On approximations, complexity, and applications for copositive programming

Gijben, Luuk

Chapter 4

Scaling relationship between the copositive cone and Parrilo's first level approximation [1]

[1]Published as [DDGH13b] Peter J.C. Dickinson, Mirjam Dür, Luuk Gijben and Roland Hildebrand. Scaling relationship between the copositive cone and Parrilo's first level approximation. *Optimization Letters*, 7(8):1669-1679, 2013.

For the purpose of this chapter recall from (1.14) that the hierarchy of inner approximations for the copositive cone introduced by Parrilo [Par00] is defined as

$$\mathcal{K}_n^r = \Big\{ A \in \mathbb{S}^n \mid \Big(\sum_{i=1}^{n} x_i^2 \Big)^r (\mathbf{x} \circ \mathbf{x})^\mathsf{T} A (\mathbf{x} \circ \mathbf{x}) \text{ is SOS } \Big\}.$$

For $\mathcal{K}_n^1$, the combined proofs of Parrilo [Par00] (sufficient) and Bomze and de Klerk [BdK02] (necessary) show that $A \in \mathcal{K}_n^1$ if and only if the following system of LMIs has a feasible solution $M^1, \dots, M^n$:

$$A - M^i \in \mathcal{S}_+^n \qquad i = 1, \dots, n \tag{4.1a}$$

$$(M^i)_{ii} = 0 \qquad i = 1, \dots, n \tag{4.1b}$$

$$(M^i)_{jj} + 2(M^j)_{ij} = 0 \qquad i, j = 1, \dots, n \quad \text{s.t. } i \neq j \tag{4.1c}$$

$$(M^i)_{jk} + (M^j)_{ik} + (M^k)_{ij} \geq 0 \qquad i, j, k = 1, \dots, n \quad \text{s.t. } i < j < k. \tag{4.1d}$$

Note that these LMIs can be directly obtained from the alternative definition (1.18) of $\mathcal{K}_n^r$ that was given in [PVZ07].

Furthermore recall again that $\mathcal{K}_n^0 = \mathcal{S}_+^n + \mathcal{N}^n$ and that $\mathcal{COP}^n = \mathcal{K}_n^0$ if and only if $n \leq 4$. In the 5×5 case, as was noted before, the Horn-matrix (1.10) is in $\mathcal{COP}^5 \setminus \mathcal{K}_5^0$, moreover an explicit certificate was given by Parrilo [Par00] that shows that the Horn matrix in fact is in $\mathcal{K}_5^1$. The natural question that was then posed was to determine the smallest n for which $\mathcal{COP}^n \neq \mathcal{K}_n^1$. In this chapter, we answer this question and show that in fact already $\mathcal{COP}^5 \neq \mathcal{K}_5^r$ for all $r \in \mathbb{Z}_+$. A central ingredient of the proof that we will present below is the observation that if we are given a diagonal matrix D with strictly positive diagonal, then for any matrix class $\mathcal{X} \in \{\mathcal{COP}^n, \mathcal{S}_+^n, \mathcal{N}^n, \mathcal{S}_+^n + \mathcal{N}^n\}$ we have that $A \in \mathcal{X} \;\Leftrightarrow\; DAD \in \mathcal{X}$, see Propositions 1.2 and 1.3. However, as we will see later on this property does not hold for $\mathcal{X} = \mathcal{K}_n^r$ when $r \geq 1$. We will show in fact that for any matrix $A \in \mathcal{COP}^n \setminus (\mathcal{S}_+^n + \mathcal{N}^n)$ and for any $r \in \mathbb{Z}_+$ there exists a diagonal matrix D with strictly positive diagonal, such that $DAD \notin \mathcal{K}_n^r$. In general we can only show that such scaling matrices D exist, but we do not know how to construct them. For the case when $r = 1$

however we give an explicit way to construct a scaling matrix that scales a given matrix in $\mathcal{K}_n^1 \setminus (\mathcal{S}_+^n + \mathcal{N}^n)$ out of $\mathcal{K}_n^1$. Furthermore we provide some remarks concerning scalings in the opposite direction, that is scaling such that $DAD \in \mathcal{K}_n^r$ for a matrix $A \in \mathcal{COP}^n \setminus \mathcal{K}_n^r$.

For the 5×5 case, we will show the more surprising result that for any matrix $A \in \mathcal{COP}^5$ scaling it in such a way that $(DAD)_{ii} \in \{0, 1\}$ will yield $DAD \in \mathcal{K}_5^1$. Our main result of this chapter (Theorem 4.18) is a complete characterization of $\mathcal{COP}^5$ in terms of $\mathcal{K}_5^1$. We will conclude this chapter by formulating several conjectures and open problems regarding scalings of copositive matrices with respect to the hierarchy of Parrilo cones.

Notation

We recall that given a vector $\mathbf{d} \in \mathbb{R}^n$, we denote by $\mathrm{Diag}(\mathbf{d})$ the diagonal matrix with the entries of $\mathbf{d}$ on its diagonal. Conversely, given a matrix A, we denote by $\mathrm{diag}(A)$ the vector of diagonal entries of A. We shall denote the set of scalings by

$$\mathcal{D} := \{\mathrm{Diag}(\mathbf{d}) \mid \mathbf{d} \in \mathbb{R}_{++}^n\}.$$

During this chapter we will use the modulo operator in such a way that it maps to $\{1, \ldots, n\}$ rather than $\{0, \ldots, n-1\}$. That is, $n \bmod n \equiv 0 \bmod n \equiv n$, and $i \bmod n \equiv i$ for $0 < i < n$. The reason for this is that it will make notation much more convenient later on improving readability.

4.1 Scaling a matrix out of $\mathcal{K}_n^r$

In this section we will show that for $n \geq 5$ we have that $\mathcal{K}_n^r \neq \mathcal{COP}^n$ for all $r \geq 0$. Instead of giving a specific example of a matrix in $\mathcal{COP}^n$ but not in $\mathcal{K}_n^r$, we will show that in fact any matrix $A \in \mathcal{COP}^n \setminus (\mathcal{S}_+^n + \mathcal{N}^n)$ can be "scaled out" of $\mathcal{K}_n^r$. As $\mathcal{S}_+^n + \mathcal{N}^n \neq \mathcal{COP}^n$ if and only if $n \geq 5$, this will then give us the required result.

First we shall show an auxiliary result on the relationship between the cones $\mathcal{K}_n^0$ and $\mathcal{K}_n^r$ for $r \geq 1$.

Lemma 4.1. *Let $n \geq 1$ and $r \geq 0$ be integers. Then*

$$\{A \in \mathbb{S}^n \mid DAD \in \mathcal{K}_n^r \text{ for all } D \in \mathcal{D}\} = \mathcal{K}_n^0.$$

Proof. Since the cone $\mathcal{K}_n^0 = \mathcal{S}_+^n + \mathcal{N}^n$ is invariant under arbitrary scalings, we have that $A \in \mathcal{K}_n^0$ implies $DAD \in \mathcal{K}_n^0 \subset \mathcal{K}_n^r$ for all $D \in \mathcal{D}$. This proves one inclusion, and the whole statement for $r = 0$.

We now prove the other inclusion for $r \geq 1$. Let $A \in \mathbb{S}^n$ be such that $DAD \in \mathcal{K}_n^r$ for all $D \in \mathcal{D}$. Then for all $d_1, \ldots, d_n > 0$, the polynomial

$$\left(\sum_{i,j=1}^{n} A_{ij} d_i d_j x_i^2 x_j^2\right)\left(\sum_{i=1}^{n} x_i^2\right)^r$$

is a sum of squares of polynomials in the variables $x_1, \ldots, x_n$.

Equivalently, $\left(\sum_{i,j=1}^n A_{ij} z_i^2 z_j^2\right)\left(\sum_{i=1}^n d_i^{-1} z_i^2\right)^r$ is a sum of squares of polynomials in the variables $z_i = \sqrt{d_i} x_i$, $i = 1, \ldots, n$. Let us now fix $d_1 = 1$ and let $d_i \to +\infty$ for $i > 1$. Since the cone of sums of squares polynomials is closed (this result is attributed to Robinson [Rob73]; a more accessible reference where a proof can be found is [Lau09, Section 3.8]), the limit polynomial $\left(\sum_{i,j=1}^n A_{ij} z_i^2 z_j^2\right)\left(z_1^2\right)^r$ must also be a sum of squares of polynomials in $z_1, \ldots, z_n$, say $\left(\sum_{i,j=1}^n A_{ij} z_i^2 z_j^2\right) z_1^{2r} = \sum_{k=1}^N q_k^2(\mathbf{z})$. But then for all k we must have $q_k(\mathbf{z}) = 0$ whenever $z_1 = 0$. It follows that z_1 can be factored out of q_k, i.e., $\left(\sum_{i,j=1}^n A_{ij} z_i^2 z_j^2\right) z_1^{2(r-1)}$ is also a sum of squares. After repeatedly carrying out this factoring out process, we arrive at the conclusion that $\left(\sum_{i,j=1}^n A_{ij} z_i^2 z_j^2\right)$ is a sum of squares, i.e., $A \in \mathcal{K}_n^0$. This concludes the proof. □

Theorem 4.2. *For any matrix $A \in \mathcal{COP}^n \setminus (\mathcal{S}_+^n + \mathcal{N}^n)$ and any $r \geq 0$, there exists $D \in \mathcal{D}$ such that $DAD \in \mathcal{COP}^n \setminus \mathcal{K}_n^r$.*

Proof. For any $A \in \mathcal{COP}^n$ and $D \in \mathcal{D}$ we have that $DAD \in \mathcal{COP}^n$. Therefore we need only show that for any $A \in \mathcal{COP}^n \setminus (\mathcal{S}_+^n + \mathcal{N}^n)$ there exists $D \in \mathcal{D}$ such that $DAD \notin \mathcal{K}_n^r$.
Assume that such a D does not exist. Then for all $D \in \mathcal{D}$ we have $DAD \in \mathcal{K}_n^r$, and by Lemma 4.1 it follows that $A \in \mathcal{K}_n^0 = \mathcal{S}_+^n + \mathcal{N}^n$, a contradiction. □

Corollary 4.3. *Let $r \geq 0$ be an integer. Then $\mathcal{COP}^n = \mathcal{K}_n^r$ if and only if $n \leq 4$.*

Although we now know from Theorem 4.2 that for every matrix $A \in \mathcal{COP}^n \setminus \mathcal{K}_n^0$ there exists a scaling matrix in $\mathcal{D}$ that scales A out of any cone $\mathcal{K}_n^r$ for $r \geq 1$, $n \geq 5$, finding such a scaling matrix for an arbitrary matrix A is not obvious. For practical purposes however it can be interesting to be able to construct such explicit examples. A result in this direction that generates scalings that scale matrices in $\mathcal{COP}^n \setminus \mathcal{K}_n^0$ out of $\mathcal{K}_n^1$ is presented next. This result comes from correspondence with Peter J.C. Dickinson.

Theorem 4.4. *Let $A \in \mathcal{COP}^n \setminus (\mathcal{S}_+^n + \mathcal{N}^n)$ and let $V^1 \in \operatorname{int}(\mathcal{S}_+^n \cap \mathcal{N}^n)$ be such that $\langle A, V^1 \rangle < 0$. For $i = 2, \ldots, n$ define*

$$V^i = \sum_{l=2}^{n} \left(\sqrt{\frac{(V^1)_{1i}}{n-1}}\, \mathbf{e}_1 + \sqrt{\frac{n-1}{(V^1)_{1i}}}\, (V^1)_{il}\mathbf{e}_l \right) \left(\sqrt{\frac{(V^1)_{1i}}{n-1}}\, \mathbf{e}_1 + \sqrt{\frac{n-1}{(V^1)_{1i}}}\, (V^1)_{il}\mathbf{e}_l \right)^{\mathsf{T}}$$
$$+ \sum_{\substack{l=2:\\ l\neq i}}^{n} \left(\mathbf{e}_l + \left(\frac{n-1}{(V^1)_{1l}} (V^1)_{li}^2 + 1 \right) \mathbf{e}_i \right) \left(\mathbf{e}_l + \left(\frac{n-1}{(V^1)_{1l}} (V^1)_{li}^2 + 1 \right) \mathbf{e}_i \right)^{\mathsf{T}}. \tag{4.2}$$

Then for any $D \in \mathcal{D}$ such that $\sum_{i=1}^n \frac{1}{D_{ii}} \langle A, V^i \rangle < 0$ we have $DAD \in \mathcal{COP}^n \setminus \mathcal{K}_n^1$.

Proof. For any $A \in \mathcal{COP}^n$ and $\bar{D} \in \mathcal{D}$ we have that $\bar{D}A\bar{D} \in \mathcal{COP}^n$. Therefore we need only show that for our particular choice of D we have $DAD \notin \mathcal{K}_n^1$. A certificate for $A \notin \mathcal{K}_n^1$ would be a matrix $Y \in \mathcal{K}_n^{1*}$ such that $\langle A, Y \rangle < 0$ and we recall from [Don13] that

$$\mathcal{K}_n^{1*} = \left\{ \sum_{i=1}^n Y^i \;\mid Y^i \in \mathcal{S}_+^n \cap \mathcal{N}^n,\ (Y^i)_{jk} = (Y^j)_{ki} \text{ for all } i,j,k = 1, \ldots, n \right\}.$$

As $V^1 > 0$ the matrices $V^2, \ldots, V^n$ are well defined and we can immediately see that $V^i \in \mathcal{S}_+^n \cap \mathcal{N}^n$ for all i.

Next we show that $(V^i)_{jk} = (V^j)_{ki}$ for all i, j, k in order to show that in fact we have $\sum_{i=1}^n V^i \in \mathcal{K}_n^{1*}$. These equalities can equivalently be written as

$$\begin{aligned} &(V^k)_{ij} = (V^i)_{jk} = (V^j)_{ki} && \text{for all } i,j,k = 1,\ldots,n \text{ s.t. } i<j<k,\\ &(V^i)_{jj} = (V^j)_{ji} && \text{for all } i,j = 1,\ldots,n \text{ s.t. } i \neq j. \end{aligned}$$

We can now split up the conditions in to the following five cases:

1. For $i, j, k = 2, \ldots, n$ with $i < j < k$ we have that $0 = (V^k)_{ij} = (V^i)_{jk} = (V^j)_{ki}$.

2. For $i = 1$ and $j, k = 2, \ldots, n$ with $i < j < k$ we have that $(V^k)_{1j} = (V^1)_{jk} = (V^j)_{k1}$.

3. For $i, j = 2, \ldots, n$ such that $i \neq j$ we have that

$$(V^i)_{jj} = \frac{n-1}{(V^1)_{1i}} (V^1)_{ij}^2 + 1 = (V^j)_{ji}$$

4. For $j = 1$ and $i = 2, \dots, n$ we have that

$$(V^i)_{11} = \sum_{l=2}^{n} \frac{(V^1)_{1i}}{n-1} = (V^1)_{1i}.$$

5. For $i = 1$ and $j = 2, \dots, n$ we have that $(V^j)_{j1} = (V^1)_{jj}$.

We are now ready to prove the proposed scaling out result. Define the matrices

$$Y^i = \frac{1}{D_{ii}} D^{-1} V^i D^{-1} \text{ for all } i = 1, \dots, n.$$

It follows from the discussion above that for all $i, j, k = 1, \dots, n$ we have that $Y^i \in \mathcal{S}_+^n \cap \mathcal{N}^n$ and

$$(Y^i)_{jk} = \frac{1}{D_{ii}} \frac{1}{D_{jj}} (V^i)_{jk} \frac{1}{D_{kk}} = \frac{1}{D_{jj}} \frac{1}{D_{kk}} (V^j)_{ki} \frac{1}{D_{ii}} = (Y^j)_{ki}.$$

Therefore we have that $\sum_{i=1}^{n} Y^i \in \mathcal{K}_n^{1*}$ and

$$\left\langle DAD, \sum_{i=1}^{n} Y^i \right\rangle = \sum_{i=1}^{n} \frac{1}{D_{ii}} \langle A, V^i \rangle < 0,$$

which shows that $DAD \notin \mathcal{K}_n^1$ as claimed. □

We will now illustrate this result with an example that scales the Horn matrix (1.10) out of $\mathcal{K}_5^1$ (recall that we know the Horn matrix to be in $\mathcal{K}_5^1$).

Example 4.5. Consider the Horn matrix from (1.10). Next, let

$$V^1 = \begin{pmatrix} 18 & 12 & 2\frac{1}{2} & 2\frac{1}{2} & 12 \\ 12 & 18 & 12 & 2\frac{1}{2} & 2\frac{1}{2} \\ 2\frac{1}{2} & 12 & 18 & 12 & 2\frac{1}{2} \\ 2\frac{1}{2} & 2\frac{1}{2} & 12 & 18 & 12 \\ 12 & 2\frac{1}{2} & 2\frac{1}{2} & 12 & 18 \end{pmatrix} \in \operatorname{int}(\mathcal{S}_+^5 \cap \mathcal{N}^5).$$

It is easily verified that $\langle H, V^1 \rangle = -5 < 0$. Then via 4.2 we obtain the following inequality:

$$\sum_{i=1}^{5} \frac{1}{D_{ii}} \langle H, V^i \rangle \approx \frac{-5}{D_{11}} + \frac{53405}{D_{22}} + \frac{56428}{D_{33}} + \frac{56428}{D_{44}} - \frac{53405}{D_{55}} < 0. \tag{4.3}$$

Note that we have an infinite number of scalings that scale the Horn matrix out of $\mathcal{K}_5^1$. One such scaling can be obtained by setting $D_{22} = D_{33} = D_{44} = D_{55} = 1$ and $D_{11} = \frac{1}{50000}$, which clearly satisfies (4.3). We then obtain a matrix DHD which is in $\mathcal{COP}^5$ by Proposition 1.3, but for which it can easily be verified that it is not in $\mathcal{K}_n^1$ using the LMIs (4.1a) - (4.1d).

4.2 Non-decreasing scalings

An immediate question following the result of Theorem 4.2 is whether or not this result also implies the reverse. That is, given $r > s \geq 1$ and a matrix $A \in \mathcal{K}_n^r \setminus \mathcal{K}_n^s$ does there always exist a scaling matrix $D \in \mathcal{D}$ such that $DAD \in \mathcal{K}_n^s$? At the moment of writing this is still an open question, however we can state the result that if such scalings were to exist then the problem of finding them would have to be NP-hard. To prove this result we will initially direct out focus on the cones $\mathcal{C}_n^r$ (1.12) instead. Showing the claimed complexity result for the cones $\mathcal{C}_n^r$ then immediately implies the stated result for the cones $\mathcal{K}_n^r$. We begin by defining the following hypothetical algorithm.

Definition 4.6. Let $r > s \geq 1$. We denote by $\Gamma(\bullet)$ an algorithm that given a matrix $A \in \mathcal{C}_n^r \setminus \mathcal{C}_n^s$ returns a scaling matrix $D \in \mathcal{D}$ such that $DAD \in \mathcal{C}_n^s$.

The proof that we present in this section for the claimed complexity result will go by the means of establishing a polynomial time Turing reduction from the stable set problem to the problem of constructing scaling matrices as produced by $\Gamma(\bullet)$. In particular we will show that if there exists a polynomial time algorithm $\Gamma(\bullet)$ then we can solve the stable set problem in polynomial time.

Note that if such an algorithm $\Gamma(\bullet)$ were to exist, regardless of its complexity, it would imply that the interior of $\mathcal{COP}^n$ reduces to $\mathcal{K}_n^1$ under some scalings $\mathcal{D}' \subseteq \mathcal{D}$, as $D_1 D_2$ is again a diagonal matrix for $D_1, D_2 \in \mathcal{D}$.

Next, recall from [dKP02] that the stability number can be formulated as a copositive program via (1.30). Furthermore recall from [dKP02] the following result regarding relaxations of (1.30) using the hierarchy of cones $\mathcal{C}_n^r$.

Theorem 4.7 (Theorem 4.1, [dKP02]). *Let G be a graph with stability number α_G, and let $\zeta^{(r)} := \min\{\lambda \mid (I + A_G)\lambda - E \in \mathcal{C}_n^r\}$, $r \in \mathbb{N}$. Then*

$$\zeta^{(0)} \geq \zeta^{(1)} \geq \ldots \geq \lfloor \zeta^{(r)} \rfloor = \alpha_G$$

for $r \geq \alpha_G^2$.

This result implies that in order to obtain the stability number of a graph it is sufficient to optimize over the polyhedral cone $\mathcal{C}_n^{\alpha_G^2}$. Note that we do need to round when using such relaxations over $\mathcal{C}_n^r$, which means we cannot conclude from Theorem 4.7 that $(I + A_G)\alpha_G - E \in \mathcal{C}_n^{\alpha_G^2}$. However, from [dKP02] we also have the following result.

Lemma 4.8. *Let G be a graph on n vertices with stability number α_G and adjacency matrix A_G. For $\lambda \in \mathbb{R}$ and $\varepsilon = \frac{1}{\lambda+1/(\lambda-1)} > 0$ let*

$$Q_\lambda = (1+\varepsilon)\lambda(I + A_G) - E. \tag{4.4}$$

Then $Q_{\alpha_G} \in \mathcal{C}_n^{\alpha_G^2}$, while $(1+\varepsilon)\alpha_G < 1 + \alpha_G$.

We now introduce an algorithm that, using $\Gamma(\bullet)$ as a subroutine, decides whether or not the stability number for some graph G is larger or equal to a given $\lambda \in \mathbb{N}$.

Algorithm 1 $F_1(\lambda, A_G)$

Require: $\lambda \in \mathbb{N}$ and A_G

$i \leftarrow 0$
$\varepsilon \leftarrow \frac{1}{\lambda+1/(\lambda-1)}$
$A^{(i)} \leftarrow (1+\varepsilon)(I + A_G)\lambda - E$
while $A^{(i)} \notin \mathcal{C}_n^1$ **and** $i \le n^2$ **do**
  $D^{(i)} \leftarrow \Gamma(A^{(i)})$
  $A^{(i+1)} \leftarrow D^{(i)} A^{(i)} D^{(i)}$
  $i \leftarrow i + 1$
end while
if $i = n^2$ **then**
  return 0
else
  return 1
end if

We present the following lemma showing that Algorithm 1 does exactly what we claim it to do.

Lemma 4.9. *Consider a graph G on n vertices. Then for any $\lambda \in \mathbb{N}$, Algorithm 1 decides whether or not $\alpha_G \le \lambda$, that is $F_1(\lambda, A_G) = 1$ if and only if $\alpha_G \le \lambda$.*

Proof. First, let $\lambda = \alpha_G$. Then from Lemma 4.8 we know that $A^{(0)} \in \mathcal{C}_n^{\alpha_G^2}$. Furthermore, by definition of the algorithm $\Gamma(\bullet)$ we know that there exists a $0 \le k \le \alpha_G^2 - 1 < n^2$ such that $A^{(k)} \in \mathcal{C}_n^1$ and hence $F_1(\lambda, A_G) = 1$.
Next, let $\lambda > \alpha_G$, so that $\lambda = \alpha_G + \beta$ for some $\beta > 0$. Then

$$\begin{aligned} Q_\lambda &= (1+\varepsilon)(\alpha_G + \beta)(I + A_G) - E \\ &= (1+\varepsilon)\alpha_G(I + A_G) - E + (1+\varepsilon)\beta(I + A_G) \\ &= Q_{\alpha_G} + (1+\varepsilon)\beta(I + A_G). \end{aligned}$$

Then because $I+A_G \in \mathcal{C}_n^0 \subseteq \mathcal{C}_n^{\alpha_G^2}$, we get $Q_\lambda \in \mathcal{C}_n^{\alpha_G^2}$. The fact that $F_1(\lambda, A_G) = 1$ now follows from the case $\lambda = \alpha_G$.
Finally, let $\lambda < \alpha_G$. Due to the fact that ε is increasing with λ and because $(1+\varepsilon)\alpha_G < 1+\alpha_G$ (see Lemma 4.8), the formulation of the stability number (1.30) implies that $Q_\lambda \notin \mathcal{COP}^n$. Then, because $\mathcal{COP}^n$ is closed under scaling it can never be the case that $DQ_\lambda D \in \mathcal{COP}^n \supseteq \mathcal{C}_n^1$. As a result the **WHILE** loop will conclude only when the condition $i = n^2$ is met and hence the algorithm returns 0, or $F(\lambda, A_G) = 0$. This proves the lemma. $\square$

Next, we demonstrate that we can use Algorithm 1 to determine the stability number of a graph, establishing a Turing reduction as claimed at the start of this section. In particular we define the following algorithm that uses Algorithm 1 as a subroutine:

Algorithm 2 $F_2(A_G)$

Require: A_G
$i \leftarrow 0$
$T \leftarrow 1$
while $T = 1$ **do**
  $T \leftarrow F_1(n-i, A_G)$
  $i \leftarrow i+1$
end while
return $n - i + 1$

Lemma 4.10. *Consider a graph G on n vertices with adjacency matrix A_G. Then $F_2(A_G) = \alpha_G$.*

Proof. This follows directly from Lemma 4.9 by noting that the first ever instance where $F_1(n-i, A_G)$ returns 0 is when $n - i = \alpha_G - 1$. $\square$

Using the above presented lemmas we can now prove the complexity result stated at the beginning of this section. That is, we show that the Turing reduction established by Lemma 4.10 is a polynomial time Turing reduction from the stable set problem to the problem of finding scalings as produced by the hypothetical algorithm $\Gamma(\bullet)$.

Theorem 4.11. *For $r > s \geq 1$, assume that for any matrix $A \in \mathcal{C}_n^r \setminus \mathcal{C}_n^s$ there exists at least one matrix $D \in \mathcal{D}$ such that $DAD \in \mathcal{C}_n^s$. Then the problem of finding such scaling matrices $D \in \mathcal{D}$ for arbitrary matrices $A \in \mathcal{C}_n^r \setminus \mathcal{C}_n^s$ is NP-hard.*

Proof. Let G be a graph on n vertices and assume that there exists an algorithm $\Gamma(\bullet)$ as in Definition 4.6. Furthermore assume that this is a polynomial time algorithm. We will now show that the existence of such a polynomial time algorithm together with Algorithm 2 establishes a method to compute the stability number in polynomial time, which would prove our result.

First, it should be noted that by definition the encoding lengths of the entries on the diagonal of any scaling matrix returned by $\Gamma(\bullet)$ are polynomially bounded. If this were not the case it would contradict the assumption that $\Gamma(\bullet)$ is a polynomial time algorithm. Next, the entries of the matrix

$$Q_\lambda = (1+\varepsilon)(I + A_G)\lambda - E$$

with $\varepsilon = \frac{1}{\lambda+1/(\lambda-1)}$ for $1 \leq \lambda \leq n$ can only have two possible values, $(1+\varepsilon)\lambda$ and -1. Due to the fact that $\lambda \in \mathbb{N}$ and because λ is bounded from above by n it can easily be seen that the encoding lengths for both of these values must be polynomially bounded in n. Finally, this means the encoding lengths of the entries of $D_1 \ldots D_k Q_\lambda D_k \ldots D_1$ are bounded for all matrices $D_1, \ldots, D_k$ returned by the algorithm $\Gamma(\bullet)$. We can now apply Algorithm 2 to determine the clique number of G while being assured that the encoding lengths of all numbers involved are polynomially bounded in n.

Next, observe that the **WHILE** loop in Algorithm 2 is called at most $n - (\alpha_G - 1) \leq n$ times. During every iteration the algorithm calls the subroutine defined by Algorithm 1. The **WHILE** loop of that algorithm has at most n^2 iterations, during each of which the algorithm $\Gamma(\bullet)$ is called, which by the assumption has a time complexity of order $\mathcal{O}(q)$ for some $q \in \mathbb{R}[n]$. Furthermore we need to compute $DA^{(i)}D$ for each pass through of the **WHILE** loop, which requires at most $2n^2$ multiplications due to the fact that D is diagonal. Finally at the end of every iteration we need to check whether or not $A^{(i+1)} \in \mathcal{C}_n^1$ which can be done in polynomial time with respect to n, i.e. has time complexity $\mathcal{O}(p)$ for some $p \in \mathbb{R}[n]$. In total this means that Algorithm 2 has a worst case time complexity of order $\mathcal{O}\big(n(2n^2 + q + p)\big)$ where q and p are both polynomially bounded in n. Together with Theorem 4.10 and the observation that the encoding lengths of all numbers present in Algorithm 2 at any time are polynomially bounded in n, this concludes the proof. □

Theorem 4.11 shows that there is no general computationally easy method that is guaranteed to scale matrices into lower levels of the hierarchy of the polyhedral cones $\mathcal{C}_n^r$ unless P= NP. Moreover, it almost immediately implies the same result for the hierarchy of cones $\mathcal{Q}_n^r$ and in particular $\mathcal{K}_n^r$, as claimed at the start of this section. We summarize this result in the following Corollary.

Corollary 4.12. *Let* $\mathbb{Y}_n^r \in \{\mathcal{C}_n^r, \mathcal{Q}_n^r, \mathcal{K}_n^r\}$ *and* $r > s \geq 1$. *Furthermore let* $A \in \mathbb{Y}_n^r \setminus \mathbb{Y}_n^s$. *Now assume that there exists at least one matrix* $D \in \mathcal{D}$ *such that* $DAD \in \mathbb{Y}_n^s$. *Then the problem of finding such a scaling matrix* D *is NP-hard.*

Proof. follows directly from Theorem 4.11 and the fact that $\mathcal{C}_n^r \subseteq \mathcal{Q}_n^r \subseteq \mathcal{K}_n^r$ for $n \in \mathbb{N}$ and $r \in \mathbb{Z}_+$. □

Note that this result does not imply that scaling cannot be used to move down in any given hierarchy, but rather that efficient algorithms to construct such scalings for arbitrary matrices $A \in \mathcal{C}_n^r$ do not exist. However we might still be able to construct matrices $D \in \mathcal{D}$ which at least do not scale a matrix out of $\mathcal{C}_n^r$ for a given $r \in \mathbb{N}$. Moreover if we choose such a scaling matrix in a clever way there is the possibility that it scales a given matrix downwards into any of the hierarchies. We will call such scalings *non-decreasing scalings.* A trivial example of a non-decreasing scaling matrix is the identity matrix.

Scaling a matrix is computationally cheap, making them interesting object for preprocessing purposes when optimizing over any of the hierarchies mentioned in this section. A relatively simple example of a type of non-decreasing scaling that, unlike the identity matrix, might be useful to apply concerns matrices with a nonnegative row and column.

Theorem 4.13. *Let* $A \in \mathcal{COP}^n$ *and assume that for some* $i \in \{1, \dots, n\}$ *we have* $(A)_{ij} \geq 0$ *for every* $j = 1, \dots, n$. *Then for* $D = I + \lambda \mathbf{e}_i \mathbf{e}_i^\mathsf{T}$ *with* $\lambda \geq 0$ *and using the notation from Definition 1.15 we get that*

$$r^*_{\mathbb{Y}_n^r}(DAD) \leq r^*_{\mathbb{Y}_n^r}(A)$$

for $\mathbb{Y}_n^r \in \{\mathcal{C}_n^r, \mathcal{Q}_n^r, \mathcal{K}_n^r\}$.

Proof. Note that by the properties of A and D we have $DAD \geq A$ entry-wise. This immediately implies that the coefficients of the polynomial

$$\Big(\sum_{i=1}^n x_i\Big)^r \mathbf{x}^\mathsf{T} DAD\mathbf{x}$$

are greater than or equal to the coefficients of

$$\Big(\sum_{i=1}^n x_i\Big)^r \mathbf{x}^\mathsf{T} A\mathbf{x},$$

and hence $r^*_{\mathcal{C}_n^r}(DAD) \leq r^*_{\mathcal{C}_n^r}(A)$. For the hierarchies $\mathcal{Q}_n^r$ and $\mathcal{K}_n^r$, the result follows from the fact that $\mathcal{C}_n^r \subseteq \mathcal{Q}_n^r \subseteq \mathcal{K}_n^r$. □

The idea behind this type of scaling is that the coefficients of the polynomial $f(\mathbf{x}) = (\sum_{i=1}^{n})^r \mathbf{x}^\mathsf{T} A\mathbf{x}$ are formed by linear combinations of the entries of $A \in \mathbb{S}^n$. Therefore, making several entries larger while leaving all the others unchanged could in theory decrease the lifting rank $r^*_{\mathcal{C}_n^r}(A)$. We provide the following example to illustrate this idea.

Example 4.14. Consider the following matrix

$$A = \begin{pmatrix} 2 & 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 & 1 \\ 1 & 1 & 1 & 5 & -\frac{5}{2} \\ 1 & 1 & 1 & -\frac{5}{2} & 7 \end{pmatrix} \in \mathcal{C}_5^2 \setminus \mathcal{C}_5^1.$$

Then taking $D = I + 9\mathbf{e}_1\mathbf{e}_1^\mathsf{T} + 9\mathbf{e}_2\mathbf{e}_2^\mathsf{T} + 9\mathbf{e}_3\mathbf{e}_3^\mathsf{T}$ it can easily be checked that $DAD \in \mathcal{C}_5^1$.

The case where a row and column contain both positive and negative values appears to be more complicated. Yet, in that case scalings can still be used to go down in the aforementioned hierarchies. A simple example to show this would again be the matrix A given in Example 4.14 by simply multiplying the scaling defined in that example by 1/10. That way we are technically scaling rows 4 and 5 of A which contain both positive and negative values. An example for a matrix that does not have any nonnegative rows or columns at all can be obtained by simply reversing the scaling that was found for the Horn matrix in Example 4.5.

A related interesting question is whether non-decreasing scalings could be constructed and applied in a generic way to some possibly well structured matrix variables of a copositive program. For example the entries of the matrix $(I + A_G)\lambda - E$ that appears in the copositive formulation of the stability problem (1.30) can be equal to only two different values, $\lambda - 1$ and -1. Moreover we know exactly where the negative entries occur for a given graph G, as well as their value which is constant. If it would be possible to construct non-decreasing scalings for this particular matrix it would open up the possibility of improving the bounds obtained in [dKP02] via $\mathcal{C}_n^r$ and $\mathcal{K}_n^r$, without actually having to go up in the respective hierarchies.

4.3 Scaling a matrix into $\mathcal{K}_5^1$

In this section, we will show that in the 5×5 case it is possible to scale any copositive matrix into $\mathcal{K}_5^1$. More precisely, we show that for any $X \in \mathcal{COP}^5$ there exists a scaling D such that $DXD \in \mathcal{K}_5^1$. To this end, we make use of the theory developed in Chapter 3, where we investigated matrices of the form (3.5), i.e.

$$S(\boldsymbol{\theta}) := \begin{pmatrix} 1 & -\cos\theta_1 & \cos(\theta_1+\theta_2) & \cos(\theta_4+\theta_5) & -\cos\theta_5 \\ -\cos\theta_1 & 1 & -\cos\theta_2 & \cos(\theta_2+\theta_3) & \cos(\theta_5+\theta_1) \\ \cos(\theta_1+\theta_2) & -\cos\theta_2 & 1 & -\cos\theta_3 & \cos(\theta_3+\theta_4) \\ \cos(\theta_4+\theta_5) & \cos(\theta_2+\theta_3) & -\cos\theta_3 & 1 & -\cos\theta_4 \\ -\cos\theta_5 & \cos(\theta_5+\theta_1) & \cos(\theta_3+\theta_4) & -\cos\theta_4 & 1 \end{pmatrix}$$

where $\boldsymbol{\theta} \in \Theta := \left\{\boldsymbol{\theta} \in \mathbb{R}^5_+ \;\middle|\; \mathbf{e}^\mathsf{T}\boldsymbol{\theta} < \pi\right\}$. We will show below in Theorem 4.15 that $S(\boldsymbol{\theta}) \in \mathcal{K}_5^1$. Combining this result with Corollary 3.35 will then immediately imply the proposed scaling result: Take $A \in \mathcal{COP}^5 \setminus (\mathcal{S}^5_+ + \mathcal{N}^5)$. If A has a zero diagonal entry, say $A_{11} = 0$, then copositivity of A implies that its first row and column must be nonnegative, so we can decompose A as

$$A = \begin{pmatrix} 0 & \mathbf{b}^\mathsf{T} \\ \mathbf{b} & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & C \end{pmatrix} \tag{4.5}$$

where $\mathbf{b} \in \mathbb{R}^4$ is nonnegative and $C \in \mathcal{COP}^4 = \mathcal{S}^4_+ + \mathcal{N}^4$, from which we immediately get $A \in \mathcal{K}_5^0 \subseteq \mathcal{K}_5^1$. Now assume $\operatorname{diag}(A) > 0$. Then we use Corollary 3.35 to get $A = DP^\mathsf{T}S(\boldsymbol{\theta})PD + N$. But this is equivalent to

$$D^{-1}AD^{-1} = P^\mathsf{T}S(\boldsymbol{\theta})P + D^{-1}ND^{-1}$$

which is then a decomposition of $D^{-1}AD^{-1}$ as a sum of two elements in $\mathcal{K}_5^1$ (observe that $\mathcal{K}_5^1$ is closed under permutations), so we conclude $D^{-1}AD^{-1} \in \mathcal{K}_5^1$.

Observe that the diagonal entries of $D^{-1}AD^{-1}$ are all equal 1 because of $(S(\boldsymbol{\theta}))_{ii} = 1$ and $(N)_{ii} = 0$. In case there was a zero entry $A_{ii} = 0$, we can analogously scale the submatrix C in (4.5) to 0/1 diagonal. This shows that for any matrix $A \in \mathcal{COP}^5$, scaling A to 0/1 diagonal will yield a matrix in $\mathcal{K}_5^1$.

So for the above arguments to be valid, it remains to show the following theorem.

Theorem 4.15. *For all $\boldsymbol{\theta} \in \Theta$, we have that $S(\boldsymbol{\theta}) \in \mathcal{K}_5^1$.*

Proof. Recall that $S(\boldsymbol{\theta}) \in \mathcal{K}_5^1$ if and only if the following system of LMIs has a feasible solution $M^1(\boldsymbol{\theta}), \dots, M^5(\boldsymbol{\theta})$:

$$S(\boldsymbol{\theta}) - M^i(\boldsymbol{\theta}) \in \mathcal{S}^5_+ \qquad i = 1, \dots, 5 \tag{4.6a}$$

$$(M^i(\boldsymbol{\theta}))_{ii} = 0 \qquad i = 1, \dots, 5 \tag{4.6b}$$

$$(M^i(\boldsymbol{\theta}))_{jj} + 2(M^j(\boldsymbol{\theta}))_{ij} = 0 \qquad i, j = 1, \dots, 5 \quad \text{s.t. } i \neq j \tag{4.6c}$$

$$(M^i(\boldsymbol{\theta}))_{jk} + (M^j(\boldsymbol{\theta}))_{ik} + (M^k(\boldsymbol{\theta}))_{ij} \geq 0 \quad i, j, k = 1, \dots, 5 \quad \text{s.t. } i < j < k. \tag{4.6d}$$

We claim that the following matrices constitute a feasible solution for this system:

$$M^1(\boldsymbol{\theta}) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha_1(\boldsymbol{\theta}) & 0 \\ 0 & 0 & 0 & \beta_1(\boldsymbol{\theta}) & \gamma_1(\boldsymbol{\theta}) \\ 0 & \alpha_1(\boldsymbol{\theta}) & \beta_1(\boldsymbol{\theta}) & 0 & 0 \\ 0 & 0 & \gamma_1(\boldsymbol{\theta}) & 0 & 0 \end{pmatrix},$$

$$M^2(\boldsymbol{\theta}) = \begin{pmatrix} 0 & 0 & 0 & \gamma_2(\boldsymbol{\theta}) & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \alpha_2(\boldsymbol{\theta}) \\ \gamma_2(\boldsymbol{\theta}) & 0 & 0 & 0 & \beta_2(\boldsymbol{\theta}) \\ 0 & 0 & \alpha_2(\boldsymbol{\theta}) & \beta_2(\boldsymbol{\theta}) & 0 \end{pmatrix},$$

$$M^3(\boldsymbol{\theta}) = \begin{pmatrix} 0 & 0 & 0 & \alpha_3(\boldsymbol{\theta}) & \beta_3(\boldsymbol{\theta}) \\ 0 & 0 & 0 & 0 & \gamma_3(\boldsymbol{\theta}) \\ 0 & 0 & 0 & 0 & 0 \\ \alpha_3(\boldsymbol{\theta}) & 0 & 0 & 0 & 0 \\ \beta_3(\boldsymbol{\theta}) & \gamma_3(\boldsymbol{\theta}) & 0 & 0 & 0 \end{pmatrix},$$

$$M^4(\boldsymbol{\theta}) = \begin{pmatrix} 0 & \beta_4(\boldsymbol{\theta}) & \gamma_4(\boldsymbol{\theta}) & 0 & 0 \\ \beta_4(\boldsymbol{\theta}) & 0 & 0 & 0 & \alpha_4(\boldsymbol{\theta}) \\ \gamma_4(\boldsymbol{\theta}) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & \alpha_4(\boldsymbol{\theta}) & 0 & 0 & 0 \end{pmatrix},$$

$$M^5(\boldsymbol{\theta}) = \begin{pmatrix} 0 & 0 & \alpha_5(\boldsymbol{\theta}) & 0 & 0 \\ 0 & 0 & \beta_5(\boldsymbol{\theta}) & \gamma_5(\boldsymbol{\theta}) & 0 \\ \alpha_5(\boldsymbol{\theta}) & \beta_5(\boldsymbol{\theta}) & 0 & 0 & 0 \\ 0 & \gamma_5(\boldsymbol{\theta}) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

where for all $i = 1, \ldots, 5$ (the indices being modulo 5) we define

$$\begin{aligned} \alpha_i(\boldsymbol{\theta}) &= \cos(\theta_{i-2} + \theta_{i-1} + \theta_i) + \cos(\theta_{i+1} + \theta_{i+2}), \\ \beta_i(\boldsymbol{\theta}) &= -\cos(\theta_{i+2}) - \cos(\theta_{i-2} + \theta_{i-1} + \theta_i + \theta_{i+1}), \\ \gamma_i(\boldsymbol{\theta}) &= \cos(\theta_{i-1} + \theta_i + \theta_{i+1}) + \cos(\theta_{i+2} + \theta_{i-2}). \end{aligned}$$

From the fact that $(M^i(\boldsymbol{\theta}))_{jj} = (M^i(\boldsymbol{\theta}))_{ij} = 0$ for all $i, j = 1, \ldots, 5$ it is immediately clear that (4.6b) and (4.6c) hold for $M^1(\boldsymbol{\theta}), \ldots, M^5(\boldsymbol{\theta})$.

To show (4.6a) we first note that due to the cyclic symmetry in $S(\boldsymbol{\theta})$ and the $M^i(\boldsymbol{\theta})$'s we need only prove this for a single i. Taking $i = 1$ we have the following factorization, where we use the well known formula that $\cos(a+b) = \cos a \cos b - \sin a \sin b$:

$$
\begin{aligned}
&S(\boldsymbol{\theta}) - M^1(\boldsymbol{\theta}) \\
&= \begin{pmatrix}
1 & -\cos\theta_1 & \cos(\theta_1+\theta_2) & \cos(\theta_4+\theta_5) & -\cos\theta_5 \\
-\cos\theta_1 & 1 & -\cos\theta_2 & -\cos(\theta_4+\theta_5+\theta_1) & \cos(\theta_5+\theta_1) \\
\cos(\theta_1+\theta_2) & -\cos\theta_2 & 1 & \cos(\theta_4+\theta_5+\theta_1+\theta_2) & -\cos(\theta_5+\theta_1+\theta_2) \\
\cos(\theta_4+\theta_5) & -\cos(\theta_4+\theta_5+\theta_1) & \cos(\theta_4+\theta_5+\theta_1+\theta_2) & 1 & -\cos\theta_4 \\
-\cos\theta_5 & \cos(\theta_5+\theta_1) & -\cos(\theta_5+\theta_1+\theta_2) & -\cos\theta_4 & 1
\end{pmatrix} \\
&= \begin{pmatrix} 1 \\ -\cos\theta_1 \\ \cos(\theta_1+\theta_2) \\ \cos(\theta_4+\theta_5) \\ -\cos\theta_5 \end{pmatrix}
\begin{pmatrix} 1 \\ -\cos\theta_1 \\ \cos(\theta_1+\theta_2) \\ \cos(\theta_4+\theta_5) \\ -\cos\theta_5 \end{pmatrix}^{\mathsf{T}}
+ \begin{pmatrix} 0 \\ \sin\theta_1 \\ -\sin(\theta_1+\theta_2) \\ \sin(\theta_4+\theta_5) \\ -\sin\theta_5 \end{pmatrix}
\begin{pmatrix} 0 \\ \sin\theta_1 \\ -\sin(\theta_1+\theta_2) \\ \sin(\theta_4+\theta_5) \\ -\sin\theta_5 \end{pmatrix}^{\mathsf{T}}.
\end{aligned}
$$

Hence $S(\boldsymbol{\theta}) - M^1(\boldsymbol{\theta}) \in \mathcal{S}_+^5$ and (4.6a) is shown.

Showing (4.6d) requires a bit more work. We shall do this by first noting that for most indices this condition is equivalent to $0 \geq 0$, and for the remaining indices it is equivalent to (the indices being modulo 5)

$$\alpha_j(\boldsymbol{\theta}) + \beta_{j-2}(\boldsymbol{\theta}) + \gamma_{j+1}(\boldsymbol{\theta}) \geq 0 \qquad j = 1, \ldots, 5, \quad \boldsymbol{\theta} \in \operatorname{cl}\Theta.$$

We recall that for $a, b, c \in \mathbb{R}$ we have that

$$
\begin{aligned}
\cos a + \cos b &= 2\cos\left(\tfrac{1}{2}(a+b)\right)\cos\left(\tfrac{1}{2}(b-a)\right), \\
\cos a - \cos b &= 2\sin\left(\tfrac{1}{2}(a+b)\right)\sin\left(\tfrac{1}{2}(b-a)\right).
\end{aligned}
$$

From this we also get that

$$
\begin{aligned}
&\cos(a+b-c) + \cos(a-b+c) - \cos(-a+b+c) \\
&\quad = \cos(a+b+c) + \Big(\cos(a+b-c) + \cos(a-b+c)\Big) \\
&\qquad\quad - \Big(\cos(-a+b+c) + \cos(a+b+c)\Big) \\
&\quad = \cos(a+b+c) + 2\cos a\cos(b-c) - 2\cos a\cos(b+c) \\
&\quad = \cos(a+b+c) + 4\cos a \sin b \sin c.
\end{aligned}
$$

Using these trigonometric identities, we get the following.

$$\begin{aligned}
&\alpha_j(\boldsymbol{\theta}) + \beta_{j-2}(\boldsymbol{\theta}) + \gamma_{j+1}(\boldsymbol{\theta}) \\
&\quad = \cos(\theta_{j-2} + \theta_{j-1} + \theta_j) + \cos(\theta_{j+1} + \theta_{j+2}) - \cos(\theta_j) \\
&\qquad\qquad - \cos(\theta_{j+1} + \theta_{j+2} + \theta_{j-2} + \theta_{j-1}) \\
&\qquad\qquad + \cos(\theta_j + \theta_{j+1} + \theta_{j+2}) + \cos(\theta_{j-2} + \theta_{j-1}) \\
&\quad = 2\cos\left(\tfrac{1}{2}\mathbf{e}^\mathsf{T}\boldsymbol{\theta}\right)\Big[\cos\left(\tfrac{1}{2}(\theta_j + (\theta_{j-2} + \theta_{j-1}) - (\theta_{j+1} + \theta_{j+2}))\right) \\
&\qquad\qquad - \cos\left(\tfrac{1}{2}(-\theta_j + (\theta_{j-2} + \theta_{j-1}) + (\theta_{j+1} + \theta_{j+2}))\right) \\
&\qquad\qquad + \cos\left(\tfrac{1}{2}(\theta_j - (\theta_{j-2} + \theta_{j-1}) + (\theta_{j+1} + \theta_{j+2}))\right)\Big] \\
&\quad = 2\cos\left(\tfrac{1}{2}\mathbf{e}^\mathsf{T}\boldsymbol{\theta}\right)\Big[\cos\left(\tfrac{1}{2}\mathbf{e}^\mathsf{T}\boldsymbol{\theta}\right) \\
&\qquad\qquad + 4\cos\left(\tfrac{1}{2}\theta_j\right)\sin\left(\tfrac{1}{2}(\theta_{j-2} + \theta_{j-1})\right)\sin\left(\tfrac{1}{2}(\theta_{j+1} + \theta_{j+2})\right)\Big].
\end{aligned}$$

This can now clearly be seen to be greater than or equal to zero for all $j = 1, \dots, 5$, $\boldsymbol{\theta} \in \Theta$. □

We can now summarize the discussions of this section in the following corollary and theorems:

Corollary 4.16. *Let $A \in \mathbb{S}^5$ such that $A_{ii} \in \{0, 1\}$ for all i. Then $A \in \mathcal{COP}^5$ if and only if $A \in \mathcal{K}_5^1$.*

This gives us the following interesting result.

Theorem 4.17. *Let $A \in \mathbb{S}^5$ and $D \in \mathcal{D}$ be such that $(DAD)_{ii} \in \{0, 1\}$ for all i. Then we have that $A \in \mathcal{COP}^5$ if and only if $DAD \in \mathcal{K}_5^1$.*

This means that if we want to test a matrix $A \in \mathbb{S}^5$ with nonnegative diagonal entries for copositivity, then it suffices to scale it so that its diagonal entries become binary, and to test the scaled matrix for inclusion in $\mathcal{K}_5^1$. The original matrix will be copositive if and only if the scaling is in $\mathcal{K}_5^1$.

Finally we have the following result on a characterization of $\mathcal{COP}^5$.

Theorem 4.18. *For matrices of order five, we have the following characterization of $\mathcal{COP}^5$:*

$$\mathcal{COP}^5 = \{DAD \mid A \in \mathcal{K}_5^1,\ D \in \mathcal{D}\}.$$

We will now briefly illustrate this result with the help of one of our earlier examples.

Example 4.19. Consider the matrix as defined by (**??**). We know that this matrix is in $\mathcal{COP}^5$ but not in $\mathcal{K}_5^1$. Obviously scaling this matrix in such a way that every diagonal entry is equal to 1 simply returns the Horn matrix again which we know to be in $\mathcal{K}_n^1$. In fact such a scaling is exactly the inverse of the scaling D we found in Example 4.5.

4.4 Conjectures and open problems

In this chapter we elaborated some very surprising results, and we are left with some open questions which we will discuss in this section. We will formulate a number of conjectures for which we also provide supporting evidence.

In Section 4.3 we saw that for $n \leq 5$ and $r = 1$ we have that

$$\mathcal{COP}^n = \{DAD \mid A \in \mathcal{K}_n^r,\ D \in \mathcal{D}\}. \tag{4.7}$$

It is an open question as to whether a similar result holds for higher n. It could be that (4.7) holds with $r = 1$ for all $n \geq 1$, or alternatively the weaker statement that for all $n \geq 1$ there exists an $r \geq 0$ such that (4.7) could hold.

In fact in Section 4.3 we found that for $n = 5$ and $r = 1$, it was useful to scale the matrix to have binary values on the diagonal. We extend this with the following conjecture.

Conjecture 4.20. *For all $n \geq 1$, there exists a finite $r \geq 0$ such that for any matrix $A \in \mathbb{S}^n$, with $A_{ii} \geq 0$ for all i, and any $D \in \mathcal{D}$ such that $(DAD)_{ii} \in \{0, 1\}$ for all i we have that $A \in \mathcal{COP}^n$ if and only if $DAD \in \mathcal{K}_n^r$.*

Numerical experiments with randomly generated instances carried out by the authors suggested that if we take a matrix in $\mathcal{K}_n^1$ and scale it such that the diagonal is binary then the scaled version of the matrix would also be in $\mathcal{K}_n^1$. An extension and reinterpretation of this result is the following.

Conjecture 4.21. *Let $A \in \mathbb{S}^n$ such that $A_{ii} \in \{0, 1\}$ for all i. Then we have that $A \notin \mathcal{K}_n^r$ implies that $DAD \notin \mathcal{K}_n^r$ for all $D \in \mathcal{D}$.*

This conjecture effectively says that if you have an arbitrary matrix, and you wish to use one of the Parrilo cones as an approximation for testing if this matrix is copositive, then the best thing to do is to first scale the matrix so that its diagonal entries are binary, as, if this is not in the Parrilo cone, then no scaling of it will be in the Parrilo cone.

These conjectures suggest the importance of scaling the diagonal to binary for the Parrilo cones. One piece of support for this idea for $n \geq 6$ is the following theorem.

Theorem 4.22. *Let $A \in \mathcal{COP}^n$ be such that $A_{ij} \in \{-1, +1\}$ for all $i, j = 1, \ldots, n$. Then we have that $A \in \mathcal{K}_n^1$.*

Proof. We consider an arbitrary $A \in \mathcal{COP}^n$ such that $A_{ij} \in \{-1, +1\}$ for all i, j. First note that we must have that $A_{ii} = 1$ for all i.

We now let $M^1, \ldots, M^n \in \mathbb{S}^n$ be given as follows,

$$(M^i)_{jk} = A_{jk} - A_{ij}A_{ik} \qquad \text{for all } i, j, k = 1, \ldots, n.$$

We claim that these provide a feasible solution to the system of LMIs (4.1a)–(4.1d), and thus a certificate for $A \in \mathcal{K}_n^1$.

From construction it is immediately apparent that for all i we have that $A - M^i$ is a rank 1, positive semidefinite matrix, and so (4.1a) holds.

For all $i, j = 1, \ldots, n$ we have that

$$\begin{aligned}(M^i)_{jj} &= A_{jj} - A_{ij}^2 = 1 - (\pm 1)^2 = 0,\\ (M^i)_{ij} &= A_{ij} - A_{ii}A_{ij} = A_{ij} - A_{ij} = 0.\end{aligned}$$

From this we immediately get that (4.1b) and (4.1c) hold.

We are now left to show that (4.1d) holds. Suppose for the sake of contradiction that there exists an $i < j < k$ such that

$$\begin{aligned}0 &> (M^i)_{jk} + (M^j)_{ik} + (M^k)_{ij}\\ &= A_{jk} + A_{ik} + A_{ij} - A_{ij}A_{ik} - A_{ij}A_{jk} - A_{ik}A_{jk}.\end{aligned}$$

As all the elements of A are in $\{-1, +1\}$, we must have that $-1 = A_{jk} = A_{ik} = A_{ij}$. However if we now let $\mathbf{z} \in \{0, 1\}^n \subset \mathbb{R}_+^n$ such that

$$(\mathbf{z})_l = \begin{cases} 1 & \text{if } l \in \{i, j, k\}\\ 0 & \text{otherwise,}\end{cases}$$

then we get the contradiction that

$$0 \leq \mathbf{z}^\mathsf{T} A \mathbf{z} = A_{ii} + A_{jj} + A_{kk} + 2A_{ij} + 2A_{ik} + 2A_{jk} = -3 < 0$$

and the proof is complete. □

Copositive matrices with all entries equal to ± 1 were previously studied in [HH69]. We have shown that these matrices must be in $\mathcal{K}_n^1$, even though this includes matrices not in $\mathcal{S}_+^n + \mathcal{N}^n$ (for example the Horn matrix) which from Section 4.1 we know can be scaled out of $\mathcal{K}_n^1$.

Equivalent to scaling to binary would be to scale such that all the diagonal entries are either equal to zero or equal to the same positive scalar. We shall now see that further support for this type of scaling comes from the use of the Parrilo cones in approximating the stability number of a graph. It was shown in [dKP02] that the stability number $\alpha(G)$ of a graph with n nodes and adjacency matrix A_G can be computed as

$$\alpha(G) = \min\{\lambda \in \mathbb{R} \mid \lambda(I + A_G) - E \in \mathcal{COP}^n\},$$

where E is the all ones matrix and I is the identity matrix. An approximation of this can then be provided by replacing the copositive cone with one of the Parrilo cones. This provides a very good approximation in practice.

Since A_G has a zero diagonal we have that for any λ the diagonal entries of $\lambda(I + A_G) - E$ are all equal. This means that the matrix is already scaled in the way that these conjectures would suggest is best, which could give an interpretation as to why the approximations are so good. In fact for the 5×5 case an analogous result of Corollary 4.16 applies. Consequently, our results show that for any graph G with five nodes we have that

$$\begin{aligned} \alpha(G) &= \min\{\lambda \in \mathbb{R} \mid \lambda(I + A_G) - E \in \mathcal{COP}^5\} \\ &= \min\{\lambda \in \mathbb{R} \mid \lambda(I + A_G) - E \in \mathcal{K}_5^1\}. \end{aligned} \tag{4.8}$$

This fact has been observed by De Klerk and Pasechnik [dKP02] for the case where G is the 5-cycle. All other graphs with five vertices are perfect, and it is known [PVZ07, Corollary 15] that the $\mathcal{K}_n^0$-approximation (and hence also the $\mathcal{K}_n^1$-approximation) provides an exact answer for perfect graphs. Therefore, (4.8) is implicitly known, but it seems that it has never been stated explicitly.