Semidefinite Relaxations for Copositive Optimization

Sergio Camelo Advisor: Mauricio Velasco

May 15, 2015

Abstract

Linear optimization over the copositive cone $\mathcal{C}^*$, (i.e. the cone of quadratic forms which are nonnegative on the positive orthant) has applications in polynomial optimization, graph theory and data analysis. Although convex, this problem is unfortunately not tractable. In this work we study two nested sequences of spectrahedral cones that approximate $\mathcal{C}^*$. One formulated by Barvinok, Veomett and Laserre $\{\mathcal{BVL}_n\}_{n\in\mathbb{N}}$ and the other proposed by Parrilo $\{\mathcal{SOS}_n\}_{n\in\mathbb{N}}$. Since these approximations are one from above and one from below, they can be used to calculate upper and lower bounds on the solutions of copositive programs efficiently. This proves particularly useful in bounding the independence number of a graph $\alpha(G)$. In this case, the fact that $\alpha(G)$ is an integer means the lower and upper bounds need not meet to find the optimal value of the problem. This work is divided in four parts, first, we give a comprehensive description of the geometry of the copositive cone, second, we describe its semidefinite approximations, third, we survey the applications of copositive programming, fourth, we attempt to calculate the independence number of certain DeBrujin and Hamming graphs using sdp software.

Contents

1 Introduction

The availability of large amounts of data and the development of statistical methods to learn from it make the field of optimization each day more important. While there are diverse machine learning methods, ranging from OLS linear regression to neural networks, the decision to adopt one or the other is guided by the computational power at hand. It does not make much sense, thus, to favour complex algorithms that rely on optimization problems that cannot be solved efficiently, in spite of all the statistical guarantees these algorithms may have.

The need to solve problems in data analysis, and areas like finance, biology and operations research, call for a proper study of the complexity of different optimization problems. In this sense, the theory has given a special place to convex optimization (minimizing a convex function over a convex set). There is a strong motivation in doing so, among others, the fact that such problems can be solved in polynomial time through the ellipsoid method if an efficient oracle for the membership of the feasible set exists.

A special class of convex optimization is conic programming, which comes down to minimizing a linear function, subject to linear and conic constraints

$$\min c^T x \tag{1}$$

$$s.t. Ax = b \tag{2}$$

$$x \in \mathcal{K} \tag{3}$$

The importance of conic problems lies in two aspects. First, it permits a simple framework to develop a general theory of optimization. Second, many non-convex optimization problems admit a conic relaxation that in some sense is equivalent to the original problem. And third, even when the conic problem cannot be solved in polynomial time, it allows efficient approximations that converge to an optimal solution.

In this work we focus on conic programming over the cone of copositive matrices

$$\mathcal{C}^* := \{A \in \mathcal{S}^{n\times n} : x^T A x \geq 0\} \tag{4}$$

There will be three main objectives of this thesis: First, we show that copositive programming is general enough to encode problems in polynomial optimization, graph theory and data analysis. Second, we give a comprehensive account on how the copositive cone can be approximated through two hierarchies of spectrahedra: the Sum of Squares ($\mathcal{SOS}$) approximation from above and the Barvinok-Veomett-Lasserre ($\mathcal{BVL}$) from below. The importance of these approximations lies in the impossiblility to solve copositive programs in polynomial time unless $P = NP$, whereas conic programming over spectrahedra can be done efficiently up to a given accuracy.

Third, we explicitly calculate these approximations, program them on Matlab, and evaluate their performance in calculating the independence number of Hamming and deBrujin graphs. A very interesting feature of this implementation is that the hierarchies $\mathcal{SOS}$ an $\mathcal{BVL}$ give lower and upper bounds on the independence number of a graph $\alpha(G)$, but because $\alpha(G)$ is an integer, these bounds do not have to meet to obtain the optimal solution to the independence number problem.

This work is self-contained. Many of the results are drawn from their seminal papers and appropriately referenced, but molded to maintain consistency and to show the relationship between the different results. One example where we did this is Lemmas 9 and 11.

The work is divided into ten sections. Section 1 is the introduction. Section 2 gives preliminaries in convex programming and multilinear algebra, as well as isolated results that will be important later. Section 3 studies the geometry of the copositive cone and its dual, focusing on the interior of the cones, the extremal rays and the cones in small dimensions. Section 4 describes the $\mathcal{BVL}$ spectrahedral hierarchy to approximate any proper cone, and calculates this approximation explicitly for the copositive cone. Section 5 does the same for the $\mathcal{SOS}$ hierarchy. Section 6 studies three programs in quadratic optimization and shows their equivalence with copositive programming. Section 7 studies the equivalence of copositive optimization on tensors with polynomial optimization. Section 8 studies applications in graph theory and data analysis, particularly the independence and chromatic number of a graph. Section 9 shows the computational approximations of the independence number of DeBrujin and Hamming graphs, and Section 10 gives further directions for research.

2 Preliminaries

In this section we state definitions and results from conic programming and multilinear algebra that will be relevant afterwards. This is no extensive account and many well-known results are given without a proof, so we remit the interested reader to one of the classics in convex optimization [27].

2.1 Proper Cones

Throughout this section, assume V is a finite dimensional vector space over $\mathbb{R}$ with inner product $\langle \cdot, \cdot \rangle$ and denote V^* its dual, i.e., the vector space of linear functions $l : V \to \mathbb{R}$. For any $S \subset V$ the expressions: $\text{cone}(S)$, $\text{conv}(S)$, $\text{affine}(S)$ will denote the smallest cone, convex set and affine hyperplane which contain K, respectively. We also denote $\text{conic}(S)$ as the smallest convex cone that contains S, i.e., $\text{conic}(S)$=$\text{conv}(\text{cone}(S))$.

Definition 1 (Proper Cone)**.** $K \subseteq V$ will be called a cone if for all $x \in K$ and $t \in \mathbb{R}_+$ then $tx \in K$. We call K a proper cone if it also satisfies the conditions

(i) K is convex and closed

(ii) K is pointed, i.e. $x, -x \in K$ implies $x = 0$

(iii) K has non empty interior

We will make extensive use of the dual cone

$$K^* := \{l \in V^* : l(x) \geq 0 \; \forall x \in K\} \tag{5}$$

Notice that if $V = \mathbb{R}^n$ or $V = \mathcal{S}^{n\times n}$ (the vector space of symmetric matrices), then $l \in V^*$ can be linked to a vector a_l such that $l(x) = \langle a_l, x\rangle$ or a symmetric matrix A_l such that $l(X) = \langle A_l, X\rangle$. In any case, we can identify V with V^*, which allows the equivalent description

$$K^* := \{a \in V : \langle a, x\rangle \geq 0 \; \forall x \in K\} \tag{6}$$

The next proposition gives important relationships between a cone and its dual

Proposition 1. *Let $K, K_1, K_2 \subseteq V$ be cones. Then*

(i) K^* *is always closed and convex*

(ii) Whenever $K_1 \subseteq K_2$ *then* $K_2^* \subseteq K_1^*$

(iii) If K *has nonempty interior then* K^* *is pointed*

(iv) If $cl(K)$ *is pointed then* K^* *has non-empty interior*

(v) $(K^{**}) = cl(conv(K))$

It is easy to deduce that if K *is a proper cone, then* K^* *is proper and* $K^{**} = K$

Whenever we study a cone we are in favour of an explict description of its elements. If K is the smallest convex cone that contains S, the next result assures it can be explicitly expressed as a short linear combinations of elements in S.

Theorem 1 (Caratheodory's Theorem)**.** *If* $K = conic(S)$ *is a convex cone in a vector space of dimension* n*, then for every* $x \in K$ *there exist* $x_1, \ldots, x_n \in S$ *and* $\lambda_1, dots, \lambda_n \in \mathbb{R}_+$ *such that*

$$x = \lambda_1 x_1 + \cdots + \lambda_n x_n \tag{7}$$

Very relevant to our discussion will also be the extreme vectors of a cone

Definition 2 (Extreme Vectors and Rays)**.** We say that $x \in K$ is a extreme vector of a pointed cone K if whenever we write $x = x_1 + x_2$ for $x_1, x_2 \in K$, necessarily $x_1, x_2 \in \text{span}(x)$. We also say that R is an extreme ray generated by x if x is an extreme vector and $R = \{\lambda x : \lambda \geq 0\}$. Notice that any vector of an extreme ray will be an extreme vector.

Intuitively, the extreme rays of K allow the simplest description of a cone K. Their definition shows they cannot be resonstructed through linear combinations after removing them from the cone. The next theorem illustrates how we can build a proper cone through a description of its extreme rays.

Theorem 2 (Krein-Milman Theorem I)**.** *If* K *is a proper cone and* $T \subseteq K$ *is a set of extreme vectors of* K *that can generate all extreme rays, then* $K = cl(cone(T))$*.*

Let $\text{Ext}(K)$ be the set of extreme vectors of K. Then the last result implies that we only need the extreme rays to define the dual cone K^*

$$K^* = \{x \in V : \langle x, y \rangle \geq 0 \ \forall y \in \text{Ext}(K)\} \tag{8}$$

We will also make use of the extreme points of a set

Definition 3 (Extreme Points of a Convex Set)**.** Let $S \subset \mathbb{R}^n$ be a convex set. Then $x \in S$ is an extreme point of S if the convex combination

$$x = \lambda_1 x_1 + \cdots + \lambda_n x_n \text{ for } x_i \in S \tag{9}$$

implies that $x = x_1 = \cdots = x_n$

The following theorem is the equivalent of Theorem 3 for extreme points

Theorem 3 (Krein-Milman Theorem II)**.** *If* S *is a convex compact set then* $Convex(Ext(S)) = S$

2.2 Conic Programming

In this work we will concern ourselves mainly with conic programming, particularly in its application to polynomial optimization. Let $K \subseteq \mathbb{R}^n$ be a proper cone and A a linear function $A : \mathbb{R}^n \to \mathbb{R}^m$. We define a conic program as the problem

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \ & c^T x \\ s.t. \ & Ax = b \\ & x \in K \end{aligned} \tag{PRIMAL}$$

where $c \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$. We will refer to this problem as the Primal. Denote $A^T : \mathbb{R}^m \to \mathbb{R}^n$ as the adjoint operator to A and define the Dual conic program as

$$\max_{y \in \mathbb{R}^m} \ b^T y \tag{10}$$

$$s.t. \ c - A^T y \in K^* \tag{DUAL}$$

The importance of the dual problem lies in the following result

Theorem 4 (Duality and Slater's Conditions)**.** *Let p^* and d^* be the optimal values of PRIMAL and DUAL. We always have*

$$d^* \leq p^* \tag{11}$$

which we call weak duality. A sufficient condition for equality (or strong duality) is that either of Slater's Conditions is satisfied

(i) There is an x such that $Ax = b$ and $x \in int(K)$

(ii) There is a y such that $c - A^T y \in int(K^*)$

Many other sufficient conditions can be found in the literature, but this is the only we will make use of.

2.3 Conic Relaxations

We will consider optimization problems of the form

$$\inf_x \ f(x) \tag{P}$$

$$s.t \ \ x \in \mathcal{F}_P \tag{12}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is called the objective function and $\mathcal{F}_P \subset \mathbb{R}^n$ the feasible set. If v_P^* is the solution of (P), we define the set of optimal vectors of (P) as

$$\mathcal{O}_P = \{x \in \mathcal{F}_P : f(x) = v_P^*\} \tag{13}$$

We say that (P) is convex if both the objective function and the feasible set are convex. In the literature, a lot emphasis has been given to such problems, the reason being that under some conditions on the complexity of $\mathcal{F}_P$[1], these problems can be solved to a specified accuracy in polynomial time (on the size of the problem). This motivates convex approximations (or relaxations) to non-convex problems.

[1]The existence of an oracle of polynomial order

Definition 4 (Convex Relaxation)**.** We say that a problem (R) of the form

$$\inf_x \; g(x) \tag{R}$$

$$s.t \;\; x \in \mathcal{F}_R \tag{14}$$

is a convex relaxation of (P) if it satisfies the two conditions

(i) (R) is convex, i.e., both g and $\mathcal{F}_R$ are convex

(ii) There exists a function $F : \mathbb{R}^n \to \mathbb{R}^m$ such that if $x \in \mathcal{F}_P$ then $F(x) \in \mathcal{F}_R$ and $f(x) = g(F(x))$

We will call such F an equivalence relation.

If v_R^* and v_P^* are the optimal values of these problems, it follows easily from the definition of a convex relaxation that $v_R^* \leq v_P^*$. We will be interested in the cases when there is equality, and additionally, when we can translate from the optimal vectors of one problem to the optimal vectors of the other.

Definition 5 (Equivalence of a Problem and its Convex Relaxation)**.** We say that problem (P) and its convex relaxation (R) are equivalent if

(i) Their optimal values are equal, i.e, $v_P^* = v_R^*$

(ii) One problem attains its optimal value if and only if the other problem does, i.e., $\mathcal{O}_P \neq \emptyset$ iff $\mathcal{O}_R \neq \emptyset$

(iii) Either the set of optimal solution of (R) can be written as

$$\mathcal{O}_R = \text{Conv}(\{F(x^*) : x^* \in \mathcal{O}_P\}) \tag{15}$$

or there exists a convex function $G : R^m \to \mathbb{R}^n$ such that

$$\text{Conv}(\mathcal{O}_R) = G(\mathcal{O}_P) \tag{16}$$

Conditions (i) and (ii) are quite natural. Notice that the optimal set of (R) will always be convex, but the optimal set of (P) not necessarily, so we need to take convex hulls in condition (iii).

2.4 Recession and Horizon Cones

The concepts of the recession and horizon cones will be important in showing the equivalence between polynomial and copositive programs.

Definition 6 (Recession Cone)**.** We define the recession cone of $S \subseteq \mathbb{R}^n$ as the set

$$S^H := \{y \in \mathbb{R}^n : \forall x \in S \; \forall \lambda \in \mathbb{R}_+ \;\; x + \lambda y \in S\} \tag{17}$$

S^H indicates the directions we can move within S without leaving it. The Horizon Cone S^∞ (also known as the asymptotic cone), generalizes this definition

Definition 7 (Horizon Cone)**.** For $S \subset \mathbb{R}^n$, the horizon cone S^∞ is the set of y for which there exist sequences $\{x^k\} \subset S$ and $\{\lambda^k\} \subset \mathbb{R}_+$ such that $\lambda^k \to 0$ and $\lambda^k x^k \to y$. If $S = \emptyset$, we define $S^\infty = \{0\}$

Notice that both S^H and S^∞ are cones, convex whenever S is convex. We also have the inclusion $S^H \subseteq S^\infty$. The next proposition proves properties of S^∞ which we will use many times.

Proposition 2. *The horizon cone satisfies*

(i) If S is bounded, then $S^\infty = \{0\}$

(ii) If $S \subseteq T$ then $S^\infty \subseteq T^\infty$

(iii) $(S \cup T)^\infty = S^\infty \cup T^\infty$

(iv) $(S \cap T)^\infty \subseteq S^\infty \cap T^\infty$ and the reverse inclusion is not necessarily true.

(v) $S^H \subseteq S^\infty$

(vi) S^∞ is closed

(vii) If S is a cone then $S^\infty = cl(S)$

(viii) If S is closed and convex then $S^H = S^\infty$

(ix) Let $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ and $S = \{x \in \mathbb{R}^n : Ax \geq b\}$. Then if $S \neq \emptyset$, $S^\infty = \{y \in \mathbb{R}^n : Ay \geq 0\}$

Proof.

(i) Notice that S^∞ will always contain the zero. If S is bounded take M s.t. $\|x\| \leq M\ \forall x \in S$. Then assume $\{x^k\} \subset S$ and $\{\lambda^k\} \to 0$. We have that $\left\|\lambda^k x^k\right\| \leq |\lambda^k| M \to 0$, so necessarily, if $\lambda^k x^k$ converges it will converge to zero.

(ii) Any sequence in S will also be a sequence in T.

(iii) Using (ii): $S^\infty, T^\infty \subseteq (S \cup T)^\infty$, so $S^\infty \cup T^\infty \subseteq (S \cup T)^\infty$. To show the other inclusion, take $y \in (S \cup T)^\infty$ and a sequence $\{x^k\} \subset S \cup T$ such that $\lambda^k x^k \to y$. Since $\{x^k\} \subset S \cup T$, there exists a subsequence $\{x^{k_n}\}$ contained solely on S or in T. Notice that for this subsequence $\lambda^{k_n} x^{k_n} \to y$, so y will live either in S^∞ or in T^∞

(iv) Use (ii). To see that the reverse inclusion is not always true take $S = \{(x, 1/x) : x \in \mathbb{R}_{++}\}$ and $T = \{(x, -1/x) : x \in \mathbb{R}_{++}\}$. Then $S \cap T = \emptyset$, so $(S \cap T)^\infty = \{0\}$, but $S^\infty = T^\infty = S^\infty \cap T^\infty = \{(x, 0) : x \in \mathbb{R}_+\}$

(v) If S is empty this is trivial. Let $z \in S^H$ and choose $x \in S$. Then $y_n = x + nz \in S$ for all $n \in \mathbb{N}$. Define $\lambda_n = \frac{1}{n}$ and note that

$$\lambda_n y_n = \frac{1}{n} x + z \to z \tag{18}$$

so $z \in S^\infty$.

(vi) Assume $\{z^n\} \subseteq S^\infty$ is a convergent sequence with limit z. Without loss of generality assume $\|z^n - z\| \leq \frac{1}{2n}$. Since $z^n \in S^\infty$ there exist sequences $\lambda_k^n \geq 0$ and $\{x_k^n\} \subset S$ such that $\lim_{k\to\infty} \lambda_k^n = 0$ and $\lim_{k\to\infty} \lambda_k^n z_k^n = z^n$. Without loss of generality assume $\|\lambda_k^n z_k^n - z^n\| \leq \frac{1}{2k}$ and $|\lambda_k^n| \leq \frac{1}{k}$. Define the sequences $\mu_n = \lambda_n^n$ and $s_n = \mu_n z_n^n$. Since

$$|\mu_n| = |\lambda_n^n| \leq \frac{1}{n} \tag{19}$$

then $\mu_n \geq 0$ and $\mu_n \to 0$. Also

$$\|s_n - z\| = \|\lambda_n^n z_n^n - z\| \leq \|\lambda_n^n z_n^n - z^n\| + \|z^n - z\| \leq \frac{1}{2n} + \frac{1}{2n} = \frac{1}{n} \tag{20}$$

so $s_n = \mu_n z_n^n \to z$. Since $z_n^n \in S$, this shows $z \in S^\infty$

(vii) It is easy to check that $S \subseteq S^\infty$ and since S^∞ is closed then $\mathrm{cl}(S) \subseteq S^\infty$. To see the other direction take $z \in S^\infty$. Assume $\lambda_n z_n \to z$ for $\lambda_n \geq 0$, $\lim_{n\to\infty} \lambda_n = 0$ and $z_n \in S$. Since S is a cone then $\lambda_n z_n \in S$, so its limit z lives in $\mathrm{cl}(S^H)$.

(viii) We already showed $S^H \subseteq S^\infty$. To show the other inclusion take $z \in S^\infty$, take a pair $x \in S$, $\mu \geq 0$ and define $y = x + \mu z$. Since $z \in S^\infty$ there exists sequences $\{z^k\} \subset S$ and $\{\lambda^k\} \subset \mathbb{R}_+$ such that $\lambda^k \to 0$ and $\lambda^k z^k \to z$. For k large enough $0 \leq \mu\lambda_k \leq 1$, and because of convexity

$$(1 - \mu\lambda_k)x + (\mu\lambda_k)z_k \in S \tag{21}$$

Note that $\lim\limits_{k\to\infty} (1 - \mu\lambda_k)x + (\mu\lambda_k)z_k = x + \mu z$. Since S is closed we conclude $x + \mu z \in S$. As $x \in S$ and $\mu \geq 0$ were arbitrary, then $z \in S^\infty$, which shows the inclusion.

(ix) First, assume $y \in S^\infty$ and take a sequence $\{x^k\} \subset S$ and $\lambda^k \to 0$ so that $\lambda^k x^k \to y$. This implies $\lambda^k(Ax^k) \to Ay$, and from the definition of S, $\lambda^k(Ax^k) \geq \lambda^k b \to 0$. Hence $Ay \geq 0$.

To show the other inclusion, take y such that $Ay \geq 0$. If $Ay = M > 0$, define $x^k = ky$ and $\lambda^k = 1/k$. Clearly $\lambda^k x^k \to y$, and since $Ax^k = kM$, then for k large enough we will have $Ax^k \geq b$. If on the contrary we have $Ay = 0$, then find a solution x to the equation $Ax = b$ (which exists since $S \neq \emptyset$). Define $x^k = ky + x$ and $\lambda^k = 1/k$. Certainly $x^k \in S$ and $\lambda^k x^k \to y$.

□

Definition 8. For a polynomial $h \in \mathbb{R}_n^d[x]$ define $\tilde{h}(x)$ as the polynomial obtained after removing the terms of degree less than $\deg(h)$, so that $\tilde{h}(x)$ is homogeneous and $\deg(\tilde{h}) = \deg(h)$.

The next results states that the zeros of h recede to the zeros of $\tilde{h}$

Proposition 3. *We have the inclusion*

$$h^{-1}(0)^\infty \subseteq \tilde{h}^{-1}(0) \tag{22}$$

Proof. Assume h has degree d. Take a non-zero element $y \in h^{-1}(0)^\infty$. Then, there exists a sequence $\{x^k\} \subset h^{-1}(0)$ and a sequence of non-negative numbers $\{\lambda_k\} \subset \mathbb{R}_+$ such that $\lambda_k x^k \to y$. Since $y \neq 0$, we can assume that $x_i^k \neq 0$ for some i and for all k.

Decompose h into homogeneous polynomials h^D of degree D as

$$h = h^d + h^{d-1} + \cdots + h^1 + h^0 \tag{23}$$

then dividing by $(x_i^k)^d$ we obtain

$$\frac{h(x^k)}{(x_i^k)^d} = \frac{h^d(x^k)}{(x_i^k)^d} + \frac{h^{(d-1)}(x^k)}{(x_i^k)^d} + \cdots + \frac{h^0(x^k)}{(x_i^k)^d} \tag{24}$$

$$= h^d\left(\frac{x^k}{x_i^k}\right) + \frac{1}{(x_i^k)^1}h^{d-1}\left(\frac{x^k}{x_i^k}\right) + \cdots + \frac{1}{(x_i^k)^{d-1}}h^1\left(\frac{x^k}{x_i^k}\right) + \frac{1}{(x_i^k)^d}h^0\left(\frac{x^k}{x_i^k}\right) \tag{25}$$

Notice that

$$\lim_{k\to\infty} h^D\left(\frac{x^k}{x_i^k}\right) = h^D\left(\frac{y}{y_i}\right) \tag{26}$$

Also, since $\lambda_i^k x_i^k \to y_i \neq 0$, then $|x_i^k| \to \infty$. Using these two results and taking limits on both sides of (24) only the d term survives, so we obtain

$$\lim_{k\to\infty} \frac{h(x^k)}{(x_i^k)^d} = \lim_{k\to\infty} h^d\left(\frac{x^k}{x_i^k}\right) = h^d\left(\frac{y}{y_i}\right) = \frac{\tilde{h}(y)}{y_i^d} \tag{27}$$

Finally, since $h(x^k) = 0$, we get $\tilde{h}(y) = 0$. That completes the proof. □

2.5 Tensors

Any symmetric matrix $A \in \mathcal{S}^{n\times n}$ gives rise to an n variable quadratic polynomial through $q_A(x) := x^T Ax$. Many properties of q_A can be translated to properties of A, for example, q_A is non-negative iff $X \succeq 0$.

To bring this idea to polynomials of higher degree we need a generalization of the set of symmetric matrices. Here we use the vector space of symmetric tensors. We do not concern ourselves with a proper definition, so for a very good discussion of tensor products and multilinear algebra we refer the reader to the first chapter of Greub's book [28].

Denote the space of tensors of dimension d on n variables as

$$\mathcal{T}_n^d = \underbrace{\mathbb{R}^n \otimes \ldots \otimes \mathbb{R}^n}_{d \text{ times}} \tag{28}$$

so that any element $x \in \mathcal{T}_n^d$ can be described as $x = [x_{i_1,\ldots,i_d}]_{1\leq i_k \leq n}$. And denote $\mathcal{S}_n^d$ as the vector space of symmetric tensors, that is, the set of tensors x such that for any permutation $\sigma \in S^d$

$$x_{i_1,\ldots,i_d} = x_{\sigma(i_1,\ldots,i_d)} \tag{29}$$

Notice that each component of a symmetric tensor x is indexed by a sequence $(i_1, \ldots, i_d)$ of numbers taken from $\{1, \ldots, n\}$ and the symmetry means we do not care about the order. This means the dimension of $\mathcal{S}_d^n$ is given by the multiset coefficient

$$\dim(\mathcal{S}_d^n) = \binom{n+d-1}{d} \tag{30}$$

We endow this vector space with the frobenius inner product

$$\langle x, y\rangle = \sum_{1\leq i_1,\ldots,i_d\leq n} x_{i_1,\ldots,i_d} y_{i_1,\ldots,i_d} \tag{31}$$

and define the map $M_d : \mathbb{R}^n \to \mathcal{S}_n^d$.

$$M_d(x) = \underbrace{x \otimes \ldots \otimes x}_{d \text{ times}} \tag{32}$$

Notice that $M_d(x)_{i_1,\ldots,i_d} = x_{i_1} \ldots x_{i_d}$. In the case $d = 2$ we are in the familiar space of symmetric matrices with the frobenius product and $M_2(x) = xx^T$.

2.6 Spectrahedral cones

Definition 9 (Spectrahedron)**.** We say that $G \subset \mathbb{R}^n$ is a spectrahedron if there exist symmetric matrices $Q_0, Q_1, \ldots, Q_n \in \mathcal{S}^{d\times d}$ such that

$$G = \{x \in \mathbb{R}^n : Q(x) := Q_0 + x_1 Q_1(x) + \cdots + x_n Q_n \succeq 0\} \tag{33}$$

In the case that $Q_0 = 0$ we call G a conic spectrahedron and Q the linear operator of G.

The main objective of this section is to find a way to describe the dual G^* of a conic spectrahedron, a result that will be useful in finding spectrahedral approximations of the completely positive cone. For an extensive study of the geometry of spectrahedra we remit the reader to the excellent paper by Ramana and Goldman [30].

Lemma 1. *Let G be a conic spectrahedron with operator $Q : \mathbb{R}^n \to \mathcal{S}^{d\times d}$*

$$Q(x) = x_1 Q_1 + \cdots + x_n Q_n \tag{34}$$

Then the adjoint operator $L : \mathcal{S}^{d\times d} \to \mathbb{R}^n$ of Q is given as

$$[L(Y)]_i = \langle Q_i, Y\rangle \textit{ for } 1 \le i \le n \tag{35}$$

Proof. The result follows easily from the fact that

$$\langle Q(x), Y\rangle = \langle x, L(Y)\rangle \ \ \forall x \in \mathbb{R}^n, \forall Y \in \mathcal{S}^{d\times d} \tag{36}$$

□

Proposition 4. *Let $G \subset \mathbb{R}^n$ be a conic spectrahedron with operator Q and dual operator L. Define*

$$\hat{G} := \{L(Y) : Y \succeq 0\} \tag{37}$$

Then $G^* = Cl(\hat{G})$

Proof. Let $x \in G$ and $Y \succeq 0$. Since $Q(x), Y \in \mathcal{S}^{d\times d}$ and $\mathcal{S}^{d\times d}$ is self-dual, then

$$\langle x, L(Y)\rangle = \langle Q(x), Y\rangle \ge 0 \tag{38}$$

implying that $L(Y) \in G^*$. This shows that $\hat{G} \subseteq G^*$ and since G^* is closed then $\mathrm{cl}(\hat{G}) \subseteq G^*$. To show the other inclusion define $H = Cl(\hat{G})$ and take $w \in H^*$. This means that

$$\langle w, L(Y)\rangle \ge 0 \text{ whenever } Y \succeq 0 \tag{39}$$

hence $\langle w, L(Y)\rangle = \langle Q(w), Y\rangle \ge 0$ for all $Y \succeq 0$, implying that $Q(w) \succeq 0$. Since w was taken from H^*, then $H^* \subseteq G$. If we take duals and use the fact that $H^{**} = H$ we obtain the other inclusion. □

2.7 Integrating over the Sphere

To calculate a spectrahedral approximation of the copositive cone we must integrate polynomials over the nonnegative portion of the sphere S^{n-1}. A slight modification of a result by Folland [8] allows us to calculate these integrals in closed form.

Theorem 5. *Let X be the intersection of S^{n-1} and the nonnegative orthant. For a polynomial $P(x) = x^{\alpha_1}x^{\alpha_2}\dots x^{\alpha_n}$ with $\alpha_i \in \mathbb{N}$*

$$\int_X P(x)d\mu = \frac{1}{2^{n-1}}\frac{\Gamma(\gamma_1)\dots\Gamma(\gamma_n)}{\Gamma(\gamma_1+\dots\gamma_2)} \tag{40}$$

where $\gamma_j = \frac{1}{2}(\alpha_j + 1)$.

Proof. Consider the integral

$$I = \int_{\mathbb{R}^d_+} P(x)e^{-|x|^2} \tag{41}$$

First, evaluate I in cartesian coordinates to obtain

$$I = \prod_{j=1}^{n}\int_0^\infty x^{\alpha_j}e^{-x_j^2}dx_j = \prod_{j=1}^{n}\frac{1}{2}\int_0^\infty u_j^{\gamma_j-1}e^{-u_j}du_j \tag{42}$$

where we used the substitution $u_j = x_j^2$. From this we get

$$I = \frac{1}{2^n}\Gamma(\gamma_1)\dots\Gamma(\gamma_n) \tag{43}$$

Second, since $P(r\hat{x}) = r^{\alpha_1+\dots+\alpha_n}P(\hat{x})$, we calculate I using polar coordinates and obtain

$$\begin{aligned} I &= \int_X\int_0^\infty P(r\hat{x})e^{-r^2}r^{n-1}drd\mu(\hat{x}) && (44)\\ &= \int_X P(\hat{x})d\mu(\hat{x})\int_0^\infty r^{\alpha_1+\dots+\alpha_n+n-1}e^{-r^2}dr && (45)\\ &= \frac{\Gamma(\gamma_1+\dots\gamma_2)}{2}\int_X P(x)d\mu && (46)\end{aligned}$$

Equating both expressions for I give us the result. □

3 The Copositive and Completely Positive Cones

In this section we will study the cone of completely positive matrices and its dual, the cone of copositive matrices. We will focus on their interiors, their extreme rays and their geometry in small dimensions. Define the cones of matrices

$$\mathcal{C}_n := \{BB^T : \exists k \in \mathbb{N} : B \in \mathbb{R}^{n\times k} \text{ and } B_{ij} \geq 0\} \tag{47}$$

$$\mathcal{C}_n^* := \{A \in S^{n\times n} : x^TAx \geq 0\ \forall x \in \mathbb{R}^n, x \geq 0\} \tag{48}$$

We call $\mathcal{C}_n$ the cone of completely positive matrices and $\mathcal{C}_n^*$ the cone of copositive matrices. Normally we will omit the n and refer to these cones as $\mathcal{C}$ and $\mathcal{C}^*$. Notice that the definition of the copositive cone is equivalent to

$$\mathcal{C}_n := \left\{\sum_{i=1}^{k} b_ib_i^T : b_i \in \mathbb{R}^n, k \in \mathbb{N}, b_i \geq 0, 1 \leq i \leq k\right\} \tag{49}$$

since we can take the b_i as the columns of B.

As the notation suggests they are dual to each other, to show it we need the following result.

Lemma 2. *The cone of copositive matrices $\mathcal{C}^*$ is a proper cone*

Proof.

(i) Convexity is immediate from the definition

(ii) Assume $A_k \to A$ and $A_k \in \mathcal{C}^*$. Because of continuity $x^T A_n x \to x^T A x$, so we must have $x^T A x \geq 0$ for all $x \in \mathbb{R}_+$. This implies $\mathcal{C}^*$ is closed.

(iii) If $A, -A \in \mathcal{C}^*$ then $x^T A x = 0 \ \forall x \geq 0$. Because $x^T A x$ is analytic we must have $x^T A x = 0$, implying $A = 0$.

(iv) Finally, $\mathcal{C}^*$ has non-empty interior because it contains the positive-semidefinite cone, which has non-empty interior.

□

Theorem 6. *The dual of the cone of completely positive matrices $\mathcal{C}$ is the cone of copositive matrices $\mathcal{C}^*$*

Proof. To avoid confusions, denote COP and COM as the cones of copositive and completely positive matrices, respectively. We must show $COM^* = COP$. To prove $COM^* \subseteq COP$ take $A \in COM^*$ and $x \geq 0$. Since $x^T A x = \langle A, xx^T \rangle$ and $xx^T \in COM$, then $x^T A x \geq 0$ for all $x \geq 0$. This makes $A \in COP$.

Because COP is proper, we can prove $COM \subseteq COP^*$ and then take the dual at both sides to obtain the other inequality. So let us take $BB^T = \sum b_i b_i^T \in COM$ with $b_i \geq 0$ and show that it lives in COP^*. If $A \in COP$, then $\langle BB^T, A \rangle = \langle \sum b_i b_i^T, A \rangle = \sum b_i^T A b_i \geq 0$, being A copositive. Since A was arbitrary then $BB^T \in COP^*$. □

3.1 The interior of the cones

The next theorem characterizes the interior of $\mathcal{C}^*$.

Theorem 7. *The interior of the copositive cone, called the strictly copositive cone, can be described as*

$$int(\mathcal{C}^*) = \{A : x^T A x > 0 : \forall x \geq 0, x \neq 0\} \tag{50}$$

Proof. The condition $x^T A x > 0 \ \forall x \geq 0, x \neq 0$ is equivalent to the form $x^T A x$ being positive on $D := \{x \in \mathbb{R}^n : \|x\| = 1, x \geq 0\}$. Consider the continuous function $f : \mathcal{S}^{n \times n} \times D \to \mathbb{R}$ defined as $f(A, x) = x^T A x$. Since D is compact, then the function defined as

$$g(A) = \min_{x \in D} f(A, x) \tag{51}$$

is continuous. Since $x^T A x$ is positive on D, then $g(A) > 0$. Using the continuity of g, there exist a neighbourhood of A, call it U_A, where g is positive. All matrices on U_A will be copositive, showing

$A \in \text{int}(\mathcal{C}^*)$

Now, assume that there is a matrix $A \in \text{int}(\mathcal{C}^*)$ and an $x \geq 0$, $x \neq 0$ such that $x^T Ax = 0$. Define $B_\lambda = A + \lambda xx^T$ and take a neghborhood U_A of A. There exists a negative λ small enough that $B_\lambda \in U_A$. However, $x^T B_\lambda x = x^T Ax + \lambda(x^T x)^2 < 0$. This contradicts the fact there is some neighborhood of A composed exclusively of copositive matrices. □

One of the first characterizations of $\text{int}(\mathcal{C})$ was given by Dür and Still [10]. Then, Dickinson [9] came up with the more compact description

$$\text{int}(\mathcal{C}_n) = \{AA^T : A_i > 0 \text{ for some } i \text{ and } rank(A) = n\} \tag{52}$$

To show this, first define the zeros of $A \in \mathbb{R}^{n\times n}$ in the positive orthant as

$$\mathcal{Z}^A := \{x \in \mathbb{R}^n_+ : x^T Ax = 0\} \tag{53}$$

Lemma 3. *If $A \in \mathcal{C}^*$ and there exists a vector $x \in \mathcal{Z}^A$ such that $x > 0$, then $A \in \mathcal{S}_+$*

Proof. Assume $A \neq 0$. Take an arbitrary $u \in \mathbb{R}^n$ and notice that for $\epsilon > 0$ small enough $y := x + \epsilon u > 0$. Thus

$$0 \leq y^T Ay = x^T Ax + 2\epsilon u^T Ax + \epsilon^2 u^T Au = \epsilon(2u^T Ax + \epsilon u^T Au) \tag{54}$$

Dividing by ϵ and letting $\epsilon \to 0$ we obtain $u^T Ax \geq 0$, and since u was arbitrary, it must be that $Ax = 0$. Replacing back we obtain

$$0 \leq y^T Ay = \epsilon^2 u^T Au \tag{55}$$

for $\epsilon > 0$. This means $u^T Au \geq 0$ and since u was arbitrary, $A \in \mathcal{S}_+$. □

The following lemma is an intermediate characterization of the interior of $\mathcal{C}^*$

Lemma 4. *Let $U = \sum_{i=1}^k u_i u_i^T \in \mathcal{C}$ for $u_i \in \mathbb{R}^n_+$. Then U is in the interior of $\mathcal{C}$ if and only if there is no copositive matrix $A \neq 0$ such that $u_i \in \mathcal{Z}^A$ for all i.*

Proof. From the geometry of cones we know that if $\mathcal{K} \subset \mathbb{R}^n$ is a proper cone, then

$$\text{int}(\mathcal{K}^*) = \{B \in \mathbb{R}^n : \langle A, B\rangle > 0,\ \forall A \in \mathcal{K},\ A \neq 0\} \tag{56}$$

Taking $\mathcal{K} = \mathcal{C}^*$, $U = \sum_{i=1}^k u_i u_i^T$ will be in the interior of $\mathcal{C}$ if and only if

$$\forall A \in \mathcal{C}^* : A \neq 0,\ \langle A, U\rangle = \sum u_i A u_i^T > 0 \tag{57}$$

and this is equivalent to

$$\nexists A \in \mathcal{C}^* : A \neq 0 \text{ and } \forall i : u_i A u_i^T = 0 \tag{58}$$

□

Theorem 8. *Let $U \in \mathcal{C}$. Then $U \in int(\mathcal{C})$ if and only if it can be written as $\sum_{i=1}^k u_i u_i^T$ for $u_i \in \mathbb{R}^n_+$, and the next conditions are satisfied*

1. *$span(u_1, \ldots, u_m) = \mathbb{R}^n$*

2. For some j, $v := u_j > 0$

Equivalently

$$int(\mathcal{C}) = \mathcal{M} := \{AA^T : rank(A) = n,\ A = [a|C], a > 0, C \geq 0\} \tag{59}$$

Proof. To show that the condition is sufficient let us argue by contradiction, so assume $U \in \mathcal{M}$ but $U \notin \text{int}(\mathcal{C})$. Lemma 4 says there must be a copositive matrix $A \neq 0$ such that $u_i \in \mathcal{Z}^A$ for all i. In particular, $v > 0$ and $v \in \mathcal{Z}^A$. Lemma 3 then implies that $A \in \mathcal{S}_+$. Recall that for a psd matriz A, $x^T Ax = 0$ implies $Ax = 0$. Since $u_i^T Au_i = 0$, then $Au_i = 0$ for all i. But $\text{span}(u_1, \ldots, u_m) = \mathbb{R}^n$, so it must be that $A = 0$ and this is a contradiction.

To prove that the condition is necessary take $B \in \text{int}(\mathcal{C})$, we will show that B admits the description of $\mathcal{M}$. To do this take any matrix $F \neq B$ such that $F \in \mathcal{M} \subseteq \mathcal{C}$ and construct

$$P = B - \beta F \in \mathcal{C} \tag{60}$$

Choose β positive and small enough that $P \in \mathcal{C}^*$. This is possible since B is the interior of $\mathcal{C}$. Notice that

$$B = P + \beta F = P + \hat{F} \tag{61}$$

Since $P \in \mathcal{C}$ then $P = YY^T$ for some $Y \geq 0$. Also, since $\hat{F} \in \mathcal{M}$ then $\hat{F} = [a|\hat{C}]$ for $a > 0$ and $\text{rank}(\hat{C}) = n$. Now we can define $A = [a|\hat{C}|Y]$ to find that

$$AA^T = [a|\hat{C}][a|\hat{C}]^T + YY^T = B \tag{62}$$

It remains to show that $C := [\hat{C}|Y]$ has rank n, but this is true since $\hat{C}$ does. □

3.2 The extreme rays of the cones

As we mentioned in the Preliminaries, the extreme rays of a cone are important because they provide the smallest description of a cone that it is possible to come up with. We will first characterize the extreme rays of $\mathcal{C}$. We need the next lemma:

Lemma 5. *If $C = A + B$ and $A, B \in S_+$ then $Null(A) \cap Null(B) = Null(C)$*

Proof. C must be psd. Assume that $x \in Null(C)$, so that

$$0 = x^T Cx = x^T Ax + x^T Bx \tag{63}$$

Since $x^T Ax, x^T Bx \geq 0$ it must be that $x^T Ax = x^T Bx = 0$, and since A, B are psd, we must have $Ax = Bx = 0$. To show the other direction assume

$$x \in \text{Null}(A) \cap \text{Null}(B) \tag{64}$$

It is automatic that $Cx = Ax + Bx = 0$ and it must be that $x \in \text{Null}(C)$ □

Theorem 9. *The extreme matrices of the completely positive cone $\mathcal{C}$ are the matrices xx^T with $x \geq 0$.*

Proof. Assume xx^T can be written as the sum of two matrices $A, B \in \mathcal{C}$

$$xx^T = A + B \tag{65}$$

Since $\mathcal{C} \subset S_+$, Lemma 5 implies that $\text{Null}(A) \cap \text{Null}(B) = \text{Null}(xx^T)$. Which implies that $\text{Null}(A) \supset \text{Null}(xx^T)$. The only way this can happen is if $\text{Null}(A) = \mathbb{R}^n$ or $\text{Null}(A) = \text{Null}(xx^t)$, and implies that for some $\lambda_A \geq 0$ then $A = \lambda_A(xx^T)$. Analogously, $B = \lambda_B(xx^T)$ for some λ_B, so xx^T must be an extreme matrix. Since any matrix in $\mathcal{C}$ can be written as the sum of elements xx^T with $x \geq 0$ then the matrices xx^T can be the only extreme matrices. □

Characterizing the extreme rays of the copositive cone is a much more difficult problem, one that is still open. In this section we describe some of the well-known extreme vectors of $\mathcal{C}^*$. We draw this discussion from [11]. First we need the lemma

Lemma 6. *If A is copositive then*

(i) A has a nonnegative diagonal

(ii) If $a_{ii} = 0$ for some i, then the i-th row and i-th column of A are nonnegative

Proof. Notice that if A is copositive, $a_{ii} = e_i^T A e_i \geq 0$, so A must have a non-negative diagonal. To show (ii) assume $a_{ii} = 0$ and define $x = e_i + \epsilon e_j$ for $j \neq i$ and $\epsilon > 0$. Then

$$0 \leq x^T A x = a_{ii} + 2\epsilon a_{ij} + \epsilon^2 a_{jj} = \epsilon(2a_{ij} + \epsilon a_{jj}) \tag{66}$$

Dividing by ϵ we get $2a_{ij} + \epsilon a_{jj} \geq 0$ for all $\epsilon > 0$, so it must be that $a_{ij} \geq 0$. The result follows from symmetry of A and the fact that j was arbitrary. □

Theorem 10. *The following are extreme vectors of the copositive cone:*

(i) E_{ii} for $1 \leq i \leq n$

(ii) $E_{ij} + E_{ji}$ for $1 \leq i, j \leq n$

(iii) xx^T where $x \in \mathbb{R}^n$ has at least one positive component and one negative component.

Proof. Assume that $E_{ii} = A + B$ for $A, B \in \mathcal{C}^*$. Then it must be that $a_{jj} + b_{jj} = 0$ for $j \neq i$. Using Lemma 6

$$a_{jj}, b_{jj} \geq 0 \implies a_{jj} = b_{jj} = 0 \tag{67}$$

Using the lemma again, for any $j \neq i$ and any k, $a_{jk} \geq 0$ and $b_{jk} \geq 0$. But since $a_{jk} + b_{jk} = 0$ we conclude $a_{jk} = b_{jk} = 0$. Since A and B are symmetric, all their entries except for the ii must be 0, hence

$$A = a_{ii}E_{ii}, \;\; B = b_{ii}E_{ii} \tag{68}$$

which shows E_{ii} is an extremal vector.

To show that $E_{ij} + E_{ji}$ is also an extreme vector, assume $E_{ij} + E_{ji} = A + B$. Using Lemma 6 (i) we conclude again that $a_{ll} = b_{ll} = 0$ for all l, and using part (ii), it must be that all elements of A and B are nonnegative. Since $E_{ij} + E_{ji} = A + B$, and both $A, B \geq 0$, it must be that the only non-zero entries of A and B are the ij and ji. Finally, symmetry implies that A and B are multiples of $E_{ij} + E_{ji}$.

Let us prove part (iii). First assume that C is an invertible matrix and $Cx \geq 0$ iff $x \geq 0$. It is easy to see that X is an extreme matrix if an only if $C^{-1}XC$ is an extreme matrix. Notice that if C is a permutation or C is a strictly positive diagonal matrix, it satisfies these conditions. This allows us to restrict our proof to the case of matrices xx^T that satisfy

$$x_1 = \cdots = x_k = 1, \;\; x_{k+1} = \cdots = x_m = -1, \;\; x_{m+1} = \cdots = x_n = 0 \tag{69}$$

For $n = 2$, the proof reduces to showing that the matrix xx^T with $x = [1, -1]^T$ is an extreme ray. Assume

$$A + B = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{70}$$

We can assume, without lose of generality, that

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \tag{71}$$

with $a, c \geq 0$ and $b < 0$. Since $0 = e^T xx^T e = e^T Ae + e^T Be$ and $e^T Ae, e^T Be \geq 0$, then $e^T Ae = 0$, which means

$$a + 2b + c = 0 \tag{72}$$

This implies that for any $y = (y_1, y_2)^T$

$$\begin{aligned} y^T Ay &= y_1^2 a + 2by_1y_2 + y_2^2 c && (73) \\ &= y_1^2 a + 2by_1y_2 - y_2^2(a + 2b) && (74) \\ &= (y_1 - y_2)(a(y_1 + y_2) + 2by_2) && (75) \end{aligned}$$

so taking $y = (1 + \epsilon, 1)$ we obtain

$$y^T Ay = \epsilon(2(a + b) + \epsilon a) \tag{76}$$

If $a+b > 0$, choose $\epsilon < 0$ small enough that $2(a+b)+\epsilon a > 0$. If $a+b < 0$, choose $\epsilon > 0$ small enough that $2(a + b) + \epsilon a < 0$. In both cases $y^T Ay < 0$, (contradicting the copositivity of A) so we must have that $a+b = 0$. This fact, along with (72), implies that $a = c = -b$, so necessarily A is a multiple of xx^T. Analogously, B must be a multiple of xx^T, and we conclude that xx^T is an extreme ray.

Now, let us go back to the general case, i.e., x is as in (69). We have

$$xx^T = \begin{bmatrix} X_{11} & X_{12} & 0 \\ X_{12}^T & X_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{77}$$

where $X_{11} \in \mathbb{R}^{k\times k}$ and $X_{22} \in \mathbb{R}^{(m-k)\times(m-k)}$ are matrices of 1s, and $X_{12} \in \mathbb{R}^{k\times(m-k)}$ is a matrix of -1s.

Assume $xx^T = A + B$. Using Lemma 6 it is easy to show that $a_{ij} = b_{ij} = 0$ whenever $i > m$ or $j > m$. Taking the submatrix $[i, j]$ of xx^T, for $i \leq k$ and $j > k$ we obtain

$$xx^T[i, j] = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{78}$$

Since this is an extreme ray, and the matrices in the equation $xx^T[i,j] = A[i,j] + B[i,j]$ are all copositive, we obtain

$$A[i,j] = a_{ii} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{79}$$

therefore, $a_{ii} = a_{jj} = -a_{ij}$ whenever $i \leq k$ and $j > k$. If we vary the indices i and j, we obtain a series of equations that necessarily imply $A = a_{11}xx^T$. Since $xx^T = A + B$, then B is also a multiple of xx^T, and we conclude that xx^T is an extreme ray. □

3.3 The geometry in small dimensions

We define $\mathcal{DN} := \mathcal{S}_+ \cap \mathcal{N}$ and call it the cone of doubly nonnegative matrices, i.e., matrices which are both nonnegative and psd. The inclusion

$$\mathcal{C} \subseteq \mathcal{S}_+ \cap \mathcal{N} \tag{80}$$

follows easily from the definition of $\mathcal{C}$. It is easy to show that if A, B are convex cones, then $(A \cap B)^* = A^* + B^*$. Applying this to $\mathcal{DN}$ we obtain $\mathcal{DN}^* := \mathcal{S}_+ + \mathcal{N}$ and the inequality

$$\mathcal{S}_+ + \mathcal{N} \subseteq \mathcal{C}^* \tag{81}$$

An interesting question is how close are the cones $\mathcal{C}$ and $\mathcal{DN}$. A good way to answer this is trying to understand the extreme rays of $\mathcal{DN}$. Hamilton and Lee [12], for example, give an algorithm to determine if a matrix is an extreme ray of $\mathcal{DN}$. One of their results states

Theorem 11. *There exist extreme matrices in $\mathcal{DN}_n$ of rank k if and only if*

$$k \leq \begin{cases} max(1, n-3) & n \text{ is even} \\ max(1, n-2) & n \text{ is odd} \end{cases} \tag{82}$$

From Theorem 11, we can deduce that if $n \leq 4$, the extreme rays of $\mathcal{DN}_n$ are all of rank 1. Since the only non-negative, psd matrices of rank 1 are of the type xx^T with $x \in \mathbb{R}^n_+$, then all the extreme matrices of $\mathcal{DN}_n$ are completely positive. This implies that $\mathcal{DN}_n = \mathcal{C}_n$ for $n \leq 4$. This result was first discovered by Diananda [14] and we state it for reference.

Theorem 12. *For $n \leq 4$, the cone of doubly nonnegative matrices $\mathcal{DN}_n = \mathcal{S}^n_+ \cap \mathcal{N}_n$ is equal to the cone of completely positive matrices $\mathcal{C}$. From duality, $\mathcal{C}^* = \mathcal{S}^n_+ + \mathcal{N}_n$*

This theorem states that for $n \leq 4$, any doubly nonnegative matrix A admits a factorization $A = XX^T$ with X is nonnegative. For $n = 2$, for example, we can take the matrix

$$X = \begin{bmatrix} \sqrt{a_{11}} & \frac{a_{12}}{\sqrt{a_{11}}} \\ 0 & \frac{\sqrt{|A|}}{\sqrt{a_{11}}} \end{bmatrix} \tag{83}$$

We may ask ourselves what happens with $n \geq 0$. Unfortunately, for the case $n = 5$ Horn [29] gave an example of a matrix which is in $\mathcal{C}^*$ but not in $\mathcal{S}_+ + \mathcal{N}$. This is proved next

Example 1. The Horn matrix

$$H = \begin{bmatrix} 1 & -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 & 1 \end{bmatrix} \tag{84}$$

is an extreme matrix of $\mathcal{C}_5^*$, but it is not the sum of a psd and a nonnegative matrix.

Proof. The form $x^T Hx$ can be written as

$$\begin{aligned} Q(x) = x^T Hx &= (x_1 + x_2 + x_3 + x_4 + x_5)^2 - 4x_1x_2 - 4x_2x_3 - 4x_3x_4 - 4x_4x_5 - 4x_5x_6 && (85) \\ &= (x_1 - x_2 + x_3 + x_4 - x_5)^2 + 4x_2x_4 + 4x_3(x_5 - x_4) && (86) \\ &= (x_1 - x_2 + x_3 - x_4 + x_5)^2 + 4x_2x_5 + 4x_1(x_4 - x_3) && (87) \end{aligned}$$

Equation shows $Q(x)$ is nonnegative on $x \in \mathbb{R}_+$ when $x_5 \geq x_4$ and Equation shows $Q(x)$ is nonnegative when $x_5 \leq x_4$. This means H is copositive.

To show that H is an extreme matrix of $\mathcal{C}^*$ assume $H = A + B$ with $A, B \in \mathcal{C}^*$. Notice that if we choose $i, j \in \{1, \ldots, 5\}$ and remove the i, j columns and rows of H we obtain a 3×3 submatrix of the form xx^T for $x \in \{[1, 1, -1]^T, [1, -1, 1]^T, [-1, 1, -]^T\}$. As was shown in Theorem 10, all of these submatrices are extreme. This implies that the 3×3 submatrices of A and B are all a multiple of the corresponding submatrices of H. The only way this can happen is if A and B are a multiple of H.

Finally, notice that H is not nonegative. Also, if $x = [1, 1, 1, -1, -1]^T$ then $x^T Hx = -3$, so H is not psd. Since the extreme rays of $\mathcal{S}_+ + \mathcal{N}$ either belong to $\mathcal{S}_+$ or to $\mathcal{N}$, then H is a copositive matrix not on $\mathcal{DN}$. □

4 The BVL Approximation

Recall that if we have a polynomial oracle for the membership of a cone K, then we can solve linear conic programs efficiently through the ellipsoid method. For cones like the copositive cone of matrices such oracles are unknown and their existence would imply $P = NP$.

In this section we try to go around this problem by constructing spectrahedral approximations of proper cones (remember that semidefinite problems can be solved in polynomial time through interior point methods). Specifically, we show that any proper cone K can be approximated by a hierarchy of conic spectrahedra that converges to K pointwise. We also perform explicit calculations for the hierarchy of the copositive cone and its dual.

These ideas were proposed independently by Barvinok and Veomett [4] and Laserre [2]. The first two sections follow Romero and Velasco [3].

4.1 Admissible pairs and proper cones

First we need the technical definition of an admissible pair.

Definition 10 (Admissible Pairs)**.** Let X be a compact topological space with a finite Borel measure μ supported on X. Assume there exists a continuous embedding $\phi : X \to V$ into a finite dimensional vector space, and a linear function $g : V \to \mathbb{R}$. We say that the pair (ϕ, g) is admissible if

$$\mathrm{Aff}(\phi(X)) = g^{-1}(1) \tag{88}$$

To any admissible pair (ϕ, g) we associate the convex cone $\mathcal{K} := \mathrm{Conic}(\phi(X))$. The role of ϕ will be in translating the measure μ over X to a measure over the cone $\mathcal{K}$. The role of g is in making the affine hull of $\phi(X)$ a hyperplane, ensuring that $\mathcal{K}$ is a proper cone. This is shown in the next theorem.

Proposition 5. *The associated cone $\mathcal{K}$ of an admissible pair (ϕ, g) is a proper cone*

Proof.

(i) $\mathcal{K}$ is the continuous image of a compact topological space, so it must be closed

(ii) The fact that $g(\mathcal{K}) = \{1\}$ insures that $\mathcal{K}$ is pointed

(iii) Notice that $g^{-1}(1)$ is an affine hyperplane, therefore, $\mathrm{Aff}(\phi(X))$ is an affine hyperplane. In order to check whether $\mathcal{K}$ has non-empty interior it suffices to check that $\mathrm{Conv}(\phi(X))$ has non-empty relative interior. Notice that $\int_X \phi \, d\mu \in \mathrm{RelInt}(C)$.

□

Since $\mathcal{K}$ is proper, $\mathcal{K}^*$ will also be proper. It is easy to see that g is in the interior of C^*. The reason is $\phi(X)$ is compact, so we can move g slightly and still obtain a positive values for $g(\phi(X))$. The next theorem show that we can always think of proper cones as cones associated with an admissible pair.

Proposition 6. *For any proper cone K on a finite dimensional vector space, there exists an admissible pair (ϕ, g) associated with K.*

Proof. Take a proper cone K on a finite dimensional vector space V. Choose $x \in \mathrm{int}(K)$ and $y \in \mathrm{int}(K^*)$, and define $\mathcal{H}$ as the affine hyperplane that passes through x and is orthogonal to y. Let us first show that the intersection $X := K \cap \mathcal{H}$ must be bounded. Assume by contradiction that there exists an unbounded sequence $\{v_k\} \subseteq X$, where $v_k = x + h_k$ for some h_k that is orthogonal to y. Choose a subsequence $\{h_{k_n}\}$ such that

$$\frac{h_{k_n}}{\|h_{k_n}\|} \to h \tag{89}$$

which can be done from the compactness of the unit ball. Notice that since X is convex and $x, x + h_{k_n} \in X$, then

$$h_n^\lambda := x + \lambda \frac{h_{k_n}}{\|h_{k_n}\|} \in X \tag{90}$$

for $\lambda < \|h_{k_n}\|$. Define

$$h^\lambda := x + \lambda h = \lim_{n \to \infty} h_n^\lambda \tag{91}$$

and notice that for n large enough $h_n^\lambda \in X$. From the closedness of X, we conclude that $h_\lambda \in X$ for all $\lambda \geq 0$. Now choose $\epsilon > 0$ small enough that $y_\epsilon := y - \epsilon h \in K^*$ and notice that

$$\left\langle y_\epsilon, h^\lambda \right\rangle = \langle y, x \rangle - \epsilon \langle h, x \rangle - \epsilon\lambda \|h\|^2 \tag{92}$$

and we can make this product negative for λ large enough, contradicting the fact that $h^\lambda \in K$ and $y_\epsilon \in K^*$.

We showed that X is bounded. It is also closed and has non-empty relative interior ($x \in \mathrm{RelInt}(X)$). To finish the proof, we endow X with the subspace topology and give it the Lebesgue measure μ as a subset of its affine hull. Since X is closed and bounded, it will be compact, and since X has non-empty relative interior, μ will be supported on X. We define $\phi : X \to V$ as the identity function and $g : V \to \mathbb{R}$ as the unique linear function that takes the value of 1 on the affine hyperplane $\mathcal{H}$. To complete the proof, notice that

$$\mathrm{Aff}(\phi(X)) = g^{-1}(1) \tag{93}$$

□

The next examples show that we can treat the traveling salesman cone, the cone of semidefinite tensors and the cone of copositive tensors as cones associated to an admissible pair. This will be essential in constructing efficient approximations to these cones.

Example 2. Let X be the set of all hamiltonian paths in N cities and μ the discrete uniform measure. Define $\phi : X \to V$ as the adjacency matrix of each path and $V = span(\phi(x))$. Let $g : V \to \mathbb{R}$ be the map that assigns to each matrix the sum of its entries divided by $2n$. Then (ϕ, g) is admissible and $\mathcal{K}$ is the conic hull of the traveling salesman polytope

Proof. To see that (ϕ, g) is admissible notice that $\mathcal{H} := g^{-1}(1) \supseteq \phi(X)$ so that

$$\mathcal{H} \supseteq \mathrm{Aff}(\phi(X)) \tag{94}$$

Since $span(\phi(X)) = V$ then $\dim(\mathrm{Aff}(\phi(X))) \geq n-1$. Since (94) also implies that $\dim(\mathrm{Aff}(\phi(X))) \leq n-1$, it must be that

$$\mathrm{Aff}(\phi(X)) = g^{-1}(1) \tag{95}$$

Since $\phi(X)$ is the TSP, then $\mathcal{K}$ is the cone over the TSP. □

Example 3. Let $X \subseteq \mathbb{R}^n$ be the $n-1$ sphere with the uniform measure. Define $\phi : X \to \mathcal{S}_n^{2d}$ as the map M_{2d} (defined in the preliminaries) that sends x to the symmetric tensor $x^{\otimes 2d}$ and let $g : \mathcal{S}_n^{2d} \to \mathbb{R}$ be the trace of the tensor. Then (ϕ, g) is admissible and $\mathcal{K}^*$ is the cone of semidefinite tensors (or nonnegative homogeneus polynomials of degree $2d$ in n variables).

Proof. $Tr(x^{\otimes 2d}) = \|x\|^2 = 1$, so $g^{-1}(1) \supseteq \phi(X)$. An argument analogous to the previous example shows (ϕ, g) is admissible. Notice that $l \in \mathcal{K}^*$ if an only if l is linear on $\mathcal{S}_n^{2d}$ and nonnegative over the set of elements $x^{\otimes 2d}$. Since a linear function on $\mathcal{S}_n^{2d}$ can be thought as a tensor $A^l \in \mathcal{S}_n^{2d}$, then being nonnegative over the elements $x^{\otimes 2d}$ is equivalent to

$$\left\langle A^l, M_{2d}(x) \right\rangle \geq 0 \quad \forall x \in \mathbb{R}^n \tag{96}$$

So the cone $\mathcal{K}^*$ can be thought as the cone of semidefinite tensors (or the cone of nonnegative polynomials of degree $2d$ in n variables) . □

The next example characterizes the cone of copositive tensors (or copositive polynomials). The proof is analogous.

Example 4. Let X be the intersection of the $n-1$ sphere and the nonnegative orthant, and give X the uniform measure. Choose $\phi : X \to \mathcal{S}_n^{2d}$ to be the map M_{2d} (defined in the preliminaries) that sends x to the symmetric tensor $x^{\otimes 2d}$ and let $g : \mathcal{S}_n^{2d} \to \mathbb{R}$ be the trace of the tensor. Then (ϕ, g) is admissible and $\mathcal{K}^*$ is the cone of copositive tensors (or homogeneus polynomials of degree $2d$ in n variables nonnegative on the nonnegative orthant).

4.2 Approximating cones with spectrahedra

In this section we construct and prove the convergence of a sequence of spectrahedra that can approximate any proper cone.

Definition 11. [BVL Aproximation] Let $\phi : X \to V$ and $g : V \to \mathbb{R}$ be an admissible pair and $\mathcal{K}$ its associated cone. Let $\mathcal{F}$ be a finite vector space of continuous real valued functions on X. For $\lambda \in V^*$ we define the quadratic bilinear form $Q_\lambda : \mathcal{F} \times \mathcal{F} \to \mathbb{R}$

$$Q_\lambda(p, q) = \int_X \lambda(\phi(u))p(u)q(u)d\mu(u) \tag{97}$$

We define the cones $\mathcal{K}^*(\mathcal{F}) = \{\lambda \in V^* : Q_\lambda \text{ is psd}\}$ and $\mathcal{K}(\mathcal{F}) = (\mathcal{K}^*(\mathcal{F}))^*$ and call them the BVL approximations to $\mathcal{K}^*$ and $\mathcal{K}$ respectively.

Notice that if $\lambda \in \mathcal{K}^*$ then $\lambda(\phi(u)) \geq 0$ for any $u \in X$, so $\lambda \in \mathcal{K}^*(\mathcal{F})$. This implies

$$\mathcal{K}^* \subseteq \mathcal{K}^*(\mathcal{F}) \text{ and } \mathcal{K}(\mathcal{F}) \subseteq \mathcal{K} \tag{98}$$

so $\mathcal{K}^*(\mathcal{F})$ approximates $\mathcal{K}^*$ from above and $\mathcal{K}(\mathcal{F})$ approximates $\mathcal{K}$ from below.

Now we show that this approximation is indeed a spetrahedron. To ease the notation, we will denote $[x]_\alpha = [x_1, \ldots, x_m]^T$ as the vector x expressed with respect to the basis α.

Proposition 7 (BVL Explicit Construction)**.** *Choose a basis $\beta = \{\beta_1, \ldots, \beta_d\}$ for $\mathcal{F}$, a basis $\gamma = \{\gamma_1, \ldots, \gamma_n\}$ for V, and give V^* the dual basis $\gamma^* = \{\gamma_1^*, \ldots, \gamma_n^*\}$. Then $\lambda \in \mathcal{K}^*(\mathcal{F})$ if and only if $[\lambda]_{\gamma^*} = [\lambda_1, \ldots, \lambda_n]$ is an element of the spectrahedron*

$$\Lambda^* := \left\{ x \in \mathbb{R}^n : \hat{Q}(x) = \sum_{1 \leq i \leq n} x_i \hat{Q}^{(k)} \succeq 0 \right\} \tag{99}$$

where $\hat{Q}^{(k)}$ is the symmetric matrix

$$[\hat{Q}^{(k)}]_{vw} = \int_X [\phi(u)]_k \beta_v(u)\beta_w(u)d\mu(u) \quad \textit{for } 1 \leq v, w \leq d \tag{100}$$

Similarly, let $L : \mathcal{S}^{n\times n} \to \mathbb{R}^n$ be the adjoint operator of Q, shown in Lemma 1 to be

$$L(Y)_i = \langle Y, Q_i \rangle \quad \textit{for } 1 \leq i \leq n \tag{101}$$

Then $y \in \mathcal{K}(\mathcal{F})$ if and only if $[y]_\gamma = [y_1, \ldots, y_n]$ is an element of the cone

$$\Lambda = \{L(U) : U \in \mathcal{S}^{d\times d}, U \succeq 0\} \tag{102}$$

Proof. Recall that the bilinear form Q_λ is positive semidefinite iff

$$Q_\lambda(p,p) = \int_X \lambda(\phi(u))p^2(u)d\mu(u) \geq 0 \quad \forall p \in \mathcal{F} \tag{103}$$

Define the matrix

$$[\hat{Q}(\lambda)]_{vw} = \int_X \lambda(\phi(u))\beta_v(u)\beta_w(u)d\mu(u) \tag{104}$$

and notice that

$$Q_\lambda(p,p) = [p]_\beta^T \hat{Q}(\lambda)[p]_\beta \tag{105}$$

so Q_λ is psd (i.e. $\lambda \in \mathcal{K}(\mathcal{F})^*$) if an only if $\hat{Q}(\lambda) \succeq 0$. Finally, notice that since $\lambda(\phi(u)) = \lambda_1\phi(u)_1 + \cdots + \lambda_n\phi(u)_n$, then (104) and (100) show that we can write $\hat{Q}(\lambda)$ as

$$[\hat{Q}(\lambda)]_{vw} = \sum_{1\leq k\leq n} \lambda_k [\hat{Q}^{(k)}]_{vw} \tag{106}$$

so $\lambda \in \mathcal{K}(\mathcal{F})^*$ if and only if

$$\hat{Q}(\lambda) = \sum_{1\leq k\leq n} \lambda_k \hat{Q}^{(k)} \succeq 0 \tag{107}$$

To show the second part of the proposition, simply use Proposition 4 to calculate the dual of the spectrahedron Λ^* □

4.3 Convergence of the BVL approximation

For a family of nested vector spaces $\mathcal{F}_1 \subset \cdots \subset \mathcal{F}_n$ note that

$$\mathcal{K}^* \subset \mathcal{K}^*(\mathcal{F}_n) \subset \cdots \subset \mathcal{K}^*(\mathcal{F}_1) \tag{108}$$

$$\mathcal{K} \supset \mathcal{K}(\mathcal{F}_n) \supset \cdots \supset \mathcal{K}(\mathcal{F}_1) \tag{109}$$

An important question is under which conditions on the sequence $\mathcal{F}_i$ we get convergence to $\mathcal{K}$ and $\mathcal{K}^*$. Barvinok [1] shows that there is convergence if $\mathcal{F}_n$ is the set of polynomials of degree less or equal to n. The following theorem by Romero and Velasco [3] generalizes this result

Theorem 13. *Assume X is a compact Hausdorff topological space and $(\mathcal{F}_i)_{i\in\mathcal{N}}$ is an increasing sequence of vector subspaces on the algebra of continuous functions on X. If $\bigcup \mathcal{F}_j$ is a subalgebra which separates points and contains the constant functions, then*

$$\bigcap_{i\in\mathbb{N}} \mathcal{K}^*(\mathcal{F}_i) = \mathcal{K}^* \quad \textit{and} \quad \overline{\bigcup_{i\in\mathbb{N}} \mathcal{K}(\mathcal{F}_i)} = \mathcal{K} \tag{110}$$

Proof. The second equality follows from the first if we take duals at both sides. To prove the first equality, notice that $\mathcal{K}^* \subseteq \mathcal{K}(\mathcal{F}_n)$ implies $\mathcal{K}^* \subseteq \bigcap \mathcal{K}(\mathcal{F}_n)$. If $\lambda \notin \mathcal{K}^*$, there exists a vector $x \in \mathcal{K}$ where $\lambda(x) < 0$, so there must exist $u \in X$ such that $\lambda(\phi(u)) = \epsilon < 0$. We can take a non-empty open set U where $\lambda \circ \phi$ is negative, namely

$$U = \{x \in X : \lambda(\phi(x)) < \epsilon/2\} \tag{111}$$

Since X is a normal topological space, there exists a continuous function $p : X \to [0,1]$ and a non-empty set $A \subseteq \bar{A} \subseteq U$ such that $p(\bar{A}) = 1$ and $p(X\backslash U) = 0$. It follows that

$$\int_X \lambda(\phi(u))p(u)^2 d\mu =: \beta \leq \epsilon\mu(A) < 0 \tag{112}$$

Using the hypotheses, the Stone-Weierstrass Theorem implies that the algebra $\bigcup \mathcal{F}_j$ is dense in the algebra of continuous functions on X with the uniform norm, so we can approximate p through a sequence of functions $p_k \in \mathcal{F}_{n_k}$ that converges to p uniformly. We obtain

$$\lim_{k\to\infty} \int_X \lambda(\phi(u))p_k(u)^2 d\mu = \beta < 0 \tag{113}$$

so for some k this integral is negative and $\lambda \notin \mathcal{K}^*(\mathcal{F}_{n_k})$. This shows the equality. □

Notice that if X is finite, then any function over X is a polynomial of degree at most N, for N large enough. In such case the cone $\mathcal{K}^*$ can be approximated in finite steps, i.e., $\mathcal{K}^* = \mathcal{K}^*(\mathcal{F}_N)$

4.4 Approximating $\mathcal{C}$ and $\mathcal{C}^*$

As Proposition 7 shows, calculating the BVL approximation $\mathcal{K}(\mathcal{F})$ lies in being able to calculate the integrals of Equation (100). This is not always possible, but in the case of the copositive cone we are lucky. Here the problem reduces to integrating polynomials over the nonnegative sphere, and Theorem 5 shows this can be done in closed form.

We will use the associated pair of Example 4, which has $\mathcal{C}$ as associated cone when $d = 1$. In this case

$$\phi(u) = u \otimes u = uu^T \tag{114}$$

To guarantee the convergence of the approximation we use $\mathcal{F}_d$ equal to the space of polynomials in n variables of degree up to d. We associate to this vector space the basis β of monomials with leading coefficient one. For $V = \mathcal{S}^{d\times d}$ we use the canonical basis E_{ii} and $E_{ij} + E_{ji}$. Define

$$\mathcal{BVL}_d := \mathcal{C}(\mathcal{F}_d) \text{ and } \mathcal{BVL}_d^* := \mathcal{C}^*(\mathcal{F}_d) \tag{115}$$

If we replace this information into Equation (100), then the matrices $\hat{Q}^{(ij)}$ must be filled up with the integrals

$$[\hat{Q}^{(ij)}]_{vw} = \int_X u_i u_j \beta_v(u)\beta_w(u) d\mu(u) \quad \text{for } 1 \le v, w \le D \text{ and } 1 \le i, j \le n \tag{116}$$

where $D = \dim(\mathcal{F}_d)$ and X is the intersection of the S^{n-1} sphere and the nonnegative orthant.

After the calculation of these matrices using Folland's Theorem 5 we can approximate any copositive linear program (COP) with a semidefinite program (COP_d), simply replacing the condition $X \in \mathcal{C}^*$ with the condition $X \in \mathcal{C}^*(\mathcal{F}_d)$. After using Proposition 7 and the definition of Λ^*, the copositive program and its relaxation are given by

$$\begin{aligned} c := \min_X \langle D, X\rangle \qquad &\text{(COP)} \qquad & c^d_{\mathcal{BVL}} := \min_{X,P} \langle D, X\rangle \qquad &(\text{COP}_d) \\ s.t\ \langle A_i, X\rangle = b_i \ \ 1 \le i \le S & & s.t\ \langle A_i, X\rangle = b_i \ \ 1 \le i \le S & \\ X \in \mathcal{C}^*_n & & \sum_{1\le i,j\le n} X_{ij}\hat{Q}^{ij} = P \succeq 0 & \end{aligned} \tag{117}$$

Notice that $c^d_{\mathcal{BVL}}$ will be a lower bound on c. This approximation will have $D(D+1)/2 + S$ linear restrictions and $n(n+1)/2 + D(D+1)/2$ unknowns, were

$$D = dim(\mathcal{F}_n) = \binom{n+d}{d} \tag{118}$$

which grows polinomially on d (but still very fast).

For any completely positive program (POS) we can obtain an even simpler relaxation (POS^d) According to Proposition 7, $y \in \Lambda$ iff $y = L(Y)$ for some $Y \succeq 0$. This means that for any matrix P

$$\langle y, P\rangle = \langle L(Y), P\rangle = \langle Y, Q(P)\rangle \tag{119}$$

where

$$[Q(P)]_{vw} = \sum_{1\leq i,j\leq n} P_{ij}Q^{(ij)} \tag{120}$$

So a completely positive program (POS) and the relaxation obtained by replacing the condition $X \in \mathcal{C}$ with the condition $X \in \mathcal{C}(\mathcal{F}_d)$ will be given by

$$\begin{array}{lll} p := \max\limits_{X} \langle D, X\rangle & \text{(POS)} & p^d_{\mathcal{BVL}} := \max\limits_{Y} \langle Q(D), Y\rangle \qquad \text{(POS}_d\text{)} \\ s.t\ \ \langle A_i, X\rangle = b_i \ \ 1 \leq i \leq S & & s.t\ \ \langle Q(A_i), Y\rangle = b_i \ \ 1 \leq i \leq S \\ \quad X \in \mathcal{C}_n & & \quad Y \succeq 0 \end{array} \tag{121}$$

Notice that p_d is an upper bound on p. The BVL semidefinite relaxation has S equalities and $D(D+1)/$ unknowns. Because of its simplicity, we will use (POS$_d$) instead of (COP$_d$) in numerical experiments.

5 The SOS Aproximation

In this section we will present an overview of the hierarchy proposed by Parrilo in [5] to approximate the copositive cone $\mathcal{C}^*$. Remember $\mathbb{R}[x]^k_n$ is the set of polynomials in n variables with degree at most k and denote P^k_n the set of such polynomials that are globally non-negative.

Definition 12. We say that $p \in \mathbb{R}[x]^{2d}_n$ is a sum of squares if there exist polynomials $f_1, \ldots, f_m \in \mathbb{R}[x]^d_n$ such that

$$p(x) = \sum_i f_i(x)^2 \tag{122}$$

and denote Σ^{2d}_n the set of such polynomials

It is easy to show that requiring f_i to have degree at most d does not change Σ^{2d}_n.

Notice that a sum of squares polynomial must be non-negative, which means $\Sigma^{2d}_n \subseteq P^{2d}_n$. Hilbert's celebrated result from 1888 [31] shows exactly when equality holds

Theorem 14 (Hilbert)**.** *$\Sigma^{2d}_n \subseteq P^{2d}_n$ holds only in the cases*

(i) $n = 1$, i.e, the univariate case

(ii) $2d = 2$, i.e, quadratic polynomials

(iii) $n = 2$, $2d = 4$ i.e, quartic polynomials in two variables

For all the other cases there exist non-negative polynomials that are not sum of squares.

Determining whether p is globally nonnegative is known to be an NP-Hard problem [32], whereas determining if p is a sum of squares boiles down to determining if certain semidefinite program is feasible. To show this, define $[x]_\beta$ as the vector of monomials in $\mathbb{R}[x]_n^d$ and consider the next result.

Theorem 15. *$p \in \Sigma_n^{2d}$ can be written as a sum of squares if and only if there is a psd matrix $Q \succeq 0$ such that $p(x) = [x]_\beta^T Q[x]_\beta$*

Proof. Assume $p \in \Sigma_n^{2d}$ can be written as a sum of squares $p(x) = \sum_i f_i(x)^2$, then $f_i(x) = V_i^T[x]_\beta$ for some vector V_i and we can write

$$p(x) = \sum_i f_i(x)^2 = [x]_\beta^T V^T V[x]_\beta = [x]_\beta^T Q[x]_\beta \tag{123}$$

where V is the matrix with rows V_i and $Q = V^T V$ is a psd matrix. For the other direction assume

$$p(x) = [x]_\beta^T Q[x]_\beta \text{ for } Q \succeq 0 \tag{124}$$

We can factor $Q = V^T V$ and then define $f_i(x) = (V[x]_\beta)_i$, obtaining

$$p(x) = [x]_\beta^T Q[x]_\beta = \sum_i f_i(x)^2 \tag{125}$$

□

Notice that the equality $p(x) = [x]_\beta^T Q[x]_\beta$ is a series of linear equations on the entries of Q, which we summarize in the next result.

Corollary 1. *Membership of Σ_n^{2d} can be determined through semidefinite programming. Namely, the polynomial $p(x) = \sum_\alpha p_\alpha x^\alpha$ is a sum of squares if and only if there exists a symmetric matrix Q of size $\binom{n+d}{d}$ that satisfies*

$$p_\alpha = \sum_{\beta+\gamma\leq\alpha} Q_{\beta\gamma} \text{ with } Q \succeq 0 \tag{126}$$

5.1 Verifying Copositivity

To verify that $M \in \mathcal{C}$ we must show that $x^T M x \geq 0$ whenever $x \geq 0$. Using the substitution $x_i = z_i^2$, this is equivalent to showing the global nonnegativity of the quartic polynomial

$$P_M(z) = \sum_{1\leq i,j\leq n} z_i^2 M_{ij} z_j^2 \tag{127}$$

Surely if $P_M \in$ is SOS then M will be copositive, but because $\Sigma_n^{2d} \neq P_n^{2d}$ this is not a sufficient condition. The following result by Reznick [7] motivates more versatile sufficient conditions.

Theorem 16 (Reznick)**.** *Let P by a homogenous polynomial that is positive on the nonnegative orthant $\mathbb{R}_+^n$. Then there exists an r such that*

$$P^r := \left(\sum_{1\leq i\leq n} z_i^2\right)^r \sum_{1\leq i,j\leq n} P \tag{128}$$

is a sum of squares

This motivates the following hierarchy of cones

Definition 13 (SOS Hierarchy)**.** For a symmetric matrix $M \in \mathcal{S}^{n\times n}$, denote the polynomial

$$P_M^r(z) = \left(\sum_{1\le i\le n} z_i^2\right)^r \sum_{1\le i,j\le n} z_i^2 M_{ij} z_j^2 \tag{129}$$

The r-th SOS approximation to the copositive cone $\mathcal{C}$ is defined as the cone

$$\mathcal{SOS}_r = \{M : P_M^r \text{ is a sum of squares}\} \tag{130}$$

This hierarchy satisfies

$$\mathcal{SOS}_0 \subseteq \mathcal{SOS}_1 \subseteq \mathcal{SOS}_2 \subseteq \cdots \subseteq \mathcal{C} \tag{131}$$

and Theorem 16 guarantees that $\mathcal{K}_n \to \mathcal{C}^*$ pointwise. Unfortunately, for the copositive cone to be reached in finite steps of the hierarchy we would need $P = NP$.

To gain some intuition from this hierarchy, remember we defined in Section 3.3 the dual of the cone of doubly nonnegative matrices

$$\mathcal{DN}_n^* := \mathcal{S}_+^n + \mathcal{N}_n \tag{132}$$

and showed that $\mathcal{C}_n^* \subseteq \mathcal{DN}_n^*$, with equality when $n \le 4$. The next theorem proved by Parrilo shows that $\mathcal{SOS}_0$ corresponds with $\mathcal{DN}^*$

Theorem 17. *The $\mathcal{SOS}_0$ approximation of the copositive cone is equal to $\mathcal{DN}_n^* := \mathcal{S}_+^n + \mathcal{N}_n$*

Proof. Assume $M \in \mathcal{DN}^*$, so that $M = P + N$ for $P \succeq 0$ and $N \ge 0$, so that

$$P_M^0 = \sum_{1\le i,j\le n} z_i^2 M_{ij} z_j^2 = \sum_{1\le i,j\le n} z_i^2 P_{ij} z_j^2 + \sum_{1\le i,j\le n} z_i^2 N_{ij} z_j^2 \tag{133}$$

The first polynomial can be written as a sum of squares because of Theorem 15 and the second is a sum of squares because $N_{ij} \ge 0$, implying $M \in \mathcal{SOS}_0$. Assume now that $M \in \mathcal{SOS}_0$, meaning the polynomial

$$P(z) = \sum_{i,j} z_i^2 M_{ij} z_j^2 \tag{134}$$

can be written as a sum of squares. From Theorem 15 there must exist a matrix $Q \succeq 0$ such that

$$P(z) = [z]_2^T Q [z]_2 \tag{135}$$

where $[z]_2$ is the vector of monomials in n variables up to degree 2. Since P is homogeneous of degree 4, the entries in Q that correspond to monomials of degree 1 or 0 cancel each other, so we may assume $[z]_2 = \{z_1^2, \ldots, z_n^2, z_1 z_2, \ldots, z_{n-1} z_n\}$.

$P(z)$ contains only monomials of the form $z_i^2 z_j^2$, so only the terms in the product $[z]_2^T Q [z]_2$ obtained from the multiplications $(z_i z_j)(z_i z_j)$ and $(z_i^2)(z_j^2)$ will not cancel each other. This means we may put the entries of Q not corresponding to these products equal to 0 to get

$$Q = \begin{bmatrix} Q^*_{11} & Q^*_{12} & \dots & Q^*_{1n} & 0 & 0 & 0 & 0 \\ Q^*_{21} & Q^*_{22} & \dots & Q^*_{2n} & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ Q^*_{n1} & Q^*_{n2} & \dots & Q^*_{nn} & 0 & 0 & \vdots & 0 \\ 0 & 0 & \dots & 0 & \lambda_{12} & 0 & \vdots & 0 \\ 0 & 0 & \dots & 0 & 0 & \lambda_{13} & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \vdots & \lambda_{n-1,n} \end{bmatrix} \tag{136}$$

For the equality $\sum_{i,j} z_i^2 M_{ij} z_j^2 = P(z) = [z]_2^T Q [z]_2$ to hold, it must be that $Q^*_{ij} + Q^*_{ji} + \lambda_{ij} = 2M_{ij}$, so the submatrix Q^* needs to have the structure

$$Q^* = \begin{bmatrix} M_{11} & M_{12} - \lambda_{12} & \dots & M_{1n} - \lambda_{1n} \\ M_{21} - \lambda_{21} & M_{22} & \dots & M_{2n} - \lambda_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n1} - \lambda_{n1} & M_{n2} - \lambda_{n2} & \dots & M_{nn} \end{bmatrix} \tag{137}$$

Since $Q \succeq 0$, Equation (136) indicates that $Q^* \succeq 0$ and $\lambda_{ij} \geq 0$. If we define the nonnegative symmetric matrix N as $N_{ij} = \lambda_{ij}$, Equation (137) shows that $M - N \succeq 0$. We conclude that $M = (M - N) + P \in \mathcal{S}_+ + \mathcal{N}$ and end the proof. □

5.2 Solving copositive programs

Just as we did in Section 4.4 for the BVL approximation, we may approximate any copositive linear program (COP)

$$c := \min_Y \langle D, Y \rangle \tag{138}$$

$$s.t \;\; \langle A_i, Y \rangle = b_i \;\; 1 \leq i \leq S \tag{COP}$$

$$Y \in \mathcal{C}_n^* \tag{139}$$

through a semidefinite program (COP^r) that uses the $\mathcal{SOS}$ hierarchy

$$\begin{aligned} c^r_{\mathcal{SOS}} &:= \min_{Y,Q} \langle D, Y \rangle \\ s.t \;\; \langle A_i, Y \rangle &= b_i \;\; 1 \leq i \leq S \\ p_\alpha &= \sum_{\beta+\gamma \leq \alpha} Q_{\beta\gamma} \\ \sum_\alpha p_\alpha x^\alpha &= \left(\sum z_i^2\right)^r \sum z_i^2 Y_{ij} z_j^2 \\ Q &\succeq 0 \end{aligned} \qquad (\text{COP}^r) \tag{140}$$

This is a semidefinite problem with $S + \frac{D(D+1)}{2}$ equalities and $\frac{n(n+1)}{2} + \frac{D(D+1)}{2}$ unknowns (notice the similarity with the BVL approximation). In this case $c^r_{\mathcal{SOS}}$ will be an upper bound on c.

6 Quadratic Optimization

In this section we show the equivalence between three formulations of quadratic optimization and copositive programming. Later we will emphasize its usefulness in solving difficult problems of data analysis and graph theory.

6.1 Continuous Quadratic Optimization

Consider the problem (QUAD) of minimizing a quadratic function on the non-negative orthant and its copositive relaxation ($\text{QUAD}_{\mathcal{C}}$)

$$\begin{aligned} &\min_x \ x^T Q x \\ &s.t. \ x^T A x = b \quad \text{(QUAD)} \\ &\quad x \geq 0 \end{aligned} \qquad \begin{aligned} &\min_X \ \langle Q, X \rangle \\ &s.t. \ \langle X, A \rangle = b \quad (\text{QUAD}_{\mathcal{C}}) \\ &\quad X \in \mathcal{C} \end{aligned} \tag{141}$$

where the matrix A is strictly copositive. Though simple this problem may appear, it finds enormous applications. Notice that if we take $A = aa^T$ for $a > 0$ then $x^T A x = 1$ is equivalent to $a^T x = 1$, so a special case of (QUAD) is the standard quadratic problem

$$\begin{aligned} &\min_x \ x^T Q x \\ &s.t. \ a^T x = 1 \quad \text{(STAND)} \\ &\quad x \geq 0 \end{aligned} \tag{142}$$

As we will show, (STAND) can encode the NP-Hard problem of finding the stability number of a graph, which gives an idea of the difficulty in solving (QUAD).

To see that ($\text{QUAD}_{\mathcal{C}}$) is a relaxation in the sense of Definition 4 define the equivalence function $F(x) := xx^T$. Notice that both $F(x) \in \mathcal{C}$ and $x^T A x = \langle F(x), A \rangle$ imply $F(x)$ is feasible for ($\text{QUAD}_{\mathcal{C}}$) whenever x is feasible for (QUAD). Moreover $x^T Q x = \langle F(x), Q \rangle$ implies x and $F(x)$ achieve the same value. This means ($\text{QUAD}_{\mathcal{C}}$) is a relaxation and we conclude

$$v_{\text{QUAD}} \geq v_{\text{QUAD}_{\mathcal{C}}} \tag{143}$$

These problems are equivalent in the sense of Definition 5. To show it first we need a lemma that characterizes the extreme points of the feasible set $\mathcal{F}_{\text{QUAD}_{\mathcal{C}}}$

Lemma 7. *The extreme points of the feasible set $\mathcal{F}_{QUAD_{\mathcal{C}}} = \{X \in \mathcal{C} : \langle X, A \rangle = b\}$ are the matrices $X \in \mathcal{F}_{QUAD_{\mathcal{C}}}$ of the form $X = xx^T$*

Proof. Assume $X \in \mathcal{F}_{\text{QUAD}_{\mathcal{C}}}$ is of the form $X = xx^T$. To show that it is extreme take $X_1, X_2 \in \mathcal{F}_{\text{QUAD}_{\mathcal{C}}}$ feasible matrices such that $X = \lambda_1 X_1 + \lambda_2 X_2$. Theorem 9 shows that X is an extreme vector of the cone $\mathcal{C}$, so $X = \mu_1 X_1 = \mu_2 X_2$ for $\mu_1, \mu_2 \geq 0$. Since

$$b = \langle X, A \rangle = \mu_i \langle X_i, A \rangle = \mu_i b \ \Rightarrow \ \mu_1 = \mu_2 = 1 \tag{144}$$

then $X = X_1 = X_2$ and we conclude that X is an extreme point of $\mathcal{F}_{\text{QUAD}_{\mathcal{C}}}$. To prove the other direction assume X is an extreme point of $\mathcal{F}_{\text{QUAD}_{\mathcal{C}}}$ and decompose it as

$$X = x_1 x_1^T + \cdots + x_n x_n^T \tag{145}$$

Now define

$$u_i = \sqrt{\frac{b}{x_i^T A x_i} x_i} \tag{146}$$

and notice that $u_i u_i^T \in \mathcal{F}_{\text{QUAD}_\mathcal{C}}$. It is easy to check that

$$X = \sum_{i=1}^{k} \frac{x_i^T A x_i}{b} \, u_i u_i^T \tag{147}$$

Moreover, this is a convex combination, since

$$\sum_{i=1}^{k} \frac{x_i^T A x_i}{b} = \frac{\langle A, X \rangle}{b} = 1 \text{ and } \frac{x_i^T A x_i}{b} \geq 0 \tag{148}$$

We showed that X can be written as a convex combination of other feasible matrices, but since X is extreme, the only way this can happen is if $X = u_i u_i^T$, showing that X is of the form xx^T. □

Theorem 18. *The problems QUAD and QUAD$_\mathcal{C}$ are equivalent in the sense of Definition 5, that is, QUAD$_\mathcal{C}$ is a relaxation of QUAD and*

(i) $v_{QUAD} = v_{QUAD_\mathcal{C}}$

(ii) $\mathcal{O}_{QUAD} \neq \emptyset$ *if and only if* $\mathcal{O}_{QUAD_\mathcal{C}} \neq \emptyset$

(iii) $\mathcal{O}_{QUAD_\mathcal{C}} = convex(\{F(x) : x \in \mathcal{O}_{QUAD})$

Proof. We already showed that QUAD$_\mathcal{C}$ is a relaxation of QUAD, which gave the inequality

$$v_{\text{QUAD}} \geq v_{\text{QUAD}_\mathcal{C}} \tag{149}$$

If $v_{\text{QUAD}} = -\infty$ then the equality is automatic, so we will assume that QUAD is bounded and we will divide the proof in the cases $b < 0$, $b = 0$ and $b > 0$. In the case that $b < 0$, then the fact that A is copositive implies $\mathcal{F}_{\text{QUAD}} = \emptyset$ and since any $X \in \mathcal{C}^*$ can be written as a sum

$$X = x_1 x_1^T + \cdots + x_n x_n^T \quad x_i \geq 0 \tag{150}$$

then $\langle X, A \rangle \geq 0$ and we also have $\mathcal{F}_{\text{QUAD}_\mathcal{C}} = \emptyset$.

In the case that $b = 0$, then $\mathcal{F}_{\text{QUAD}_\mathcal{C}}$ is a cone. If $\langle Q, X \rangle < 0$ for some $X \in \mathcal{F}_{\text{QUAD}_\mathcal{C}}$, then $v_{\text{QUAD}} = -\infty$ and we automatically get the equality $v_{\text{QUAD}} = v_{\text{QUAD}_\mathcal{C}}$ from (149). If $\langle Q, X \rangle \geq 0$ for all feasible X then

$$v_{\text{QUAD}_\mathcal{C}} = \langle Q, 0 \rangle = 0 \tag{151}$$

since $X = 0$ is feasible. Since $x = 0$ is also feasible for $(QUAD)$ and achieves a value of 0, then $v_{\text{QUAD}} \leq v_{\text{QUAD}_\mathcal{C}}$ and (149) implies equality.

In the case that $b > 0$, notice that the feasible set $\mathcal{F}_{\text{QUAD}_\mathcal{C}}$ is convex and closed, so the only way it can be unbounded is if it contains a line. This cannot be, because in such a case there would be a feasible matrix X and matrix H such that for all positive λ then $X + \lambda H \in \mathcal{C}$ and $\langle X + \lambda H, A \rangle = b$. These two conditions would imply that $H \in \mathcal{C}$ and $\langle H, A \rangle = 0$, contradicting the strict copositivity

of A. We conclude that $\mathcal{F}_{\text{QUAD}_\mathcal{C}}$ is convex and compact.

Now choose a feasible $X \in \mathcal{F}_{\text{QUAD}_\mathcal{C}}$. Because $\mathcal{F}_{\text{QUAD}_\mathcal{C}}$ is convex and compact we can use the Krein-Milmal Theorem II to show that X is a convex combination of the extreme points of $X \in \mathcal{F}_{\text{QUAD}_\mathcal{C}}$. This means that at least one extreme point Y achieves a value $\langle Q, Y\rangle \leq \langle Q, X\rangle$, and since Y is extreme, Lemma 7 implies that $Y = yy^T$ for some $y \in \mathcal{F}_{\text{QUAD}}$. Since for all $X \in \mathcal{F}_{\text{QUAD}_\mathcal{C}}$ we could find $y \in \mathcal{F}_{\text{QUAD}}$ with the same value, then $v_{\text{QUAD}} \leq v_{\text{QUAD}_\mathcal{C}}$. Along with (149) this shows (i).

To show (ii) and (iii) notice that if $x \in \mathcal{O}_{\text{QUAD}}$ then $F(x) = xx^T$ achieves the same value, and because of (i) then $F(x) = xx^T \in \mathcal{O}_{\text{QUAD}_\mathcal{C}}$. On the other hand, if $X \in \mathcal{O}_{\text{QUAD}_\mathcal{C}}$, then the Krein-Milman Theorem II and Lemma 7 imply that X can be written as a convex combination

$$X = \lambda_1 x_1 x_1^T + \cdots + \lambda_n x_n x_n^T \tag{152}$$

Since $x_i x_i^T \in \mathcal{F}_{\text{QUAD}_\mathcal{C}}$ and X achieves the minimum value of a linear function on $\mathcal{F}_{\text{QUAD}_\mathcal{C}}$, then all of the matrices $x_i x_i^T$ should achieve the minimum value, implying $x_i x_i^T \in \mathcal{O}_{\text{QUAD}_\mathcal{C}}$ and $x_i \in \mathcal{O}_{\text{QUAD}}$ □

Using Equation 10, notice that the dual to ($\text{QUAD}_\mathcal{C}$) must be

$$\max_y \ by \tag{153}$$

$$s.t. \ \ Q - yA \in \mathcal{C}^* \tag{$\text{QUAD}^*_\mathcal{C}$}$$

which is a copositive optimization problem. The next result shows there is strong duality

Theorem 19. *There is strong duality for the completely positive relaxation $QUAD_\mathcal{C}$, that is*

$$v_{QUAD_\mathcal{C}} = v_{QUAD^*_\mathcal{C}} \tag{154}$$

Proof. Notice that since the sphere S^{n-1} is compact, then there are α and β such that

$$\alpha = \min_{x \in S^{n-1}} x^T A x \tag{155}$$

$$\beta = \min_{x \in S^{n-1}} x^T Q x \tag{156}$$

Since A is strictly copositive, then $\alpha > 0$. Notice that if $y < 0$

$$x^T(Q - yA)x = x^T Q x - y x^T A x \geq \beta - y\alpha \tag{157}$$

which is positive for $-y$ large enough. So for y negative enough $Q - yA$ is strictly copositive. This implies strong duality from Slater's conditions of Theorem 4 □

6.2 Fractional Programming

Consider the fractional programming problem and the standard quadratic problem defined before

$$\begin{array}{ll} \max\limits_x \ \dfrac{x^T Q x}{x^T A x} & \\ s.t. \ e^T x = 1 & \text{(FRAC)} \\ x \geq 0 & \end{array} \qquad \begin{array}{ll} \min\limits_x \ x^T Q x & \\ s.t. \ e^T x = 1 & \text{(STAND)} \\ x \geq 0 & \end{array} \tag{158}$$

In this section we will argue, following [17], that solving both problems is equivalent. This result will be very important in formulating the independence number of a graph as a copositive program.

Theorem 20. *If A is strictly positive, $(STAND)$ and $(FRAC)$ are equivalent, in the sense that $v_{STAND} = v_{FRAC}$ and we can construct feasible and optimal vectors of one problem with feasible and optimal vectors of the other.*

Proof. Assume $x \in \mathcal{F}_{\text{STAND}}$. If we normalize x so that $y = \frac{x}{e^T x}$ then $e^T y = 1$ and

$$\frac{y^T Q y}{y^T A y} = \frac{x^T Q x}{x^T A x} = x^T Q x \tag{159}$$

So using $x \in \mathcal{F}_{\text{STAND}}$ we could construct $y \in \mathcal{F}_{\text{FRAC}}$ that achieves the same value in the objective function. Similarly, take $y \in \mathcal{F}_{\text{FRAC}}$ and scale it as

$$x = \frac{y}{\sqrt{y^T A y}} \tag{160}$$

It is easy to see that $x^T A x = 1$ and that

$$x^T Q x = \frac{y^T Q y}{y^T A y} \tag{161}$$

So using $y \in \mathcal{F}_{\text{FRAC}}$ we could construct $x \in \mathcal{F}_{\text{STAND}}$ that achieves the same value in the objective function. These results imply that

$$v_{\text{STAND}} = v_{\text{FRAC}} \tag{162}$$

□

6.3 Mixed Binary Quadratic Optimization

In this section we concern ourselves with copositive relaxations to the much more general quadratic binary problem (QBIN). The results of this section were first given by Burer [18]

$$\min_x \quad x^T Q x + 2c^T x \tag{163}$$

$$s.t. \quad (a^i)^T x = b_i \quad i = 1, \dots, m \tag{QBIN}$$

$$x \geq 0 \tag{164}$$

$$x_i \in \{0, 1\} \quad \text{for } i \in B \tag{165}$$

Define the set

$$L = \{x \in \mathbb{R}^n_+ : (a^i)^T x = b_i \quad i = 1, \dots, m\} \tag{166}$$

We will assume that $x \in L$ implies $0 \leq x_i \leq 1$ for $i \in B$. Notice this is not a restrictive condition, since we could always add for $i \in B$ the auxiliary variables $y_i \geq 0$ and the equalities

$$x_i + y_i = 1, \text{ which would imply } x_i \leq 1 \tag{167}$$

To construct a completely positive relaxation consider the problem

$$\min_X \quad \langle Q, X \rangle + 2c^T x \tag{168}$$

$$s.t. \quad (a^i)^T x = b_i \quad i = 1, \dots, m \tag{169}$$

$$(a^i)^T X a^i = b_i^2 \quad i = 1, \dots, m \tag{QBIN$_{\mathcal{C}^*}$}$$

$$x_i = X_{ii} \quad \text{for } i \in B \tag{170}$$

$$\begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \in \mathcal{C}^* \tag{171}$$

Proposition 8. *The problem QBIN$_{\mathcal{C}}$ is a relaxation of QBIN with the equivalence function*

$$F(x) = \begin{bmatrix} 1 & x^T \\ x & xx^T \end{bmatrix} \tag{172}$$

Proof. Assume x is feasible for QBIN. $F(x) \in \mathcal{C}^*$ since $F(x) = yy^T$ for $y = [1, x^T]^T$. All the conditions for the feasibility of $F(x)$ are easy to verify except perhaps for $x_i = X_{ii}$. Notice that this is equivalent to saying $x_i = x_i^2$, which holds because $x_i \in \{0, 1\}$. □

The next conditions, which are easy to verify, characterize the horizon cone of the feasible set.

Lemma 8. *The horizon cone of the feasible set $\mathcal{F}_{QBIN}$ is the cone*

$$\mathcal{F}^H_{QBIN} = \{z \in \mathbb{R}^n : \quad (a^i)^T z = 0 \quad i = 1, \ldots, m \tag{173}$$

$$z_i = 0 \qquad i \in B \tag{174}$$

$$z \geq 0 \qquad \} \tag{175}$$

The equivalence between QBIN and QBIN$_{\mathcal{C}}$ will follow easily from the next lemma

Lemma 9. *Let $M = \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix}$ be feasible for QBIN$_{\mathcal{C}}$. Then M admits a decomposition*

$$M = \sum_{k \in K_1} \lambda_k \begin{pmatrix} 1 \\ v^k \end{pmatrix} \begin{pmatrix} 1 \\ v^k \end{pmatrix}^T + \sum_{k \in K_0} \begin{pmatrix} 0 \\ z^k \end{pmatrix} \begin{pmatrix} 0 \\ z^k \end{pmatrix}^T \tag{176}$$

such that

(i) $\lambda_k \geq 0$ *and* $\sum \lambda_k = 1$

(ii) $z^k \in \mathcal{F}^H_{QBIN}$

(iii) $v^k \in \mathcal{F}_{QBIN}$

Proof. Take $M = \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix}$ feasible for QBIN$_{\mathcal{C}}$. Since M is completely positive, there is a decomposition

$$M = \sum_{k \in K} \begin{pmatrix} c_k \\ z^k \end{pmatrix} \begin{pmatrix} c_k \\ z^k \end{pmatrix}^T \tag{177}$$

where $c_k \in \mathbb{R}_+$ and $z^k \in \mathbb{R}^n_+$. If we split K into $K_1 = \{k \in K : c_k > 0\}$ and $K_0 = \{k \in K : c_k = 0\}$, and then define $v_k := \frac{z_k}{c_k}$, $\lambda_k := c_k^2$ we obtain

$$M = \sum_{k \in K_1} \lambda_k \begin{pmatrix} 1 \\ v^k \end{pmatrix} \begin{pmatrix} 1 \\ v^k \end{pmatrix}^T + \sum_{k \in K_0} \begin{pmatrix} 0 \\ z^k \end{pmatrix} \begin{pmatrix} 0 \\ z^k \end{pmatrix}^T \tag{178}$$

Let us first show the equality

$$b_i c_k = (a^i)^T z^k \tag{179}$$

for $k \in K_0$ and $k \in K_1$. From Equation (178), note that

$$x = \sum_{k \in K} c_k z^k \tag{180}$$

$$X = \sum_{k \in K} z^k (z^k)^T \tag{181}$$

And since $(a^i)^T x = b_i$ and $(a^i)^T X a^i = b_i^2$ we have the equations

$$b_i = (a^i)^T x = \sum_{k \in K} c_k [(a^i)^T z^k] \tag{182}$$

$$b_i^2 = (a^i)^T X a^i = \sum_{k \in K} [(a^i)^T z^k]^2 \tag{183}$$

If we define the vector $w_k^i = (a^i)^T z^k$, these equations are saying that

$$b_i = \langle c, w^i \rangle \tag{184}$$

$$b_i = \left\| w^i \right\| \tag{185}$$

And since $\|c\| = 1$, we obtain $\langle c, w^i \rangle = \|c\| \left\| w^i \right\|$. From Cauchy inequality, we now this can only be true if c and w_i are a multiple of the other. The appropriate constant is b_i, which shows equation (179).

Let us check that all properties are satisfied.

(i) Notice that

$$1 = M_{11} = \sum_{k \in K} c_k^2 = \sum_{k \in K_1} c_k^2 = \sum_{k \in K_1} \lambda_k \tag{186}$$

and $\lambda_k = c_k^2 \geq 0$.

(ii) First, by construction $z^k \geq 0$. Second, whenever $k \in K_0$ then $c_k = 0$, so (179) implies $(a^i)^T z^k = 0$ for $k \in K_0$. Third, notice this also means that whenever $x \in L$ then for all $\mu \geq 0$

$$y = x + \mu z^k \in L \tag{187}$$

But remember, we assumed $y \in L$ implies $0 \leq y \leq 1$ for $j \in B$. So the only way this fact and Equation (187) can hold for any $\mu \geq 0$ is if $z_j^k = 0$.

(iv) Let us check that v^k is feasible for QBIN. First, $v^k \geq 0$ by constrution. Second, (179) and the definition $v^k = z^k / c_k$ show that

$$(a^i)^T v^k = b_i \tag{188}$$

Third, the constrain $x_i = X_{ii}$ means that for $j \in B$

$$\sum_{k \in K_1} \lambda_k v_j^k = \sum_{k \in K_1} \lambda_k (v_j^k)^2 + \sum_{k \in K_0} (z_j^k)^2 \tag{189}$$

$$\Rightarrow \quad \sum_{k \in K_1} \lambda_k v_j^k = \sum_{k \in K_0} \lambda_k (v_j^k)^2 \qquad \text{[using (iii)]} \tag{190}$$

$$\Rightarrow \quad \sum_{k \in K_1} \lambda_k [v_j^k - (v_j^k)^2] = 0 \tag{191}$$

Since $v^k \in L$, then $0 \leq v_j^k \leq 1$, and this means $v_j^k - (v_j^k)^2 \geq 0$. Also $\lambda_k > 0$, so we conclude from the last equation that $v_j^k - (v_j^k)^2 = 0$. This shows that $v_j^k \in \{0, 1\}$.

□

Theorem 21. *$QBIN_{\mathcal{C}}$ is a convex relaxation of QBIN. In fact, they are equivalent in the sense of Definition 5, where the equivalence function that takes a feasible vector into a feasible completely positive matrix is*

$$F(x) = \begin{pmatrix} 1 & x^T \\ x & xx^T \end{pmatrix} \tag{192}$$

And the function G to recover a vector from a matrix is

$$G\begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} = x \tag{193}$$

The next conditions are satisfied

(i) $v_{QBIN_{\mathcal{C}}} = v_{QBIN}$

(ii) $\mathcal{O}_{QBIN} \neq \emptyset$ *iff* $\mathcal{O}_{QBIN_{\mathcal{C}}} \neq \emptyset$

(iii) $conv(\mathcal{O}_{QBIN_{\mathcal{C}}}) = G(\mathcal{O}_{QBIN})$

Proof. From Proposition 8 we already know that $\text{QBIN}_{\mathcal{C}}$ is a relaxation, so

$$v_{\text{QBIN}} \geq v_{\text{QBIN}_{\mathcal{C}}} \tag{194}$$

is automatic, and in the case that $v_{\text{QBIN}} = -\infty$ we already have equality of the optimum values, which would show part (i) of the equivalence. In case problem QBIN is bounded, take a feasible $M = \begin{pmatrix} 1 & x^T \\ x & xx^T \end{pmatrix}$ and write its decomposition as in Lemma 9.

$$M = \sum_{k \in K_1} \lambda_k \begin{pmatrix} 1 \\ v^k \end{pmatrix} \begin{pmatrix} 1 \\ v^k \end{pmatrix}^T + \sum_{j \in K_0} \begin{pmatrix} 0 \\ z^k \end{pmatrix} \begin{pmatrix} 0 \\ z^k \end{pmatrix}^T \tag{195}$$

If we take any $\mu \geq 0$ and any pair $p \in K_0$, $l \in K_1$, then Lemma 9 (ii) shows the vector $y = v^l + \mu z^p$ is feasible. Consider then the value of problem QBIN when we use the feasible vector y

$$y^T Q y + c^T y = (v^l)^T Q v^l + c^T v^l + \lambda[(z^p)^T Q y + c^T z^p] \tag{196}$$

Since QBIN is bounded, it must be that $(z^p)^T Q y + c^T z^p \geq 0$ for all $p \in K_0$, and this means that the matrix

$$\hat{M} = \sum_{k \in K_1} \lambda_k \begin{pmatrix} 1 \\ v^k \end{pmatrix} \begin{pmatrix} 1 \\ v^k \end{pmatrix}^T \tag{197}$$

achieves a smaller or equal optimal value than M. Notice that since each vector v^k is feasible for QBIN then each matrix

$$V^k := \begin{pmatrix} 1 \\ v^k \end{pmatrix} \begin{pmatrix} 1 \\ v^k \end{pmatrix}^T \tag{198}$$

is feasible for $\text{QBIN}_{\mathcal{C}}$, an since part (i) of the lemma shows $\hat{M}$ is a convex combination, then it will also be feasible. If we find the s that solves

$$\min_s \ (v^s)^T Q v^s + c^T v^s \tag{199}$$

then the matrix V^s will be feasible and achieve a value that at most is the value of M. In a nutshell, for any feasible M we are able to come up with a feasible vector v^s that achieves a smaller or equal value. This shows the inequality $v_{\text{QBIN}_{\mathcal{C}}} \geq v_{\text{QBIN}}$, which completes the proof of (i).

To prove (ii) and (iii) notice that if x is optimal for QBIN then (i) implies that $F(x)$ is optimal for $\text{QBIN}_{\mathcal{C}}$. On the other hand, if M is optimal for QBIN then it admits a decomposition as in Equation (195) and the discussion of the previous paragraph, along with the fact that M is optimal, shows that $\hat{M}$ is optimal and each of the vectors v^k should be optimal. Since the first row of M is a convex combination of the vectors $[1, (v^k)^T]$, we obtain (iii). □

7 Completely Positive Tensors

In this section we study the cone of completely positive tensors, which is a natural generalization of $\mathcal{C}^*$ first introduced by Dong [15] that will allow us to formulate completely positive relaxations of polynomial optimization problems. We will follow Peña [16] for most of the discussion, but will emphasise the relation between this relaxation and $\text{QBIN}_{\mathcal{C}}$.

Recall from the preliminaries the operator $M_d : \mathbb{R}^n \to \mathcal{S}_n^d$.

$$M_d(x) = \underbrace{x \otimes \ldots \otimes x}_{d \text{ times}} \tag{200}$$

We define the set of completely positive tensors as the cone

$$\mathcal{C}_{n,d} = \text{conv}(M_d(x) : x \in \mathbb{R}_+^n) \tag{201}$$

Its dual $\mathcal{C}_{n,d}^*$, which we call the set of copositive tensors, is the set of $T \in \mathcal{S}_n^d$ such that $\langle T, M_d(x) \rangle \geq 0$ for all $x \in \mathbb{R}_+^n$. Since any symmetric tensor T can be thought as a homogeneous polynomial p_T defined as

$$p_T(x) = \langle T, M_d(x) \rangle \tag{202}$$

it is easy to see from (201) that $T \in \mathcal{C}_{n,d}^*$ if and only if p_T is nonnegative on $\mathbb{R}_+^n$. In Proposition 10 we show that p_T is indeed a bijection between symmetric tensors and homogeneous polynomials, so we can think of $\mathcal{C}_{n,d}^*$ exactly as the cone of homogeneous polynomials of degree d on n variables that are nonnegative on $\mathbb{R}_+^n$.

Proposition 9. *$\mathcal{C}_{n,d}$ is a proper cone*

Proof.

(i) Closedness and pointedness are immediate from the definition.

(ii) To show that $\mathcal{C}_{n,d}$ is closed, take a sequence $\{X^k\} \subseteq \mathcal{C}_{n,d}$ such that $X^k \to X$. If N is the dimension of $\mathcal{S}_d^n$, Caratheodory's Theorem implies that there exists a decomposition $X^k = \sum_{l=1}^N M_d(u^{(k,l)})$. Let us show that the sequence $U^l = \{u^{(k,l)}\}_{k \in \mathbb{N}} \subset \mathbb{R}_+^n$ is bounded. Consider

the product

$$\left\langle X^k, M_d(e)\right\rangle = \sum_{l=1}^{N} \left\langle M_d(u^{(k,l)}), M_d(e)\right\rangle \tag{203}$$

$$= \sum_{l=1}^{N} \left\langle u^{(k,l)}, e\right\rangle^d \tag{204}$$

Since $\left\langle X^k, M_d(e)\right\rangle$ converges to $\langle X, M_d(e)\rangle$, then $\sum_{l=1}^{N} \left\langle u^{(k,l)}, e\right\rangle^d$ must be bounded, and since all the vectors $u^{(k,l)}$ are non-negative, each of the sequences $U^l = \{u^{(k,l)}\}$ must be bounded. From compacity, we can choose increasing indices n_k such that all the subsequences $u^{(k_n,l)}$ converge, to a limit we call $u^l \in \mathbb{R}_+^n$. It is easy to see that

$$X = \sum_{l=1}^{N} M_d(u^l) \tag{205}$$

showing that $\mathcal{C}_{n,d}$ is indeed closed.

(iii) Showing that $\mathcal{C}_{n,d}$ is non-empty is equivalent to showing that its dual $\mathcal{C}_{n,d}^*$ is pointed. Assume $T, -T \in \mathcal{C}_{n,d}^*$, which means p_T is a homogeneous polynomial of degree d such that $p_T(x) = 0$ when $x \geq 0$. Because p_T is analytic, this can only happen if p_T is the zero polynomial.

□

Proposition 10. *Let $q(x) = \sum_\alpha q_\alpha x^\alpha$ be a homogeneous polynomial in n variables of degree d. Define the symmetric tensor $T \in \mathcal{S}_n^d$*

$$T_{i_1,\ldots,i_d} = \frac{\alpha_1! \ldots \alpha_n!}{|\alpha|} q_\alpha \tag{206}$$

where α_k is the number of times that k occurs in $\{i_1, \ldots, i_d\}$. Then $p_T(x) = q(x)$

Proof. The component T_σ of the tensor contributes to the coefficient of x^α in the polynomial $p_T(x) = \langle T, M_d(x)\rangle$ if and only if

$$\sigma = (i_1, \ldots, i_d) \tag{207}$$

and $k = 1, \ldots, n$ appears exactly α_k times in σ. The number of such σ is the multinomial coefficient

$$\binom{|\alpha|}{\alpha_1, \ldots, \alpha_n} = \frac{|\alpha|}{\alpha_1! \ldots \alpha_n!} \tag{208}$$

and each of such σ will have a value of

$$\frac{\alpha_1! \ldots \alpha_n!}{|\alpha|} q_\alpha \tag{209}$$

Multiplying these two numbers we obtain the coefficient of x^α in the polynomial $p_T(x)$, which ends up being is q_α. This shows the proposition. □

Any polynomial of degree up to d in n variables can be homogenized adding one more variable, and dehomogenized by replacing it with a 1. So we can associate to any polynomial $p \in \mathbb{R}_n^d[x]$, the tensor of its homogeneous polynomial and recover it evaluating the tensor in $M_d([1, x])$. We state this in the following proposition

Proposition 11. *There exists a bijection between polynomials and symmetric tensors $C_d : \mathbb{R}^d_n[x] \to \mathcal{S}_{n+1,d}$, which Proposition 10 shows should be defined as*

$$(C_d(p))_{i_1,\dots,i_d} = \frac{\alpha_0!\alpha_1!\dots\alpha_n!}{|\alpha|} p_\alpha \tag{210}$$

where α_k is the number of times that k occurs in $\{i_1,\dots,i_d\}$ and p_α is the coefficient that accompanies the monomial $x_1^{\alpha_1}\dots x_n^{\alpha_n}$ in p. We also have a way to recover p and $\tilde{p}$ from $C_d(p)$, given by

$$p(x) = \langle C_d(p), M_d(1,x)\rangle \tag{211}$$

$$\tilde{p}(x) = \langle C_d(p), M_d(0,x)\rangle \tag{212}$$

Example 5. Take $p(x_1,x_2) = x_1^3 - x_2^3 + 2x_1x_2 - 3x_1^2 + 3x_2^2 - x_1 + 2$. Then

$$C_{(0,\cdot,\cdot)} = \begin{bmatrix} 2 & -1/3 & 0 \\ -1/3 & -1 & 1/3 \\ 0 & 1/3 & 1 \end{bmatrix} \qquad C_{(1,\cdot,\cdot)} = \begin{bmatrix} -1/3 & -1 & 1/3 \\ -1 & 1 & 0 \\ 1/3 & 0 & 0 \end{bmatrix} \tag{213}$$

$$C_{(2,\cdot,\cdot)} = \begin{bmatrix} 0 & 1/3 & 1 \\ 1/3 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix} \tag{214}$$

For example, to determine $C_{(2,0,1)}$ we look for the monomial $x_2x_0x_1 = x_0^1x_1^1x_2^1$, which has $\alpha = (1,1,1)$ and $p_\alpha = 2$. So

$$C_{(2,0,1)} = 2\frac{1!1!1!}{3!} = \frac{1}{3} \tag{215}$$

7.1 Polynomial optimization

Consider the optimization problem

$$\begin{aligned} &\min_x \ q(x) \\ &s.t.\ h_i(x) = 0 \quad (POP) \\ &z \geq 0 \end{aligned} \tag{216}$$

where q and h_i $(1 \leq i \leq m)$ are polynomials in n variables of degree up to d. The previous discussion motivates the completely positive relaxation (CPP)

$$\begin{aligned} &\min_Y \ \langle C_d(q), Y\rangle \\ &s.t.\ \langle C_d(h_i), Y\rangle = 0 \quad (CPP) \\ &\langle C_d(1), Y\rangle = 1 \\ &Y \in \mathcal{C}^*_{n+1,d} \end{aligned} \tag{217}$$

Notice that if x is a feasible solution of (CPP), then $Y = M_d(x)$ is feasible for (POP) and achieves the same value, so (CPP) is a relaxation of (POP) in the sense of Definition 4.

We would like to prove the equivalence between these problems. Notice that if we define $\mathcal{X}(Y) \in \mathbb{R}^n_+$ as $\mathcal{X}(Y)_i = Y_{0,\dots,0,i}$ for $i = 1,\dots,n$, then

$$\mathcal{X}(M_d(1,x)) = x \tag{218}$$

So to show the equivalence between (CPP) and (POP), in the sense of Definition 5, a good candidate for recovering an optimal vector x^* from an optimal tensor Y^* is the function $\mathcal{X}$.

Before characterizing the conditions of equivalence, we prove two lemmas.

Lemma 10. *If $h \in \mathbb{R}_n^d[x]$ is bounded below in S, then $\tilde{h}(x) \geq 0$ for all $x \in S^\infty$*

Proof. Take $x \in S^\infty$, $x \neq 0$, so for some i we have $x_i \neq 0$. Take a sequence $x^k \in S$ and $\lambda_k \in \mathbb{R}_+$ such that $\lambda_k \to 0$ and $\lambda_k x^k \to x$.

Now decompose h in homogeneous polynomials h^D of degree D as $h = h^d + h^{d+1} + \cdots + h^1 + h^0$, so that

$$\lambda_k^d h(x^k) = h^d(\lambda_k x^k) + \lambda_k^1 h^{d-1}(\lambda_k x^k) + \cdots + \lambda_k^{d-1} h^1(\lambda_k x^k) + \lambda_k^d h^0(\lambda_k x^k) \tag{219}$$

Notice that since h is bounded from below, then

$$\liminf_{k\to\infty} \lambda_k^d h(x^k) \geq 0 \tag{220}$$

If we take this same limit on the right side of (219), all terms die, except for the d term, so we obtain

$$0 \leq \liminf_{k\to\infty} h^d(\lambda_k x^k) = \tilde{h}(x) \tag{221}$$

□

In the context of POP define

$$\mathcal{H}_t := \{x \in \mathbb{R}_+^n : h_i(x) = 0 \text{ for } 1 \leq i \leq t\} \tag{222}$$

$$\tilde{\mathcal{H}} := \{x \in \mathbb{R}_+^n : \tilde{h}_i(x) = 0 \text{ for } 1 \leq i \leq m\} \tag{223}$$

Define $\mathcal{H}_0 = \mathbb{R}_+^n$ as well. Note the analogy between Lemmas 9 and 11, especially considering that $\mathcal{F}_{\text{CPP}}^\infty \subseteq \tilde{\mathcal{H}}$.

Lemma 11. *Let $Y \in \mathcal{C}_{n+1,d}^*$ be feasible for CPP. Then Y admits a representation*

$$Y = \sum_{k\in K_0} \lambda_k M_d(1, v_k) + \sum_{k\in K_1} M_d(0, z_k) \tag{224}$$

for $v_k, z_k \in \mathbb{R}_+^n$ such that

$$\text{(i) } \lambda_k \geq 0 \text{ and } \sum_{k\in K_1} \lambda_k = 1 \tag{225}$$

Additionally

Conditions I

1. *$deg(h_i) = d$ and $h_i(x) \geq 0$ for all $x \in \mathcal{H}_{i-1}$*
2. *$\mathcal{H}_{i-1}^\infty \cap \{x : \tilde{h}_i(x) = 0\} \subseteq \mathcal{H}_i^\infty \quad 1 \leq i \leq m$*

Imply *(ii) $z_k \in \mathcal{F}_{CPP}^\infty$ and (iii) $v_k \in \mathcal{F}_{CPP}$*

Conditions II

1. *$deg(h_i) = d$ and $h_i(x) \geq 0$ for all $x \in R_+^n$*
2. *$\tilde{q}(x) \geq 0$ when $x \in \tilde{\mathcal{H}} \qquad 1 \leq i \leq m$*

Imply *(ii) $z_k \in \tilde{\mathcal{H}}$ and (iii) $v_k \in \mathcal{F}_{CPP}$*

Proof. Since $Y \in \mathcal{C}^*_{n+1,d}$, then it can be written as a sum

$$Y = \sum_{k \in K} M_d(a_k, z_k) \tag{226}$$

where $a_k \in \mathbb{R}^+$ and $z_k \in \mathbb{R}^n_+$. We can split K into $K_1 := \{k \in K : a_k > 0\}$ and $K_0 := \{k \in K : a_k = 0\}$. Then define $v_k = \frac{z_k}{a_k}$ and $\lambda_k = a_k^d$ for $k \in K_1$ to obtain

$$Y = \sum_{k \in K_0} \lambda_k M_d(1, v_k) + \sum_{k \in K_1} M_d(0, z_k) \tag{227}$$

Since $Y \in \mathcal{F}_{CPP}$ we must have

$$1 = \langle C_d(1), Y \rangle = \sum_{k \in K_1} \lambda_k \tag{228}$$

Similarly,

$$0 = \langle C_d(h_i), Y \rangle = \sum_{k \in K_1} \lambda_k h_i(v_k) + \sum_{k \in K_0} \mu_k \tilde{h}_i(z_k) \text{ for } 1 \leq i \leq m \tag{229}$$

Under **Conditions I**, the assumpion 1) implies that h_1 is nonnegative on $\mathcal{H}_0$, and because of Lemma 10, then $\tilde{h}_1$ is nonnegative on $\mathcal{H}_0^\infty$. Notice that since $\mathcal{H}_0 = \mathcal{H}_0^\infty = \mathbb{R}^n_+$, then $v_p \in \mathcal{H}_0$ and $z_q \in \mathcal{H}_0^\infty$ for any pair $(p \in K_1, q \in K_0)$, so that $h_1(v_p) \geq 0$ and $\tilde{h}_1(z_q) \geq 0$.

Equation (229) then implies that $h_1(v_p) = 0$ and $\tilde{h}_1(z_q) = 0$. This means that $v_p \in \mathcal{H}$, and using part 2), $v_q \in \mathcal{H}_0^\infty \cap \{x : \tilde{h}_1(x) = 0\} \subseteq \mathcal{H}_1^\infty$.

We started from $v_p \in \mathcal{H}_0, z_q \in \mathcal{H}_0^\infty$ and showed that $v_p \in \mathcal{H}_1, z_q \in \mathcal{H}_1^\infty$, using 1) and 2). If we repeat this reasoning iteratively, we arrive at the end to

$$v_p \in \mathcal{H}_m = \mathcal{F}_{\text{POP}} \tag{230}$$

$$z_q \in \mathcal{H}_m^\infty = \mathcal{F}_{\text{POP}}^\infty \tag{231}$$

for any pair $(p \in K_1, q \in K_0)$. This shows the first result.

Under **Conditions II**, h_i is nonnegative on $\mathbb{R}^n_+$, so Lemma 10 implies $\tilde{h}_i$ is nonnegative on $(\mathbb{R}^n_+)^\infty = \mathbb{R}^n_+$. For any pair $(p \in K_1, q \in K_0)$ we know that $v_p, z_q \in \mathbb{R}^n_+$, so $h_i(v_p) \geq 0$ and $\tilde{h}_i(z_q) \geq 0$ for $1 \leq i \leq m$. Using (229) this implies that $h_i(v_p) = 0$ and $\tilde{h}_i(z_q) = 0$. So $v_p \in \mathcal{F}_{\text{POP}}$ and $z_q \in \tilde{\mathcal{H}}$.

□

Theorem 22. *Problems POP and CPP are equivalent in the sense of Definition 5 under any set of conditions in Lemma 11. This means that CPP is a relaxation of POP, with equivalence function $F(x) = M_d(x)$, and the following conditions are satisfied*

(i) $v_{POP} = v_{CPP}$

(ii) $\mathcal{O}_{POP} \neq \emptyset$ *iff* $\mathcal{O}_{CPP} \neq \emptyset$

(iii) $Conv(\mathcal{O}_{cpp}) = \mathcal{X}(\mathcal{O}_{POP})$ *where* $\mathcal{X} : \mathcal{S}^d_{n+1} \to \mathbb{R}^n$ *was defined as*

$$[\mathcal{X}(X)]_i := X_{0,\ldots,0,i} \tag{232}$$

Proof. We already showed that CPP is a relaxation of POP so

$$v_{POP} \geq v_{CPP} \tag{233}$$

Whenever $v_{POP} = -\infty$ we have equality and the statement of the theorem is vacuously satisfied. So assume that $q(x)$ is bounded below on $\mathcal{F}_{POP}$. Lemma 10 shows that

$$\tilde{q}(x) \geq 0 \text{ on } \mathcal{F}_{\text{POP}}^{\infty} \tag{234}$$

Now take $Y \in \mathcal{F}_{\text{CPP}}$. Using Lemma 11, Y can be decomposed as

$$Y = \sum_{k \in K_0} \lambda_k M_d(1, v_k) + \sum_{k \in K_1} M_d(0, z_k) \tag{235}$$

for $v_k, z_k \in \mathbb{R}_+^n$ and $\lambda_k > 0$ such that $\sum \lambda_k = 1$

Under any set of conditions, we have that $v_k \in \mathcal{F}_{\text{POP}}$, let us show also that $\tilde{q}(z_k) \geq 0$. Under **Conditions I** we know that $z_k \in \mathcal{F}_{\text{POP}}^{\infty}$, so equation (234) implies $\tilde{q}(z_k) \geq 0$. Under **Conditions II** we know that $z_k \in \tilde{\mathcal{H}}$, so part 2) of the conditions also implies $\tilde{q}(z_k) \geq 0$.

Now notice that

$$\begin{aligned} \langle C_d(q), Y \rangle &= \sum_{k \in K_1} \lambda_k q(v_k) + \sum_{k \in K_0} \tilde{q}(z_k) && (236) \\ &\geq \sum_{k \in K_1} \lambda_k q(v_k) \geq v_{\text{POP}} && (237) \end{aligned}$$

where the last inequality comes from the fact that $\sum \lambda_k q(v_k)$ is a convex combination of values of q on feasible vectors. This shows that $v_{\text{POP}} \leq v_{\text{CPP}}$, which implies (i) in combination with (233).

It is easy to see from equation (236) that if Y^* is optimal for CPP and v_k^* are the vectors corresponding to its decomposition in (235), we must have

$$\langle C_d(q), Y^* \rangle = q(v_k^*) \quad 1 \in K_1 \tag{238}$$

Similarly, if x^* is optimal for POP, then $M_d(x^*)$ will be optimal for CPP, which shows part (ii) of the definition of equivalence.

To prove part (iii), applying $\mathcal{X}$ to equation (235) we obtain

$$\begin{aligned} \mathcal{X}(Y^*) &= \sum_{k \in K_1} \lambda_k \mathcal{X}(M_d(1, v_k^*)) + \sum_{k \in K_0} \mathcal{X}(M_d(0, z_j^*) && (239) \\ &= \sum_{k \in K_1} \lambda_k v_k^* && (240) \end{aligned}$$

And since v_k^* is optimal, we conclude that $\mathcal{X}(Y^*)$ is a convex combination of optimal vectors. □

To finish this section we give a simple condition under which (CPP) and (POP) are equivalent. In his work, Peña et. al. give many such conditions.

Corollary 2. *The problems (POP) and (CPP) are equivalent in the sense of Theorem 22 if deg(q) $\leq d$, deg(h_i) $= d$, q is bounded below on $\mathbb{R}^n_+$ and $h_i(x) \geq 0$ on $\mathbb{R}^n_+$ for $1 \leq i \leq m$*

Proof. If q is bounded below on $\mathbb{R}^n_+$, then Lemma 10 implies $\tilde{q}(x) \geq 0$ for $x \in (\mathbb{R}^n_+)^\infty = (\mathbb{R}^n_+)$. This, and the fact that h_i is nonnegative on $\mathbb{R}^n_+$ make **Conditions II** of Lemma 11 hold. The result follows from Theorem 22. □

8 Applications

8.1 Robust Regression

Consider the problem of linear regression using least squares. We are given data $(X_i)_{1\leq i\leq n} \subset \mathbb{R}^d$ to predict $(Y_i)_{1\leq i\leq n} \subset \mathbb{R}$, and we want the linear prediction that best minimizes the squared error. Specifically, we would like to solve

$$\min_\beta \sum_{i=1}^n \|X_i\beta - Y_i\|^2 \tag{MCO}$$

The solution β^* to this problem is well known. If we form the matrix X with rows X_i, and project the vector Y onto the column space of X, then the solution β^* is the vector such that $\text{Proy}_X(Y) = X\beta^*$. Moreover, if we assume that each observation Y_i is generated through a probabilistic process given as

$$Y_i = X\beta + \epsilon_i \tag{241}$$

where the ϵ_i are independent random variables of mean 0 and variance σ^2, then the Gauss-Markov Theorem states that the unbiased estimator of β with the lowest variance is given by MCO.

In many applications, however, MCO does not give a good fit of the data. A fundamental problem might be that the relation between X and Y is not linear but since linear models tend to be easier to solve and interpret, data analysis tends to stick with them. A different problem is the presence of outliers in the data, sometimes generated by the incorrect collection of data, or because the errors ϵ_i are not identically distributed but have different volatilities.

A scenario where this is a relevant problem is in calculating risk premiums for health insurance. In a competitive market, these premiums are roughly the expected value of the total spending of each individual. To estimate these expected values, a linear model is fit to the data (X, Y) that contains medical diagnoses and other relevant factors, along with the realized spending of previous years. Spending, however, is highly volatile, and this voltility is not constant across the population (notice, for example, that there is higher volatility on the more severe patients than on healthy young individuals), so this leads to ill fitted models. Riascos [19], for example, shows that in Colombia the incorrect use of linear models leads to an inefficient repartition of the government funds to different insurance companies.

To solve the problem generated by outliers, many methods have been proposed. The simplest are in identifying outliers visually or through rules of thumb before fitting any model. These, however, can only be implemented when the dimensionality of the data allows visualization, or when there is a good knowledge of the nature of the database.

To overcome this difficulty, Zioutas [20] proposes Least Trimmed Squares (LTS) to find a good linear model and automatically eliminate a set of outliers from the data set.

$$\min_{\beta} \sum_{i=1}^{n} z_i \left\| X_i\beta - Y_i \right\|^2 \tag{LTS}$$

$$s.t.\ z_i \in \{0,1\} \tag{242}$$

$$z^T e = k \tag{243}$$

Notice that after fixing k, the problem will try to find the best fit of the data while removing $n-k$ outliers. This problem has been extensively studied. In particular, Giloni [22] shows that it can be expressed as the minimization of a concave function over a polyhedron, which is known to be NP-hard. Another very interesting read is Nguyen [23], who gives an approximate solution using semidefinite approximation. Although he gives no hierarchy that converges to the exact solution.

Notice that this problem is a Polynomial Optimization Problem. The fact that z_i is binary can be replaced by the condition $z_i(z_i - 1) = 0$ but there is no need. It is easy to show that the next relaxation is equivalent to (LTS)

$$\min_{\beta} \sum_{i=1}^{n} z_i \left\| X_i\beta - Y_i \right\|^2 \tag{LTSR}$$

$$s.t.\ 0 \leq z_i \leq 1 \tag{244}$$

$$z^T e = k \tag{245}$$

If there are n observations of dimension d, this is a POP with a polynomial of degree 3 in $n+d$ variables, subject to $n+1$ linear restrictions.

To create a CPP relaxation we change the condition $z_i \leq 1$ for $z_i + y_i = 1$ and $y_i \geq 0$. We also need all polynomials to be of the same degree, so we change each linear restriction for a cubic equivalent linear restriction

$$\begin{aligned} \min_{\beta} \sum_{i=1}^{n} z_i \left\| X_i\beta - Y_i \right\|^2 & \\ s.t.\ (z_i + y_i - 1)^2(z_i + y_i + 1) = 0 & \quad \text{(LTS2)} \\ (z^T e - k)^2(z_i + y_i + 1) = 0 & \\ z_i, y_i \geq 0 & \end{aligned} \tag{246}$$

Since the objective function is bounded below by 0 and each restriction is nonnegative on $\mathbb{R}^n_+$, Corollary 2 implies (LTS2) is equivalent to its completely positive relaxation (CPP).

8.2 Clique and Stability Numbers

Given a graph G, the maximum clique problem asks to find the size of the largest complete subgraph (clique) of G, call this number $k(G)$. Motzkin and Strauss [34] showed that $k(G)$ can be found solving the quadratic optimization problem (CLIQUE).

Theorem 23 (Motzkin and Strauss)**.** *Let $G = (V, E)$ be a graph on n vertices. Let A be its adjacency matrix and let $k(G)$ be its clique number. Then the problem*

$$\begin{aligned} f(G) := \max_x \ & x^T A x \\ s.t. \ & e^T x = 1 \qquad (CLIQUE) \\ & x \geq 0 \end{aligned} \tag{247}$$

has as solution

$$f(G) = 1 - \frac{1}{k(G)} \tag{248}$$

Proof. First, notice that

$$x^T A x = \sum_{(i,j) \in E} x_i x_j \tag{249}$$

Assume that the largest complete subgraph is given by the vertices $S \subseteq V$. If we define

$$x_i = \begin{cases} \frac{1}{k} \text{ for } i \in S \\ 0 \text{ for } i \notin S \end{cases} \tag{250}$$

then x is feasible. Moreover,

$$f(G) \geq \sum_{(i,j) \in E} x_i x_j = 2 \binom{k}{2} \frac{1}{k^2} = 1 - \frac{1}{k(G)} \tag{251}$$

To show the other inequality we will proceed by induction on n. For $n = 1$ the result is trivial. Now assume the theorem holds for graphs of size less than n and define

$$F(x) = x^T A x \tag{252}$$

so that for any optimum vector x^* then $f(G) = F(x^*)$ (note that the feasible set is compact, so at least one optimal vector exists). If for some i we have $x_i^* = 0$, then construct the graph G' removing the i-th vertex. The vector obtained from removing the i-th component of x^* will be feasible for $f(G')$ and obtain the same optimal value. This means

$$f(G) \leq f(G') = 1 - \frac{1}{k(G')} \leq 1 - \frac{1}{k(G)} \tag{253}$$

which shows the other inequality. Now, assume that all optimal solutions satisfy $x^* > 0$. Define $s(x) = e^T x$ and notice that the function

$$P(x) = \frac{F(x)}{s(x)^2} = \frac{x^T A x}{(e^T x)^2} \tag{254}$$

is homogeneous of degree 0, so x^* is optimal for the problem of maximizing $P(x)$ over $\mathbb{R}^n_+$. Since x^* is an interior point, the first order conditions must hold

$$P'(x^*) = 0 \Rightarrow s^2 F_i = 2sF \Rightarrow F_i(x^*) = 2F(x^*) \tag{255}$$

(the last inequality comes from $s(x^*) = 1$).

Now assume that G is not a complete graph. Let us say, without loss of generality, that $(1,2) \notin E$. This implies

$$F(x_1 - c, x_2 + c, x_3, \ldots, x_n) = \sum_{(i,j)\in E} x_i x_j - 2 \sum_{j:(1,j)\in E} c x_j + 2 \sum_{j:(2,j)\in E} c x_j \tag{256}$$

$$= F(x) - c(F_1(x) - F_2(x)) \tag{257}$$

So when we evaluate at an optimal point x^*, take $c = x_1$, and use $F_1(x^*) = F_2(x^*)$, then

$$F(0, x_1^* + x_2^*, x_3^*, \ldots, x_n^*) = F(x_1^*, x_2^*, x_3^*, \ldots, x_n^*) \tag{258}$$

If we construct G' removing edge 1 of G, then the vector $(x_1^* + x_2^*, x_3^*, \ldots, x_n^*)$ will be feasible for the problem $f(G')$ and achieve the same value. The induction hypothesis implies

$$f(G) \leq f(G') = 1 - \frac{1}{k(G')} \leq 1 - \frac{1}{k(G)} \tag{259}$$

which shows the other inequality. Finally, assume G is a complete graph. This means

$$F(x) = [(x_1 + \cdots + x_n)^2 - (x_1^2 + \cdots + x_n^2)] = 1 - \|x\|^2 \tag{260}$$

$$\implies \max_{e^T x = 1} F(x) = 1 - \min_{e^T x = 1} \|x\|^2 = 1 - \frac{1}{n} \tag{261}$$

and the proof is complete. □

Corollary 3. *Let $\alpha(G)$ be the stability number of G, i.e., the maximum order of an empty subgraph. Then $f(G) = \frac{1}{\alpha(G)}$ is the solution to the program*

$$\begin{aligned} f(G) := \min_x \ & x^T(A+I)x \\ s.t. \ & e^T x = 1 \qquad (STAB) \\ & x \geq 0 \end{aligned} \tag{262}$$

Proof. Let $\bar{G}$ be the complement graph of G, i.e., the graph that has the same set of vertices as G and $(i,j) \in E(\bar{G})$ iff $(i,j) \notin E(G)$). Note that a clique of $\bar{G}$ corresponds to an empty subgraph of G, so $\alpha(G) = k(\bar{G})$. Since the adjacency matrix of $\bar{G}$ is $E - A - I$, its objective function in (CLIQUE) will be

$$F(x) = x^T(E - A - I)x = 1 - x^T(A+I)x \tag{263}$$

and will give the optimal value $f(\bar{G}) = \frac{1}{k(\bar{G})}$. A proper transformation of the objective function gives the result. □

Theorem 23 states that $\frac{1}{k(G)} = 1 - f(G)$. Note that $x^T e = 1$ is equivalent to $1 = x^T E x$, so the solution to

$$\min_x \ x^T(E - A)x \tag{264}$$

$$s.t. \ x^T E x = 1 \tag{265}$$

$$x \geq 0 \tag{266}$$

should be $1 - f(G) = \frac{1}{k(G)}$. To find $k(G)$ directly we could invert the objective function and obtain the problem

$$\max_x \ \frac{x^T E x}{x^T (E - A) x} \tag{267}$$

$$s.t.\ e^T x = 1 \tag{268}$$

$$x \geq 0 \tag{269}$$

which is a fractional optimization program that using Theorem 20 is equivalent to (CLIQUE$_{\mathcal{C}}$) below. Analogously, finding the stability number of G can be done through the completely positive program (STAB$_{\mathcal{C}}$).

$$\begin{array}{lcl} k(G) = \max\limits_X \ \langle E, X \rangle & & \alpha(G) = \max\limits_X \ \langle E, X \rangle \\ s.t.\ \langle E - A, X \rangle = 1 & \text{(CLIQUE}_{\mathcal{C}}\text{)} & s.t.\ \langle A + I, X \rangle = 1 \quad \text{(STAB}_{\mathcal{C}}\text{)} \\ X \in \mathcal{C} & & X \in \mathcal{C} \end{array} \tag{270}$$

Formulations CLIQUE$_{\mathcal{C}}$ and STAB$_{\mathcal{C}}$ were first shown by De Klerk [24], and allow the calculation of the stability and clique numbers of a graph through completely positive programming. We can also write them as copositive programs. To see this, recall $\mathcal{C}^*$ from Equation (153). The dual to (CLIQUE$_{\mathcal{C}}$) and (STAB$_{\mathcal{C}}$) will be the copositive programs

$$\begin{array}{lcl} k(G) = \min\limits_y \ y & & \alpha(G) = \min\limits_y \ y \\ s.t.\ y(E - A) - E \in \mathcal{C}^* & \text{(CLIQUE}^*_{\mathcal{C}}\text{)} & s.t.\ y(A + I) - E \in \mathcal{C}^* \quad \text{(STAB}^*_{\mathcal{C}}\text{)} \end{array} \tag{271}$$

Theorem 24. *Determining if a matrix M is not copositive is an NP-Complete problem*

Proof. Determining the clique number is an NP-Complete problem and to determine the clique number, we just need to verify which matrices $y(E - A) - E$ are copositive for $y = 1, \ldots, n$. So the time it would take to find $k(G)$ is linear on the time checking copositivity of each matrix would take. □

8.3 Lovasz Theta Function

Consider again the problem of finding the stability number of a graph $G = (V, E)$. For $S \subseteq V$ define the vector

$$(x_S)_v := \begin{cases} 1 \text{ if } v \in S \\ 0 \text{ if } v \notin S \end{cases} \tag{272}$$

and define the matrix

$$X_S := \frac{x_S x_S^T}{x_S^T x_S} \tag{273}$$

The matrix X_S is both positive semidefinite and completely positive. Moreover

$$\operatorname{tr}(X_S) = \frac{1}{x_S^T x_S} \sum_{v \in V} X_{vv} = \frac{1}{\sum_{v \in V} (x_S)_v^2} \sum_{v \in V} (x_S)_v^2 = 1 \tag{274}$$

$$\langle E, X \rangle = e^T X e = \frac{(e^T x_S)^2}{x_S^T x_S} = \frac{|S|^2}{|S|} = |S| \tag{275}$$

Finally, notice that if S is an independent set, then $(X_S)_{ij} = (x_S)_i(x_S)_j = 0$ whenever $(i,j) \in E$. This suggests two natural relaxation to the problem of finding $\alpha(G)$, namely

$$\begin{aligned} \vartheta(G) := \max_X \langle E, X\rangle \quad (\text{LOV}) \qquad & \vartheta^{\mathcal{C}}(G) := \max_X \langle E, X\rangle \quad (\text{LOV}_{\mathcal{C}}) \\ s.t.\ X_{i,j} = 0\ \forall (i,j) \in E \qquad & s.t.\ X_{i,j} = 0\ \forall (i,j) \in E \\ \text{tr}(X) = 1 \qquad & \text{tr}(X) = 1 \\ X \succeq 0 \qquad & X \in \mathcal{C} \end{aligned} \tag{276}$$

The function $\vartheta(G)$ is known as the Lovásh Theta Function, and was first introduced by Lovásh in [25] to study the Shannon Capacity of a graph. The function $\vartheta^{\mathcal{C}}(G)$ was first introduced by De Klerk in [24]. We have the next result from both papers

Theorem 25. *The Lovash relaxations and the independence number are related as*

$$\alpha(G) = \vartheta^{\mathcal{C}}(G) \leq \vartheta(G) \tag{277}$$

Proof. Assume S is and independent set with cardinality $\alpha(G)$, then the previous discussion shows that X_S is feasible for $\text{LOV}_{\mathcal{C}}$ and achieves a value $k(G)$. Also, if X is feasible for $(\text{LOV}_{\mathcal{C}})$, it will be feasible for (LOV) and will achieve the same value. This implies

$$\alpha(G) \leq \vartheta^{\mathcal{C}}(G) \leq \vartheta(G) \tag{278}$$

Now, remember the completely positive relaxation to the stability number $(\text{STAB}_{\mathcal{C}})$ in Equation (270). Notice that if X is feasible for $(\text{LOV}_{\mathcal{C}})$ it will also be feasible for $(\text{STAB}_{\mathcal{C}})$ and achieve the same value. Since both are maximization problems, this implies

$$\alpha(G) \geq \vartheta^{\mathcal{C}}(G) \tag{279}$$

which concludes the proof. □

8.4 Chromatic Number

In the previous section we showed that the theta function $\vartheta(G)$ is an upper bound on the stability number. Then we strengthen the program with a completely positive formulation and showed the equality $\vartheta^{\mathcal{C}}(G) = \alpha(G)$. Lovásh also showed that the theta function was a lower bound on the chromatic number of a graph, specifically $\vartheta(\bar{G}) \leq \chi(G)$. In this section we follow Dukanovic [26] to strengthen this program through completely positive programming and see if we can get closer to $\chi(G)$. Specifically, we show that the completely positive formulation gives the fractional chromatic number $\chi_f(G)$.

An s-coloring of a graph $G = (V, E)$ is an assignment of colors $c : V \to \{1, \ldots, s\}$ such that if $(i,j) \in E$ then $c(i) \neq c(j)$, i.e., two adjacent nodes do not share the same color. Any coloring c partitions V into s independent subsets $S_i = c^{-1}(i)$. Similarly, any partition of V into s independent subsets $S_1, \ldots, S_s$ induces a coloring $c(i) = p$ where p is the unique integer with the property $i \in S_p$.

If we denote $\mathcal{S} = \{S \subset V : S \text{ is independent}\}$, then we can characterize a partition of V into s independent sets as a subfamily $\{S_1, \ldots, S_s\} \subset \mathcal{S}$ such that

$$x_{S_1} + \cdots + x_{S_s} = e \tag{280}$$

The chromatic number of G, denoted as $\chi(G)$, is defined as the minimum s such that there exists an s-coloring of G. This problem, can thus be written as

$$\chi(G) = \min_{\lambda} \sum_{S \in \mathcal{S}} \lambda_S \tag{281}$$

$$s.t. \sum_{S \in \mathcal{S}} \lambda_S x_S = e \tag{CHROM}$$

$$\lambda_S \in \{0, 1\} \tag{282}$$

The fractional chromatic number of G, denoted as $\chi_f(G)$, is the solution to the relaxation

$$\chi_f(G) = \min_{\lambda} \sum_{S \in \mathcal{S}} \lambda_S \tag{283}$$

$$s.t. \sum_{S \in \mathcal{S}} \lambda_S x_S = e \tag{CHROM$_f$}$$

$$0 \leq \lambda_S \leq 1 \tag{284}$$

A well-known result is that computing $\chi(G)$ is an NP-Hard problem, moreover, no algorithms to compute $\chi_f(G)$ in polynomial time are known yet. To motivate semidefinite and copositive relaxations of $\chi(G)$, we define a coloring matrix.

Definition 14. We say that C is a coloring matrix if there exists a $\chi(G)$-coloring of G with an associated partition $\mathcal{T} = S_1, \ldots, S_{\chi(G)}$ such that

$$C = \sum_{i=1}^{\chi(G)} x_{S_i} x_{S_i}^T \tag{285}$$

A coloring matrix uniquely determines its partition $\mathcal{T}$. To see this take $u, v \in V$ and denote $S_{[u]}, S_{[v]}$ their corresponding subsets in the partition, then, notice that

$$c_{uv} = e_u^T C e_v = \sum_{i=1}^{s} (e_u^T x_{S_i})(x_{S_i}^T e_v) = \begin{cases} 1 & \text{if } S_{[u]} = S_{[v]} \\ 0 & \text{if } S_{[u]} \neq S_{[v]} \end{cases} \tag{286}$$

Since an appropriate coloring requires that connected vertices v, w have different colors, then we must have $c_{uv} = 0$ whenever $(u, v) \in E$. Also, notice that $\text{diag}(C) = e$. These properties are summarized in Proposition 12. To motivate a semidefinite relaxation we need the following lemma

Lemma 12. *Let $[\lambda_S]_{S \subseteq V}$ a vector defined over the subsets of vertices such that $\lambda_S \geq 0$ and define*

$$X_\lambda = \sum_{S \subseteq V} \lambda_S x_S x_S^T \tag{287}$$

Then we have

$$M := \left(\sum_{S \subseteq V} \lambda_S \right) X_\lambda - diag(X_\lambda) diag(X_\lambda)^T \succeq 0 \tag{288}$$

Proof. The vector x_S has 0-1 entries, so $\text{diag}(x_S x_S^T) = x_S$. This means that

$$diag(X_\lambda) = \sum_{S \subseteq V} \lambda_S x_S \tag{289}$$

The matrix M can thus be written as

$$\begin{aligned} M &= \sum_{S,T \subset V} \lambda_S \lambda_T x_T x_T^T - \sum_{S,T \subset V} \lambda_S \lambda_T x_S x_T^T && (290) \\ &= \sum_{S,T \subset V} \lambda_S \lambda_T (x_T x_T^T - x_S x_T^T) && (291) \end{aligned}$$

In order to show that $M \succeq 0$ consider the product $y^T M y$ and denote $a_T = y^T x_T$, $a_S = y^T x_S$

$$\begin{aligned} y^T M y &= \sum_{S,T \subset V} \lambda_S \lambda_T ((y^T x_T)^2 - (y^T x_S)(y^T x_S)) && (292) \\ &= \sum_{S,T \subset V} \lambda_S \lambda_T (a_T^2 - a_T a_S) && (293) \\ &= \sum_{S,T \subset V} \lambda_S \lambda_T (a_T^2 - a_T a_S) + \left[\frac{1}{2} \sum_{S,T \subset V} \lambda_S \lambda_T (a_S^2) - \frac{1}{2} \sum_{S,T \subset V} \lambda_S \lambda_T (a_T^2) \right] && (294) \\ &= \frac{1}{2} \sum_{S,T \subset V} \lambda_S \lambda_T (a_T^2 - 2 a_T a_S + a_S^2) && (295) \\ &= \frac{1}{2} \sum_{S,T \subset V} \lambda_S \lambda_T (a_T - a_S)^2 \geq 0 && (296) \end{aligned}$$

□

The next proposition summarizes the properties of coloring matrices that we will use

Proposition 12. *If C is a coloring matrix, then C satisfies the following properties*

(i) It $t \geq \chi(G)$ then $tC - E \succeq 0$

(ii) $diag(C) = e$

(iii) $C_{uv} = 0$ if $(u, v) \in E$

(iv) $C \in \mathcal{C}$

(v) $C \succeq 0$

Proof. Properties (ii) and (iii) were already shown. Properties (iv) and (v) come from the definition of a coloring matrix. To show property (i) take the partition induced by C, $\mathcal{T} = (S_1, \ldots, S_{\chi(G)})$, and define

$$\lambda_S = \begin{cases} 1 & \text{if } S = S_i \text{ for some } i \\ 0 & \text{in other case} \end{cases} \tag{297}$$

Notice that $X_\lambda = C$. Since $\sum_{S \subset V} \lambda_S = \chi(G)$ and property (i) shows $\text{diag}(C) = e$, then Lemma 12 implies that

$$\chi(G) C - E \succeq 0 \tag{298}$$

Finally, whenever $t \geq \chi(G)$ then

$$tC - E = (t - \chi(G))C + (\chi(G)C - E) \succeq 0 \tag{299}$$

since $C \succeq 0$. $\square$

These properties induce the semidefinite and the completely positive relaxations

$$\begin{aligned} \Theta(G) := \min_{t,C} \ & t \\ s.t. \ & C_{uv} = 0 \text{ if } (u,v) \in E \\ & \operatorname{diag}(C) = e \\ & tC - E \succeq 0 \end{aligned} \quad (\text{CHROM}_P) \qquad \begin{aligned} \Theta^{\mathcal{C}}(G) := \min_{t,C} \ & t \\ s.t. \ & C_{uv} = 0 \text{ if } (u,v) \in E \\ & \operatorname{diag}(C) = e \\ & tC - E \succeq 0 \\ & C \in \mathcal{C} \end{aligned} \quad (\text{CHROM}_{\mathcal{C}}) \tag{300}$$

The next theorem, proved in [26], shows the relationship between these numbers

Theorem 26. *We have the following relationship between the fractional chromatic number and its relaxations*

$$\vartheta(\bar{G}) = \Theta(G) \leq \Theta^{\mathcal{C}}(G) = \chi_f(G) \leq \chi(G) \tag{301}$$

9 Computational Simulations

In this section we show the results of computational simulations to calculate the independence number of Hamming and DeBrujin graphs.

Definition 15 (Hamming and DeBrujin Graphs)**.** For $d, q \in \mathbb{N}$ define V as the set of words with d letters on an alphabet of q characters.

- For the Hamming Graph $H(d,q)$, $(v,w) \in E$ if and only if v and w are words that differ exactly by one character.
- For the DeBrujin graph $B(d,q)$, $([v_1,\ldots,v_d],[w_1,\ldots,w_d]) \in E$ if and only if

$$(v_2,\ldots,v_d) = (w_1,\ldots,w_{d-1}) \tag{302}$$

$$\text{or} \quad (v_1,\ldots,v_{d-1}) = (w_2,\ldots,w_d) \tag{303}$$

Assume we choose $A \subset V$ as a dictionary of words to communicate through a channel. Because of errors, it may happen that one of the letters of a word is altered. Then, an independent set A of $H(d,q)$ assures us that if there is an error, the received word will not be in the dictionary, hence, the error will be detected. The independent set is the maximum cardinality of a dictionary with such property.

DeBrujin graphs have an interesting interpretation when $d = 3$. In this case, an independent set $A \subset V$ will be a dictionary with the property that if we write one word after the other, and then take a random fragment of what we've written, we will be able to identify the exact places at which one word follows the other, without the need of commas. This is known as a comma-free code.

The next table shows the results in calculating the independence number of several Hamming and DeBrujin graphs. We approximate the number from above using the $\mathcal{SOS}$ hierarchy (140) of the copositive problem $\text{STAB}^*_{\mathcal{C}}$ given in equation (271). To approximate it from below we use the $\mathcal{BVL}$ hierarchy (121) of the completely positive problem $\text{STAB}^*_{\mathcal{C}}$ given in equation (270). We used the CVX toolbox for MATLAB to formulate the problems, and then MOSEK to solve them.

Table 1: $B(3,3)$ $\alpha(G) = 9$

Method	**Hierarchy**	**Value**	**Time**
BVL	1	3.86	0.6
SOS	1	9.41	364.0
Theta		9.41	0.7

Table 2: $B(2,3)$ $\alpha(G) = 3$

Method	**Hierarchy**	**Value**	**Time**
BVL	1	1.77	0.3
BVL	2	1.96	0.7
BVL	3	2.11	3.8
SOS	1	3.00	2.5
Theta		3.00	0.3

Table 3: $H(2,4)$ $\alpha(G) = 8$

Method	**Hierarchy**	**Value**	**Time**
BVL	1	1.77	0.3
BVL	2	1.96	0.7
BVL	3	2.11	3.8
SOS	1	3.00	2.5
Theta		3.00	0.4

Table 4: $H(2,4)$ $\alpha(G) = 8$

Method	**Hierarchy**	**BVL**	**Time**
BVL	1	2.97	0.2
BVL	2	3.09	0.6
BVL	3	3.23	9.5
SOS	1	8.00	10.8
Theta		8.00	0.4

Table 5: Random Erdos graph in 10 nodes with $p = 0.3$ $\alpha(G) = 4$

Method	**h**	**BVL**	**Time**
BVL	1	1.87	0.2
BVL	2	1.98	0.3
BVL	3	2.11	0.7
BVL	4	2.25	6.9
SOS	1	4.00	1.8
Theta		4.00	0.3

Table 6: $R(10, 0.3)$ $\alpha(G) = 8$

Method	**Hierarchy**	**BVL**	**Time**
BVL	1	1.96	0.2
BVL	2	2.21	0.2
BVL	3	2.38	1.5
BVL	4	2.51	26.0
SOS	1	4.19	1.2
SOS	2	4.00	25.2
Theta		4.19	0.2

The analysis of the computational results give interesting insights

- The $\mathcal{SOS}$ first approximation and the Lozasz Theta Function are equal. A quick search of the literature shows DeKlerk [24] was the first to show it.
- The $\mathcal{SOS}$ approximation is exact on the first or second approximation. This enters in harsh contrast with the $\mathcal{BVL}$ approximation, which was not exact for any of the problems, even the smaller ones.

- Even though the first $\mathcal{SOS}$ approximation should be equivalent to the Lovasz Theta function, solving the first problem takes significantly more time than the second. Our explanation is that solving the first $\mathcal{SOS}$ approximation with its direct formulation is inefficient. Instead, we should look for simplifications of the hierarchies, such as the fact that $\mathcal{SOS}_0 = \mathcal{DN}^* = \mathcal{S}_+ + \mathcal{N}$. The Lovasz Theta is the result of this simplification, which would explain the increase in efficiency.

- The size of the matrices involved in solving the psd problem rapidly blow up, making the computer run out of memory. Similarly, going from one hierarchy to the next hierarchy will increase the computing time from seconds to hours. In this sense, the $\mathcal{BVL}$ approximation is kinder, since we could go up to the fourth hierarchy, while we could only go up to the second for $\mathcal{SOS}$.

10 Future Research

In this work we gave a comprehensive account of copositive programming. We gave concrete semidefinite upper and lower approximation and emphasized their application in solving problems in graph theory and data analysis. Particularly, the bounds provided by the $\mathcal{SOS}$ and $\mathcal{BVL}$ hierarchies allowed to calculate bounds on the independence number of a graph.

However, we noticed that there are still strong computational limitations to making copositive programming an alternative in solving large polynomial optimization and data analysis problems. In the short run, future research should be directed towards:

- Understanding fast computational solutions to SDP problems, particularly those that do not rely on interior point methods.

- Give simpler and equivalent descriptions of the $\mathcal{SOS}$ hierarchies, to require less memory and power on the computers.

In the middle run, we think it is important to:

- Understand how to recover optimal solutions from the convex relaxations of copositive problems, similarly to the randomization procedure in the Goemans-Williamson algorithm. We know such recovery can be accomplished for the $\mathcal{SOS}$ hierarchy when the gap between the optimal value and the approximation is zero. However, more attention should be given to lower approximations such as $\mathcal{BVL}$. In particular, the lower approximation of the independence number, even when not exact, should aid the search for large independent subgraphs.

- Solve copositive programs using the capabilities of parallel computing.

In the long run, we believe that faster alternatives to sum of squares procedures should be proposed. One important route could be linear programming, which is now in a more advanced state of development than sdp.

References

[1] Barvinok, Veomett. The computational complexity of convex bodies. University of Michigan. 2007

[2] J. Lasserre. New approximations for the cone of copositive matrices and its dual. arXiv:1012.2552

[3] J. Romero, M.Velasco. Semidefinite approximations of conical hulls of measured sets. arXiv:1409.8272

[4] A. Barvinok, E. Veomett. The computational complexity of convex bodies. arXiv:math/0610325

[5] P. Parrilo. Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization, Ph.D. Dissertation, California Institute of Technology, 2000.

[6] G. H. Hardy, J. E. Littlewood, and G. P olya. Inequalities. Cambridge Univer- sity Press, Cambridge, UK, 1967.

[7] B. Reznick. Uniform denominators in Hilberts seventeenth problem. Math. Z., 220:7597, 1995.

[8] Folland G.B.: How to Integrate a Polynomial over a Sphere, The American Mathematical Monthly, Vol. 108-5, May 2001, pp. 446-448

[9] Dickinson P.J.C., An improved characterisation of the interior of the completely positive cone. Electronic Journal of Linear Algebra ISSN 1081-3810. Volume 20, pp. 723-729, November 2010

[10] Dür M., Still G., Interior points of the completely positive cone. Electronic Journal of Linear Algebra 17(2008): 4853.

[11] A. Berman and N. Shaked-Monderer. Completely Positive Matrices. World Scientific Publishing Co., River Edge, NJ, 2003.

[12] Hamilton-Jester, Crista Lee; Li, Chi-Kwong. Extreme Vectors of Doubly Nonnegative Matrices. Rocky Mountain J. Math. 26 (1996), no. 4, 1371–1383. doi:10.1216/rmjm/1181071993. http://projecteuclid.org/euclid.rmjm/1181071993.

[13] J. E. Maxfield and H. Mine. On the equation $X'X = A$. Proc. Edinburgh Math. Soc.. 13:125 129, 1962.

[14] Diananda, P. H. "On Nonnegative Forms in Real Variables Some or All of Which Are Nonnegative." Proc. Cambridge Philos. Soc. 58, 17-25, 1962.

[15] Dong, H.: Symmetric tensor approximation hierarchies for the completely positive cone. SIAM J. Optim. 23(3), 18501866 (2013)

[16] J. Peña, J. Vera, and L. Zuluaga "Completely positive reformulations for polynomial optimization," To Appear in Mathematical Programming.

[17] Preisig, J.C. 1996. Copositivity and the minimization of quadratic functions with non- negativity and quadratic equality constraints., SIAM J. Control Optimization 34, no. 4, 11351150

[18] Burer, Samuel. 2009. On the copositive representation of binary and continuous nonconvex quadratic programs, Mathematical Programming 120, no. 2, 479495.

[19] Riascos A., Romero., M. The Performance of Risk Adjustment Models in Colombian Competitive Health Insurance Market

[20] Zioutas G, Avramidis A (2005) Deleting outliers in robust regression with mixed integer programming. Acta Math Appl Sin 21:323334

[21] Rousseeuw PJ, Leroy AM (1987) Robust regression and outlier detection. Wiley, New York

[22] Giloni A, Padberg M (2002) Least trimmed squares regression, least median squares regression, and mathematical programming. Math Comput Model 35:10431060

[23] Nguyen, T. D., Welsch, R. Outlier Detection and Least Trimmed Squares Approximation Using Semi-definite Programming Comput. Stat. Data Anal. December, 2010.

[24] De Klerk, Etienne and D. V. Pasechnik. 2002. Approximation of the stability number of a graph via copositive programming, SIAM Journal on Optimization 12, no. 4, 875892.

[25] Lovsz, Lszl (1979), "On the Shannon Capacity of a Graph", IEEE Transactions on Information Theory, IT-25.

[26] Dukanovix, I., Rendl, F. Copositive programming motivated bounds on the stability and the chromatic number. Preprint. University of Klagenfurt, 2006.

[27] Stephen Boyd and Lieven Vandenberghe. 2004. Convex Optimization. Cambridge University Press, New York, NY, USA.

[28] Greub, Werner. Multilinear Algebra. 2013. Second Edition. Universitext, Springer-Verlag. New York, NY, USA.

[29] M. Jr. Hall and M. Newman. Copositive and completely positive quadratic forms. Proc. Cambridge Philos. Soc., 59:329339, 1963

[30] Motakuri Ramana, A. J. Goldman. Some geometric results in semidefinite programming. Journal of Global Optimization, 1995, Volume 7, Number 1, Page 33

[31] D. Hilbert. Uber die Darstellung definiter Formen als Summe von Formen-quadraten. Math. Ann. 32, 342350 (1888).

[32] K. G. Murty and S. N. Kabadi. Some NP-complete problems in quadratic and nonlinear programming. Mathematical Programming, 39:117129, 1987.

[33] Bundfuss S., Copositive matrices, copositive programming, and applications. Ph.D. Dissertation, TU Darmstadt 2009. Online at http://www3.mathematik.tu- darmstadt.de/index.php?id=483

[34] Motzkin, T.S., Straus, E.G., Maxima for graphs and a new proof of a theorem of Turan, Canadian Journal of Mathematics 17(1965), 533540.