications to transitive orientation, interval graph recognition and consecutive ones testing, *Theoret. Comput. Sci.* **234**, 59–84.

J. W. Helton and J. Nie (2010), Semidefinite representation of convex sets, *Math. Program.* **122**, 21–64.

J. W. Helton and V. Vinnikov (2007), Linear matrix inequality representation of sets, *Commun. Pure Appl. Math.* **60**, 654–674.

H. Ishi (2013), On a class of homogeneous cones consisting of real symmetric matrices, *Josai Math. Monogr.* **6**, 71–80.

H. Ishi (2015), Matrix realization of a homogeneous cone, in *Geometric Science of Information* (F. Nielsen and F. Barbaresco, eds), Springer International Publishing, pp. 248–256.

H. Ishi (2016), Explicit formula of Koszul–Vinberg characteristic functions for a wide class of regular convex cones, *Entropy* **18**, 383.

B. Jansen, C. Roos and T. Terlaky (1996), A polynomial primal–dual Dikin-type algorithm for linear programming, *Math. Oper. Res.* **21**, 341–353.

M. Karimi and L. Tunçel (2020*a*), Domain-driven solver (DDS) version 2.0: A MATLAB-based software package for convex optimization problems in domain-driven form. Available at arXiv:1908.03075.

M. Karimi and L. Tunçel (2020*b*), Primal–dual interior-point methods for domain-driven formulations, *Math. Oper. Res.* **45**, 591–621.

K. Khare and B. Rajaratnam (2011), Wishart distributions for decomposable covariance graph models, *Ann. Statist.* **39**, 514–555.

K. Khare and B. Rajaratnam (2012), Sparse matrix decompositions and graph characterizations, *Linear Algebra Appl.* **437**, 932–947.

M. Kočvara (2021), Decomposition of arrow type positive semidefinite matrices with application to topology optimization, *Math. Program.* **190**, 105–134.

M. Kojima, N. Megiddo, T. Noma and A. Yoshise (1991), *A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems*, Vol. 538 of Lecture Notes in Computer Science, Springer.

L. Kong, L. Tunçel and N. Xiu (2012), Existence and uniqueness of solutions for homogeneous cone complementarity problems, *J. Optim. Theory Appl.* **153**, 357–376.

L. S. Lasdon (2002), *Optimization Theory for Large Systems*, Dover Publications. First published in 1970 by the Macmillan Company.

G. Letac and H. Massam (2007), Wishart distributions for decomposable graphs, *Ann. Statist.* **35**, 1278–1323.

J. W. H. Liu (1990), The role of elimination trees in sparse factorization, *SIAM J. Matrix Anal. Appl.* **11**, 134–172.

J. W. H. Liu (1992), The multifrontal method for sparse matrix solution: Theory and practice, *SIAM Review* **34**, 82–109.

J. W. H. Liu, E. G. Ng and B. W. Peyton (1993), On finding supernodes for sparse matrix computations, *SIAM J. Matrix Anal. Appl.* **14**, 242–252.

J. Lofberg (2004), YALMIP: A toolbox for modeling and optimization in MATLAB, in *2004 IEEE International Conference on Robotics and Automation (IEEE Cat. No. 04CH37508)*, IEEE, pp. 284–289.

T. Myklebust and L. Tunçel (2014), Interior-point algorithms for convex optimization based on primal–dual metrics. Available at arXiv:1411.2129.

A. Natanzon, R. Shamir and R. Sharan (2000), A polynomial approximation algorithm for the minimum fill-in problem, *SIAM J. Comput.* **30**, 1067–1079.

A. Nemirovski (2007), Advances in convex optimization: Conic programming, in *International Congress of Mathematicians Madrid 2006*, Vol. I: *Plenary Lectures and Ceremonies*, European Mathematical Society, pp. 413–444.

A. S. Nemirovski and M. J. Todd (2008), Interior-point methods for optimization, *Acta Numer.* **17**, 191–234.

Y. Nesterov (2000), Squared functional systems and optimization problems, in *High Performance Optimization Techniques* (J. Frenk *et al.*, eds), Kluwer Academic, pp. 405–440.

Y. Nesterov (2008), Parabolic target space and primal–dual interior-point methods, *Discrete Appl. Math.* **156**, 2079–2100.

Y. Nesterov (2012), Towards non-symmetric conic optimization, *Optim. Methods Softw.* **27**, 893–917.

Y. Nesterov and A. Nemirovskii (1994), *Interior-Point Polynomial Algorithms in Convex Programming*, Vol. 13 of SIAM Studies in Applied Mathematics, SIAM.

Y. Nesterov and M. J. Todd (1997), Self-scaled barriers and interior-point methods for convex programming, *Math. Oper. Res.* **22**, 1–42.

Y. Nesterov and M. J. Todd (1998), Primal–dual interior-point methods for self-scaled cones, *SIAM J. Optim.* **8**, 324–364.

Y. Nesterov and M. J. Todd (2002), On the Riemannian geometry defined by self-concordant barriers and interior-point methods, *Found. Comput. Math.* **2**, 333–361.

Y. Nesterov and L. Tunçel (2016), Local superlinear convergence of polynomial-time interior-point methods for hyperbolicity cone optimization problems, *SIAM J. Optim.* **26**, 139–170.

D. Papp and F. Alizadeh (2013), Semidefinite characterization of sum-of-squares cones in algebras, *SIAM J. Optim.* **23**, 1398–1423.

D. Papp and S. Yıldız (2022), Alfonso: Matlab package for nonsymmetric conic optimization, *INFORMS J. Comput.* **34**, 11–19.

J. Pearl and N. Wermuth (1994), When can association graphs admit a causal interpretation?, in *Selecting Models from Data*, Springer, pp. 205–214.

J. Renegar (2001), *A Mathematical View of Interior-Point Methods in Convex Optimization*, MPS/SIAM Series on Optimization, SIAM and Mathematical Programming Society (MPS).

J. Renegar and M. Sondjaja (2014), A polynomial-time affine-scaling method for semidefinite and hyperbolic programming. Available at arXiv:1410.6734.

R. T. Rockafellar (1970), *Convex Analysis*, Princeton University Press.

D. J. Rose, R. E. Tarjan and G. S. Lueker (1976), Algorithmic aspects of vertex elimination on graphs, *SIAM J. Comput.* **5**, 266–283.

O. S. Rothaus (1963), The construction of homogeneous convex cones, *Bull. Amer. Math. Soc.* **69**, 248–250.

O. S. Rothaus (1966), The construction of homogeneous convex cones, *Ann. of Math. (2)* **83**, 358–376.

O. S. Rothaus (1968), Correction to: 'The construction of homogeneous convex cones', *Ann. of Math. (2)* **87**, 399.

S. Roy and L. Xiao (2022), On self-concordant barriers for generalized power cones, *Optim. Lett.* **16**, 681–694.

G. Sagnol and M. Stahlberg (2022), PICOS: A Python interface to conic optimization solvers, *J. Open Source Software* **7**, 3915.

M. A. Saunders (1972), Product form of the Cholesky factorization for large-scale linear programming. Technical report STAN-CS-72-301, Stanford University.

C. Scheiderer (2018), Spectrahedral shadows, *SIAM J. Appl. Algebra Geom.* **2**, 26–44.

R. B. Schnabel (1983), Quasi-Newton methods using multiple secant equations. Technical report, University of Colorado at Boulder.

A. Skajaa and Y. Ye (2015), A homogeneous interior-point algorithm for nonsymmetric convex conic optimization, *Math. Program.* **150**, 391–422.

D. C. Sorensen (1982), Collinear scaling and sequential estimation in sparse optimization algorithms, in *Algorithms and Theory in Filtering and Control* (D. C. Sorensen and R. J.-B. Wets, eds), Vol. 18 of Mathematical Programming Studies, Springer, pp. 135–159.

G. Srijuntongsiri and S. Vavasis (2004), A fully sparse implementation of a primal–dual interior-point potential reduction method for semidefinite programming. Available at arXiv:cs:0412009.

J. F. Sturm and S. Zhang (1999), Symmetric primal–dual path-following algorithms for semidefinite programming, *Appl. Numer. Math.* **29**, 301–315.

R. E. Tarjan and M. Yannakakis (1984), Simple linear-time algorithms to test chordality of graphs, test acyclicity of hypergraphs, and selectively reduce acyclic hypergraphs, *SIAM J. Comput.* **13**, 566–579.

M. J. Todd (2009), Largest dual ellipsoids inscribed in dual cones, *Math. Program.* **117**, 425–434.

V. A. Truong and L. Tunçel (2004), Geometry of homogeneous convex cones, duality mapping, and optimal self-concordant barriers, *Math. Program.* **100**, 295–316.

L. Tunçel (1998), Primal–dual symmetry and scale invariance of interior-point algorithms for convex optimization, *Math. Oper. Res.* **23**, 708–718.

L. Tunçel (2001), Generalization of primal–dual interior-point methods to convex optimization problems in conic form, *Found. Comput. Math.* **1**, 229–254.

L. Vandenberghe and M. S. Andersen (2014), Chordal graphs and semidefinite optimization, *Found. Trends Optim.* **1**, 241–433.

E. B. Vinberg (1965*a*), Structure of the group of automorphisms of a homogeneous convex cone, *Trudy Moskov. Mat. Obšč.* **13**, 56–83.

E. B. Vinberg (1965*b*), The theory of homogeneous cones, *Trans. Moscow Math. Soc.* **12**, 340–403.

E. S. Wolk (1962), The comparability graph of a tree, *Proc. Amer. Math. Soc.* **13**, 789–795.

E. S. Wolk (1965), A note on 'The comparability graph of a tree', *Proc. Amer. Math. Soc.* **16**, 17–20.

T. Yamasaki and T. Nomura (2015), Realization of homogeneous cones through oriented graphs, *Kyushu J. Math.* **69**, 11–48.

J.-H. Yan, J.-J. Chen and G. J. Chang (1996), Quasi-threshold graphs, *Discrete Appl. Math.* **69**, 247–255.

M. Yannakakis (1981), Computing the minimum fill-in is NP-complete, *SIAM J. Algebraic Discrete Methods* **2**, 77–79.

Y. Zheng, G. Fantuzzi and A. Papachristodoulou (2021), Chordal and factor-width decompositions for scalable semidefinite and polynomial optimization, *Annu. Rev. Control* **52**, 243–279.