OPTIMIZATION AND APPROXIMATION OF NC POLYNOMIALS WITH SUMS OF SQUARES

Kristijan Cafuta
University of Ljubljana, Faculty of Electrical Engineering
Tržaška 25, 1000 Ljubljana, Slovenia, E-mail: kristijan.cafuta@fe.uni-lj.si

Igor Klep[1]
University of Ljubljana/Faculty of Science, Koroška 160, 2000 Maribor
University of Ljubljana, Faculty of Mathematics and Physics
Jadranska 19, 1111 Ljubljana, Slovenia, E-mail: igor.klep@fmf.uni-lj.si

Janez Povh[2]
Faculty of Information Studies in Novo Mesto
Novi trg 5, 8000 Novo mesto, Slovenia, E-mail: janez.povh@fis.unm.si

ABSTRACT. In this paper we study eigenvalue optimization of non-commutative polynomials. That is, we compute the smallest or biggest eigenvalue a non-commutative polynomial can attain. Our algorithm is based on sums of hermitian squares. To test for exactness, the solutions of the dual SDP are investigated. When we consider the eigenvalue lower bounds we can show that attainability of the optimal value on the dual side implies that the eigenvalue bound is attained. We also show how to extract global eigenvalue optimizers with a procedure based on two ingredients:

- the first is the solution to the truncated (tracial) moment problem;
- the second is the Gelfand-Naimark-Segal (GNS) construction.

The implementation of these procedures in our computer algebra system NC-SOStools is presented and several examples pertaining to matrix inequalities are given to illustrate the results.

Key words: *noncommutative polynomial, sum of squares, semidefnite programming, trace optimization, eigenvalue optimization, free positivity*

1. INTRODUCTION

Starting with Helton's seminal paper [Hel02], *free semialgebraic geometry* is being established. Among the things that make this area exciting are its many facets of applications. A nice survey on applications to control theory, systems engineering and optimization is given in [dOHMP08], while applications to mathematical physics and operator algebras have been done in [KS08a, KS08b, CKP10].

Unlike classical semialgebraic (or real algebraic) geometry where real polynomial rings in *commuting* variables are the objects of study, free semialgebraic geometry deals with real polynomials in *noncommuting* (NC) variables and their finite-dimensional representations. Of interest are various notions of *positivity* induced by these. For instance, positivity via positive semidefiniteness or the positivity of the trace. Both of these can be reformulated and studied using sums of hermitian squares (with commutators) and semidefinite programming.

We developed `NCSOStools` as a consequence of this recent interest in free noncommutative positivity and sums of (hermitian) squares (SOHS). `NCSOStools` is an open source Matlab toolbox for solving SOHS problems using semidefinite

[1]Partially supported by the Slovenian Research Agency (program No. P1-0222 and project No. J1-3608).
[2]Supported by the Slovenian Research Agency (project no. 1000-08-210518).

programming. As a side product our toolbox implements symbolic computation with noncommuting variables in Matlab. There is a small overlap in features with Helton's `NCAlgebra` package for Mathematica [HdOMS]. However, `NCSOStools` performs only basic manipulations with noncommuting variables, while `NCAlgebra` is a fully-fledged add-on for symbolic computation with polynomials, matrices and rational functions in noncommuting variables.

Readers interested in solving sums of squares problems for commuting polynomials are referred to one of the many great existing packages, such as SOSTOOLS [PPSP05], SparsePOP [WKK+09], GloptiPoly [HLL09], or YALMIP [Löf04].

This paper is organized as follows. The first section fixes notation and introduces terminology. Then in Section 2 we introduce the central objects, *sums of hermitian squares* and use these to study positive semidefinite NC polynomials. The natural correspondence between sums of hermitian squares and semidefinite programming is explained in some details in Section 3. The main theoretical contribution here is an algorithm to extract an eigenvalue minimizer of an NC polynomial. Detailed explanation with illustrative examples are in Section 4.

1.1. **Notation.** We write $\mathbb{N} := \{1, 2, \dots\}$, $\mathbb{R}$ for the sets of natural and real numbers. Let $\langle \underline{X} \rangle$ be the monoid freely generated by $\underline{X} := (X_1, \dots, X_n)$, i.e., $\langle \underline{X} \rangle$ consists of *words* in the n noncommuting letters $X_1, \dots, X_n$ (including the empty word denoted by 1).

We consider the algebra $\mathbb{R}\langle \underline{X} \rangle$ of polynomials in n noncommuting variables $\underline{X} = (X_1, \dots, X_n)$ with coefficients from $\mathbb{R}$. The elements of $\mathbb{R}\langle \underline{X} \rangle$ are linear combinations of words in the n letters $\underline{X}$ and are called *NC polynomials*. The length of the longest word in an NC polynomial $f \in \mathbb{R}\langle \underline{X} \rangle$ is the *degree* of f and is denoted by $\deg f$. We shall also consider the degree of f in X_i, $\deg_i f$. Similarly, the length of the shortest word appearing in $f \in \mathbb{R}\langle \underline{X} \rangle$ is called the *min-degree* of f and denoted by $\operatorname{mindeg} f$. Likewise, $\operatorname{mindeg}_i f$ is introduced. If the variable X_i does not occur in some monomial in f, then $\operatorname{mindeg}_i f = 0$. For instance, if $f = X_1^3 - 3X_3X_2X_1 + 2X_4X_1^2X_4$, then

$$\deg f = 4, \quad \deg_1 f = 3, \quad \deg_2 f = \deg_3 f = 1, \quad \deg_4 f = 2,$$

$$\operatorname{mindeg} f = 3, \quad \operatorname{mindeg}_1 f = 1, \quad \operatorname{mindeg}_2 f = \operatorname{mindeg}_3 f = \operatorname{mindeg}_4 f = 0.$$

0 An element of the form aw where $0 \neq a \in \mathbb{R}$ and $w \in \langle \underline{X} \rangle$ is called a *monomial* and a its *coefficient*. Hence words are monomials whose coefficient is 1.

We equip $\mathbb{R}\langle \underline{X} \rangle$ with the *involution* $*$ that fixes $\mathbb{R} \cup \{\underline{X}\}$ pointwise and thus reverses words, e.g.

$$(X_1^2 - X_2X_3X_1)^* = X_1^2 - X_1X_3X_2.$$

Hence $\mathbb{R}\langle \underline{X} \rangle$ is the $*$-algebra freely generated by n symmetric letters. Let $\operatorname{Sym} \mathbb{R}\langle \underline{X} \rangle$ denote the set of all *symmetric elements*, that is,

$$\operatorname{Sym} \mathbb{R}\langle \underline{X} \rangle = \{f \in \mathbb{R}\langle \underline{X} \rangle \mid f = f^*\}.$$

The involution $*$ extends naturally to matrices (in particular, to vectors) over $\mathbb{R}\langle \underline{X} \rangle$. For instance, if $V = (v_i)$ is a (column) vector of NC polynomials $v_i \in$

$\mathbb{R}\langle \underline{X} \rangle$, then V^* is the row vector with components v_i^*. We shall also use V^t to denote the row vector with components v_i.

2. Positive semidefinite NC polynomials

A symmetric matrix $A \in \mathbb{R}^{s\times s}$ is positive semidefinite if and only if it is of the form B^tB for some $B \in \mathbb{R}^{s\times s}$. In this section we introduce the notion of *sum of hermitian squares* (SOHS) and explain its relation with semidefinite programming.

An NC polynomial of the form g^*g is called a *hermitian square* and the set of all sums of hermitian squares will be denoted by Σ^2. A polynomial $f \in \mathbb{R}\langle \underline{X} \rangle$ is SOHS if it belongs to Σ^2. Clearly, $\Sigma^2 \subsetneq \operatorname{Sym} \mathbb{R}\langle \underline{X} \rangle$. For example,

$$X_1X_2 + 2X_2X_1 \notin \operatorname{Sym} \mathbb{R}\langle \underline{X} \rangle, \quad X_1^2X_2X_1^2 \in \operatorname{Sym} \mathbb{R}\langle \underline{X} \rangle \setminus \Sigma^2,$$

$$2 + X_1X_2 + X_2X_1 + X_1X_2^2X_1 = 1 + (1 + X_2X_1)^*(1 + X_2X_1) \in \Sigma^2.$$

If $f \in \mathbb{R}\langle \underline{X} \rangle$ is SOHS and we substitute symmetric matrices $A_1, \ldots, A_n$ of the same size for the variables $\underline{X}$, then the resulting matrix $f(A_1, \ldots, A_n)$ is positive semidefinite. Helton [Hel02] and McCullough [McC01] proved (a slight variant of) the converse of the above observation: if $f \in \mathbb{R}\langle \underline{X} \rangle$ and $f(A_1, \ldots, A_n) \succeq 0$ for *all* symmetric matrices A_i of the same size, then f is SOHS. For a beautiful exposition, we refer the reader to [MP05].

The following proposition (cf. [Hel02, §2.2] or [MP05, Theorem 2.1]) is the noncommutative version of the classical result due to Choi, Lam and Reznick ([CLR95, §2]; see also [Par03, PW98]). The easy proof is omitted.

Proposition 2.1. *Suppose $f \in \operatorname{Sym} \mathbb{R}\langle \underline{X} \rangle$ is of degree $\leq 2d$. Then $f \in \Sigma^2$ if and only if there exists a positive semidefinite matrix G satisfying*

$$f = W_d^* G W_d = \sum_{i,j} G_{i,j} (W_d)_i^* (W_d)_j, \tag{1}$$

where W_d is a vector consisting of all words in $\langle \underline{X} \rangle$ of degree $\leq d$.

Conversely, given such a positive semidefinite matrix G with rank r, one can construct NC polynomials $g_1, \ldots, g_r \in \mathbb{R}\langle \underline{X} \rangle$ of degree $\leq d$ such that

$$f = \sum_{i=1}^{r} g_i^* g_i. \tag{2}$$

The matrix G is called *a Gram matrix* for f. Polynomials g_i are products of i-th row of F with vector W_d, where F is such that $G = F^TF$.

Example 2.2. In this example we consider NC polynomials in 2 variables which we denote by X, Y. Let

$$f = 1 - 2X + 2X^2 + Y^2 - 2X^2Y - 2YX^2 + 2YXY + 2YX^2Y.$$

Let V be the subvector $\begin{bmatrix} 1 & X & Y & XY \end{bmatrix}^t$ of W_2. Then the Gram matrix for f with respect to V is given by

$$G(a) := \begin{bmatrix} 1 & -1 & 0 & a \\ -1 & 2 & -a & -2 \\ 0 & -a & 1 & 1 \\ a & -2 & 1 & 2 \end{bmatrix}.$$

(That is, $f = V^*G(a)V$.) This matrix is positive semidefinite if and only if $a = 1$ as follows easily from the characteristic polynomial of $G(a)$. Moreover, $G(1) = C^tC$ for

$$C = \begin{bmatrix} 1 & -1 & 0 & 1 \\ 0 & 1 & -1 & -1 \end{bmatrix}.$$

From

$$CV = \begin{bmatrix} 1 - X + XY & X - Y - XY \end{bmatrix}^t$$

it follows that

$$f = (1 - X + XY)^*(1 - X + XY) + (X - Y - XY)^*(X - Y - XY) \in \Sigma^2.$$

The problem whether a given polynomial is SOHS is therefore a special instance of a semidefinite feasibility problem. Our standard reference for semidefinite programming (SDP) is [WSV00].

3. Sums of hermitian squares and SDP

In this subsection we present a *conceptual algorithm* based on SDP for checking whether a given $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X} \rangle$ is SOHS. Following Proposition 2.1 we must determine whether there exists a positive semidefinite matrix G such that $f = W_d^* G W_d$, where W_d is the vector of all words of degree $\leq d$, given in a fixed order. This is a semidefinite feasibility problem in the matrix variable G, where the constraints $\langle A_i, G \rangle = b_i$ are implied by the fact that for each product of monomials $w \in \{p^*q \mid p, q \in W_d\} = W_{2d}$ the following must be true:

$$\sum_{\substack{p,q \in W_d \\ p^*q = w}} G_{p,q} = a_w, \tag{3}$$

where a_w is the coefficient of w in f ($a_w = 0$ if the monomial w does not appear in f).

Any input polynomial f is symmetric, so $a_w = a_{w^*}$ for all w, and equations (3) can be rewritten as

$$\sum_{\substack{u,v \in W_d \\ u^*v = w}} G_{u,v} + \sum_{\substack{u,v \in W_d \\ u^*v = w^*}} G_{u,v} = a_w + a_{w^*} \quad \forall w \in W_{2d}, \tag{4}$$

or equivalently,

$$\langle A_w, G \rangle = a_w + a_{w^*} \quad \forall w \in W_{2d}, \tag{5}$$

where A_w is the symmetric matrix defined by

$$(A_w)_{u,v} = \begin{cases} 2; & \text{if } u^*v \in \{w, w^*\},\ w^* = w, \\ 1; & \text{if } u^*v \in \{w, w^*\},\ w^* \neq w, \\ 0; & \text{otherwise.} \end{cases}$$

Note: $A_w = A_{w^*}$ for all w.

As we are interested in an arbitrary positive semidefinite $G = [G_{u,v}]_{u,v \in W}$ satisfying the constraints (5), we can choose the objective function freely. However, in practice one prefers solutions of small rank leading to shorter SOHS decompositions. Hence we minimize the trace, a commonly used heuristic for matrix rank minimization [RFP10]. Therefore our SDP in the primal form is as follows:

$$(\mathrm{SOHS_{SDP}}) \qquad \begin{array}{rrcl} \inf & \langle I, G\rangle & & \\ \text{s.t.} & \langle A_w, G\rangle & = & a_w + a_{w^*} \quad \forall w \in W_{2d} \\ & G & \succeq & 0. \end{array}$$

(Here and in the sequel, I denotes the identity matrix of appropriate size.) To reduce the size of this SDP (i.e., to make W_d smaller), we may employ the following simple observation:

Proposition 3.1. *Let $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle$, let $m_i := \frac{\operatorname{mindeg}_i f}{2}$, $M_i := \frac{\deg_i f}{2}$, $m := \frac{\operatorname{mindeg} f}{2}$, $M := \frac{\deg f}{2}$. Set*

$$V := \{w \in \langle \underline{X}\rangle \mid m_i \leq \deg_i w \leq M_i \text{ for all } i,\ m \leq \deg w \leq M\}.$$

Then $f \in \Sigma^2$ if and only if there exists a positive semidefinite matrix G satisfying $f = V^*GV$.

Proof. This follows from the fact that the highest or lowest degree terms in a SOHS decomposition cannot cancel. ■

Example 3.2 (Example 2.2 revisited)**.** Let us return to

$$f = 1 - 2X + 2X^2 + Y^2 - 2X^2Y - 2YX^2 + 2YXY + 2YX^2Y.$$

We shall describe in some detail ($\mathrm{SOHS_{SDP}}$) for f. From Proposition 3.1, we obtain

$$V = \begin{bmatrix} 1 & X & Y & XY & YX \end{bmatrix}^t.$$

Thus G is a symmetric 5×5 matrix and there will be 17 matrices A_w, as $|\{u^*v \mid u, v \in V\}| = 17$. In fact, there are only 13 different matrices A_w as $A_w = A_{w^*}$. Here is a sample:

$$A_{YX} = A_{XY} = \begin{bmatrix} 0&0&0&1&1\\ 0&0&1&0&0\\ 0&1&0&0&0\\ 1&0&0&0&0\\ 1&0&0&0&0 \end{bmatrix}, \quad A_{XY^2X} = \begin{bmatrix} 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&2 \end{bmatrix}.$$

These two give rise to the following linear constraints in ($\mathrm{SOHS_{SDP}}$):

$$\begin{array}{rcl} G_{1,XY} + G_{X,Y} + G_{XY,1} + G_{1,YX} + G_{Y,X} + G_{YX,1} & = & \langle A_{XY}, G\rangle \\ & = & a_{XY} + a_{YX} = 0, \\ 2G_{YX,YX} = \langle A_{XY^2X}, G\rangle & = & 2a_{XY^2X} = 0, \end{array}$$

where we have used a_w to denote the coefficients of f and the entries of V enumerate the columns, while the entries of V^* enumerate the rows of G. Observe that the second constraint tells us that the (YX, YX) entry of G is zero. As we are looking for a positive semidefinite G, the corresponding row and column of G can be assumed to be identically zero. That is, the last entry of V is redundant (cf. Example 2.2).

A further reduction in the vector of words needed is presented in [KP10] (the so-called *Newton chip method*) and its implementation in `NCSOStools` is `NCsos`.

4. Eigenvalue optimization of NC polynomials and flat extensions

One of the features of our freely available Matlab software package NCSOStools [CKP] is NCmin which uses sum of hermitian squares and semidefinite programming to compute a global (eigenvalue) minimum of a symmetric NC polynomial f. This is discussed in detail in [KP10, §5]. Here we present the theoretical underpinning of an algorithm to extract the minimizers of f, implemented in NCopt.

The main ingredients are the noncommutative moment problem and its solution due to McCullough [McC01], and the Curto-Fialkow theory [CF96] of how flatness governs the truncated moment problem. Our results are influenced by the method of Henrion and Lasserre [HL05] for the commutative case, which has been implemented in GloptiPoly [HLL09]. For an investigation of the non-global case in the free noncommutative setting see [PNA10].

4.1. Eigenvalue optimization is an SDP. Let $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle_{\le 2d}$. We are interested in the smallest eigenvalue $f^\star \in \mathbb{R}$ of the polynomial f. That is,

$$f^\star = \inf\big\{\langle f(A)v, v\rangle \mid A \text{ an } n\text{-tuple of symmetric matrices, } v \text{ a unit vector}\big\}. \tag{6}$$

Hence $f^\star$ is the greatest lower bound on the eigenvalues $f(A)$ can attain for n-tuples of symmetric matrices A, i.e., $(f - f^\star)(A) \succeq 0$ for all n-tuples of symmetric matrices A, and $f^\star$ is the largest real number with this property. Given that a polynomial is positive semidefinite if and only if it is a sum of hermitian squares (the Helton-McCullough SOHS theorem), we can compute $f^\star$ conveniently with SDP. Let

$$(\mathrm{SDP}_{\mathrm{eig-min}}) \qquad \begin{array}{rcl} f^{\mathrm{sohs}} & = & \sup\ \lambda \\ \text{s. t.} & & f - \lambda \ \in\ \Sigma^2. \end{array}$$

Then $f^{\mathrm{sohs}} = f^\star$.

In general ($\mathrm{SDP}_{\mathrm{eig-min}}$) does not satisfy the Slater condition. That is, there does not always exist a *strictly feasible* solution. Nevertheless ($\mathrm{SDP}_{\mathrm{eig-min}}$) satisfies strong duality [KP10, Theorem 5.1], i.e., its optimal value f^{sohs} coincides with the optimal value L_{sohs} of the dual SDP:

$$(\mathrm{DSDP}_{\mathrm{eig-min}})_d \qquad \begin{array}{rcl} L_{\mathrm{sohs}} & = & \inf L(f) \\ \text{s. t.} & & L : \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle_{\le 2d} \to \mathbb{R} \quad \text{is linear} \\ & & L(1) = 1 \\ & & L(p^*p) \ge 0 \quad \text{for all } p \in \mathbb{R}\langle \underline{X}\rangle_{\le d}. \end{array}$$

4.2. Extract the optimizers. In this section we investigate the attainability of $f^\star$ and explain how to extract the minimizers A, v for f if the lower bound $f^\star$ is attained. That is, A is an n-tuple of symmetric matrices and v is a unit eigenvector for $f(A)$ satisfying

$$f^\star = \langle f(A)v, v\rangle. \tag{7}$$

Of course, in general f will not be bounded from below. Another problem is that even if f is bounded, the infimum $f^\star$ need not be attained.

Example 4.1. Let $f = Y^2 + (XY-1)^*(XY-1)$. Clearly, $f^{\mathrm{sohs}} \geq 0$. However, $f(1/\varepsilon, \varepsilon) = \varepsilon^2$, so $f^{\mathrm{sohs}} = 0$ and hence $L_{\mathrm{sohs}} = 0$. On the other hand, $f^\star$ from (6) and the dual optimum L_{sohs} are not attained.

Let us first consider $f^\star$. Suppose (A, B) is a pair of matrices yielding a singular $f(A, B)$ and let v be a nullvector. Then

$$B^2 v = 0 \quad \text{and} \quad (AB - I)^*(AB - I)v = 0.$$

From the former we obtain $Bv = 0$, whence

$$v = Iv = (AB - I)v = 0,$$

a contradiction.

We now turn to the nonexistence of a dual optimizer. Suppose otherwise and let $L : \operatorname{Sym} \mathbb{R}\langle \underline{X} \rangle_{\leq 4} \to \mathbb{R}$ be a minimizer with $L(1) = 1$. We extend L to $\mathbb{R}\langle \underline{X} \rangle_{\leq 4}$ by symmetrization. That is,

$$L(p) := \frac{1}{2} L(p + p^*).$$

We note L induces a semi-scalar product (i.e., a positive semidefinite bilinear form) $(p, q) \mapsto L(p^* q)$ on $\mathbb{R}\langle \underline{X} \rangle_{\leq 2}$ due to the positivity property. Since $L(f) = 0$, we have

$$L(Y^2) = 0 \quad \text{and} \quad L\big((XY-1)^*(XY-1)\big) = 0.$$

Hence by the Cauchy-Schwarz inequality, $L(XY) = L(YX) = 0$. Thus

$$0 = L\big((XY-1)^*(XY-1)\big) = L\big((XY)^*(XY)\big) + L(1) \geq L(1) \geq 0,$$

implying $L(1) = 0$, a contradiction.

Hence despite the strong duality holding for ($\mathrm{SDP}_{\mathrm{eig-min}}$), the eigenvalue infimum $f^\star$ and the dual optimum L_{sohs} need not be attained, so some caution is necessary. In the sequel our main interest lies in the case where $f^\star$ is attained. We shall see later below (see Corollary 4.6) that this happens if and only it the infimum $L_{\mathrm{sohs}} = f^{\mathrm{sohs}} = f^\star$ for $(\mathrm{DSDP}_{\mathrm{eig-min}})_{d+1}$ is attained.

Definition 4.2. To each linear functional $L : \mathbb{R}\langle \underline{X} \rangle_{\leq 2d} \to \mathbb{R}$ we associate a matrix M_d (called an *NC Hankel matrix*) indexed by words $u, v \in \langle \underline{X} \rangle$ of length $\leq d$, with

$$(M_d)_{u,v} = L(u^* v). \tag{8}$$

If L is *positive*, i.e., $L(p^* p) \geq 0$ for all $p \in \mathbb{R}\langle \underline{X} \rangle_{\leq d}$, then M_d is positive semi-definite. We say that L is *unital* if $L(1) = 1$.

Note that a matrix M indexed by words of length $\leq d$ satisfying the *NC Hankel condition* $M_{u_1, v_1} = M_{u_2, v_2}$ if $u_1^* v_1 = u_2^* v_2$, yields a linear functional L on $\mathbb{R}\langle \underline{X} \rangle_{\leq 2d}$ as in (8). If M is positive semidefinite, then L is positive.

Definition 4.3. Let $A \in \mathbb{R}^{s \times s}$ be a symmetric matrix. A (symmetric) extension of A is a symmetric matrix $\tilde{A} \in \mathbb{R}^{(s+\ell) \times (s+\ell)}$ of the form

$$\tilde{A} = \begin{bmatrix} A & B \\ B^t & C \end{bmatrix}$$

for some $B \in \mathbb{R}^{s \times \ell}$ and $C \in \mathbb{R}^{\ell \times \ell}$. Such an extension is *flat* if $\operatorname{rank} A = \operatorname{rank} \tilde{A}$, or, equivalently, if $B = AZ$ and $C = Z^t A Z$ for some matrix Z.

Proposition 4.4. *Let $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle_{\le 2d}$ be bounded from below. If the infimum L_{sohs} for $(\mathrm{DSDP}_{\mathrm{eig-min}})_{d+1}$ is attained, then it is attained at a linear map L that is flat over its own restriction to $\mathbb{R}\langle \underline{X}\rangle_{\le 2d}$.*

Proof. For this proof it is beneficial to work with NC Hankel matrices. Let L be a minimizer for $(\mathrm{DSDP}_{\mathrm{eig-min}})_{d+1}$. To it we associate M_{d+1} and its restriction M_d. Then

$$M_{d+1} = \begin{bmatrix} M_d & B \\ B^t & C \end{bmatrix}$$

for some B, C. Since M_{d+1} and M_d are positive semidefinite, $B = M_d Z$ and $C \succeq Z^t M_d Z$ for some Z (this is easy to verify using Schur complements; or see [CF96]). Now form a "new" M_{d+1}:

$$\tilde{M}_{d+1} = \begin{bmatrix} M_d & B \\ B^t & Z^t M_d Z \end{bmatrix} = \begin{bmatrix} I & Z \end{bmatrix}^t M_d \begin{bmatrix} I & Z \end{bmatrix}.$$

This matrix is obviously flat over M_d, positive semidefinite, and satisfies the NC Hankel condition (it is inherited from M_{d+1} since for all quadruples u, v, z, w of words of degree $d+1$ we have $u^*v = z^*w$ if and only if $u = z$ and $z = w$). So it yields a positive linear map $\tilde{L}$ on $\mathbb{R}\langle \underline{X}\rangle_{\le 2d+2}$ flat over $\tilde{L}|_{\mathbb{R}\langle \underline{X}\rangle_{\le 2d}} = L|_{\mathbb{R}\langle \underline{X}\rangle_{\le 2d}}$. Moreover, $\tilde{L}(f) = L(f) = L_{\mathrm{sohs}}$. ∎

The following theorem is a solution to the free noncommutative moment problem in the truncated case. It resembles the classical results of Curto and Fialkow [CF96] in the commutative case. For the free noncommutative moment problem see [McC01] or also [PNA10]. A similar statement (with a positive definiteness assumption) is given in [MP05].

Theorem 4.5. *Suppose $L : \mathbb{R}\langle \underline{X}\rangle_{\le 2d+2} \to \mathbb{R}$ is positive and flat over $L|_{\mathbb{R}\langle \underline{X}\rangle_{\le 2d}}$. Then there is an n-tuple A of symmetric matrices of size $s \le \dim \mathbb{R}\langle \underline{X}\rangle_{\le d}$ and a vector v such that*

$$L(p^*q) = \langle p(A)v, q(A)v\rangle \tag{9}$$

for all $p, q \in \mathbb{R}\langle \underline{X}\rangle$ with $\deg p + \deg q \le 2d$.

Proof. For this we use the Gelfand-Naimark-Segal (GNS) construction. To L we associate two positive semidefinite matrices, M_{d+1} and its restriction M_d. Since M_{d+1} is flat over M_d, there exist s linear independent columns of M_d labeled by words $w \in \langle \underline{X}\rangle$ with $\deg w \le d$ which form a basis $\mathcal{B}$ of $E = \operatorname{range} M_{d+1}$. Now L (or M_{d+1}) induces a positive definite bilinear form (i.e., a scalar product) $\langle \_, \_ \rangle_E$ on E.

Let A_i be the left multiplication with X_i on E, i.e., if $\overline{w}$ denotes the column of M_{d+1} labeled by $w \in \langle \underline{X}\rangle_{\le d+1}$, then $A_i\colon \overline{u} \mapsto \overline{X_i u}$ for $u \in \langle \underline{X}\rangle_{\le d}$. The operator A_i is well defined and symmetric:

$$\langle A_i \overline{p}, \overline{q}\rangle_E = L(p^* X_i q) = \langle \overline{p}, A_i \overline{q}\rangle_E.$$

Let $v := \overline{1}$, and $A = (A_1, \ldots, A_n)$. Note it suffices to prove (9) for words $u, w \in \langle \underline{X}\rangle$ with $\deg u + \deg w \le 2d$. Since the A_i are symmetric, there is no harm in assuming $\deg u, \deg w \le d$. Now compute

$$L(u^*w) = \langle \overline{u}, \overline{w}\rangle_E = \langle u(A)\overline{1}, w(A)\overline{1}\rangle_E = \langle u(A)v, w(A)v\rangle_E.$$ ∎

Corollary 4.6. *Let $f \in \mathbb{R}\langle \underline{X}\rangle_{\leq 2d}$. Then $f^\star$ is attained if and only if there is a feasible point L for* $(\mathrm{DSDP}_{\mathrm{eig-min}})_{d+1}$ *satisfying $L(f) = f^\star$.*

Proof. ($\Rightarrow$) If (7) holds for some A, v, then $L(p) := \langle p(A)v, v\rangle$ is the desired feasible point. ($\Leftarrow$) By Proposition 4.4, we may assume L is flat over $L|_{\mathbb{R}\langle \underline{X}\rangle_{\leq 2d}}$. Now Theorem 4.5 applies and yields A, v. By definition, $\|v\| = \sqrt{\langle v, v\rangle} = \sqrt{L(1)} = 1$. Hence $f(A)$ has (unit) eigenvector v with eigenvalue $f^\star$. ∎

4.3. **Implementing the extraction of optimizers.** Let $f \in \operatorname{Sym}\mathbb{R}\langle \underline{X}\rangle_{\leq 2d}$.

Step 1: Solve $(\mathrm{DSDP}_{\mathrm{eig-min}})_{d+1}$. If the problem is unbounded or the optimum is not attained, stop. Otherwise let L denote an optimizer.

Step 2: To L we associate the positive semidefinite matrix $M_{d+1} = \begin{bmatrix} M_d & B \\ B^t & C \end{bmatrix}$.

Modify M_{d+1}: $\tilde{M}_{d+1} = \begin{bmatrix} M_d & B \\ B^t & Z^t M_d Z \end{bmatrix}$, where Z satisfies $M_d Z = B$.

This matrix yields a positive linear map $\tilde{L}$ on $\mathbb{R}\langle \underline{X}\rangle_{\leq 2d+2}$ which is flat over $\tilde{L}|_{\mathbb{R}\langle \underline{X}\rangle_{\leq 2d}} = L|_{\mathbb{R}\langle \underline{X}\rangle_{\leq 2d}}$. In particular, $\tilde{L}(f) = L(f) = f^\star$.

Step 3: As in the proof of Theorem 4.5, use the GNS construction on $\tilde{L}$ to compute symmetric matrices A_i and a vector v with $\tilde{L}(f) = f^\star = \langle f(A)v, v\rangle$.

In Step 3, to construct symmetric matrix representations $A_i \in \mathbb{R}^{s\times s}$ of the multiplication operators we calculate their image according to a chosen basis $\mathcal{B}$ for $E = \operatorname{range}\tilde{M}_{d+1}$. To be more specific, $A_i\overline{u}_1$ for $u_1 \in \langle \underline{X}\rangle_{\leq d}$ being the first label in $\mathcal{B}$, can be written as a unique linear combination $\sum_{j=1}^{s} \lambda_j \overline{u}_j$ with words u_j labeling $\mathcal{B}$ such that $L\big((u_1X_i - \sum \lambda_j u_j)^*(u_1X_i - \sum \lambda_j u_j)\big) = 0$. Then $\begin{bmatrix}\lambda_1 & \dots & \lambda_s\end{bmatrix}^t$ will be the first column of A_i. The vector v is the eigenvector of $f(A)$ corresponding to the smallest eigenvalue.

Warning 4.7. Running the above algorithm raises several challenges in practice. Since the primal problem ($\mathrm{SDP}_{\mathrm{eig-min}}$) often has no strictly feasible point we have no guarantee that the optimal value L_{sohs} of $(\mathrm{DSDP}_{\mathrm{eig-min}})_{d+1}$ is attained. We do not know how to test for attainability efficiently, since all state-of-the-art SDP solvers return only an ε-optimal solution (a point which is feasible and gives optimal value up to some rounding error).

Detecting unboundedness of $(\mathrm{DSDP}_{\mathrm{eig-min}})_{d+1}$ seems easier. First of all, the SDP solver is likely to detect it directly. Otherwise numerical problems will be mentioned, and we then solve the (usually much smaller) primal problem ($\mathrm{SDP}_{\mathrm{eig-min}}$) to detect its infeasibility, which is equivalent to the unboundedness of $(\mathrm{DSDP}_{\mathrm{eig-min}})_{d+1}$.

In summary, the performance of our algorithm to extract the optimizers depends heavily on the quality of the underlying SDP solver.

Remark 4.8. We finish this section by emphasizing that the extraction of eigenvalue optimizers (theoretically) *always* works if the optimum for $(\mathrm{DSDP}_{\mathrm{eig-min}})_{d+1}$ is attained. This is in sharp contrast with the commutative case; cf. [Las09].

4.4. **Example for eigenvalue minimization.** In this subsection we present a toy example of eigenvalue optimization as presented in Section 4. The numerical results were obtained using our open source software `NCSOStools`, developed by the authors of this paper.

Example 4.9. Let us introduce NC variables x, y by

```
>> NCvars x y;
```

and an NC polynomial

```
>> f = (1-3*x*y+y*x)'*(1-3*x*y+y*x) + (x^2-1)^2 + (y^2-y)^2;
```

As is usual in Matlab, the prime ' denotes an involution, in our case acting on NC polynomials. By definition, f is a sum of hermitian squares. We shall compute the eigenvalue minimum $f^\star$ of f and determine the minimizers A, B, v satisfying $\langle f(A,B)v, v\rangle = f^\star$. Here A, B are symmetric matrices, and v is a unit eigenvector of $f(A,B)$, corresponding to the eigenvalue $\lambda_{\min}(f(A,B))$. Running

```
>> NCmin(f)
```

yields an eigenvalue minimum $f^\star = 0.0000$. We next run the algorithm presented in Subsection 4.3 to extract optimizers:

```
>> [X,fX,eig_val,eig_vec]=NCopt(f)
```

The output: `X` is a 2×16 matrix, whose rows represent symmetric matrices A, B; `fX` is the 4×4 matrix $f(A,B)$; `eig_val` are the eigenvalues of `fX`, and `eig_vec` are the corresponding unit eigenvectors. In our example,

$$A = \begin{bmatrix} 0.9644 & -0.0379 & -0.1276 & 0.0879 \\ -0.0379 & -0.9828 & 0.1588 & 0.0235 \\ -0.1276 & 0.1588 & 0.4923 & 0.2253 \\ 0.0879 & 0.0235 & 0.2253 & -0.9790 \end{bmatrix}$$

$$B = \begin{bmatrix} 0.8367 & 0.1790 & 0.3326 & 0.0832 \\ 0.1790 & 0.0215 & 0.1388 & 0.5320 \\ 0.3326 & 0.1388 & -0.0227 & -0.6871 \\ 0.0832 & 0.5320 & -0.6871 & -0.1778 \end{bmatrix}$$

$$f(A,B) = \begin{bmatrix} 0.7978 & 1.2130 & 0.8094 & 0.6920 \\ 1.2130 & 3.3989 & -2.6498 & -0.0064 \\ 0.8094 & -2.6498 & 10.5185 & 3.0781 \\ 0.6920 & -0.0064 & 3.0781 & 7.9733 \end{bmatrix}$$

and the smallest eigenvalue $f^\star$ of $f(A,B)$ is (only 4 decimal digits displayed) 0.0000, with the corresponding unit eigenvector

$$v = \begin{bmatrix} -0.8741 & 0.4515 & 0.1789 & 0.0072 \end{bmatrix}^t.$$

We note the minimum of f on $\mathbb{R}^2$ can be computed exactly using Mathematica. It is 0.0625.

Conclusions

In the paper we presented a sums of hermitian squares approach to detect positivity of given non-commutative polynomials with a special focus on computing lower bounds for eigenvalue minimization of such polynomials, based on semidefinite programming. To test for exactness of bounds we investigated the solution of the dual semidefinite program. We showed that attainability of the optimal value on the dual side implies that the eigenvalue bound is attained. In such a case we can extract the global eigenvalue optimizers. The procedure, based on solution of the truncated free moment problem and the GNS construction is presented and demonstrated with an example.

References

[CF96] R.E. Curto and L.A. Fialkow. Solution of the truncated complex moment problem for flat data. *Mem. Amer. Math. Soc.*, 119(568):x+52, 1996.

[CKP] K. Cafuta, I. Klep, and J. Povh. NCSOStools: a computer algebra system for symbolic and numerical computation with noncommutative polynomials. To appear in Optim. Meth. Software, available at http://ncsostools.fis.unm.si.

[CKP10] K. Cafuta, I. Klep, and J. Povh. A note on the nonexistence of sum of squares certificates for the Bessis-Moussa-Villani conjecture. *J. math. phys.*, 51(8), 2010.

[CLR95] M.D. Choi, T.Y. Lam, and B. Reznick. Sums of squares of real polynomials. In *K-theory and algebraic geometry: connections with quadratic forms and division algebras*, volume 58 of *Proc. Sympos. Pure Math.*, pages 103–126. AMS, Providence, RI, 1995.

[dOHMP08] M.C. de Oliveira, J.W. Helton, S. McCullough, and M. Putinar. Engineering systems and free semi-algebraic geometry. In *Emerging Applications of Algebraic Geometry*, volume 149 of *IMA Vol. Math. Appl.*, pages 17–62. Springer, 2008.

[HdOMS] J.W. Helton, M. de Oliveira, R.L. Miller, and M. Stankus. NCAlgebra: A Mathematica package for doing non commuting algebra. http://www.math.ucsd.edu/~ncalg/.

[Hel02] J.W. Helton. "Positive" noncommutative polynomials are sums of squares. *Ann. of Math. (2)*, 156(2):675–694, 2002.

[HL05] D. Henrion and J.-B. Lasserre. Detecting global optimality and extracting solutions in GloptiPoly. In *Positive polynomials in control*, volume 312 of *Lecture Notes in Control and Inform. Sci.*, pages 293–310. Springer, Berlin, 2005.

[HLL09] D. Henrion, J.-B. Lasserre, and J. Löfberg. GloptiPoly 3: moments, optimization and semidefinite programming. *Optim. Methods Softw.*, 24(4-5):761–779, 2009. http://www.laas.fr/~henrion/software/gloptipoly3/.

[KP10] I. Klep and J. Povh. Semidefinite programming and sums of hermitian squares of noncommutative polynomials. *J. Pure Appl. Algebra*, 214:740–749, 2010.

[KS08a] I. Klep and M. Schweighofer. Connes' embedding conjecture and sums of Hermitian squares. *Adv. Math.*, 217(4):1816–1837, 2008.

[KS08b] I. Klep and M. Schweighofer. Sums of Hermitian squares and the BMV conjecture. *J. Stat. Phys*, 133(4):739–760, 2008.

[Las09] J.B. Lasserre. *Moments, Positive Polynomials and Their Applications*, volume 1. Imperial College Press, 2009.

[Löf04] J. Löfberg. YALMIP: A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004. http://control.ee.ethz.ch/~joloef/yalmip.php.

[McC01] S. McCullough. Factorization of operator-valued polynomials in several non-commuting variables. *Linear Algebra Appl.*, 326(1-3):193–203, 2001.

[MP05] S. McCullough and M. Putinar. Noncommutative sums of squares. *Pacific J. Math.*, 218(1):167–171, 2005.

[Par03] P.A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Math. Program.*, 96(2, Ser. B):293–320, 2003.

[PNA10] S. Pironio, M. Navascues, and A. Acin. Convergent relaxations of polynomial optimization problems with non-commuting variables. *SIAM J. Optim.*, 20(5):2157–2180, 2010.

[PPSP05] S. Prajna, A. Papachristodoulou, P. Seiler, and P.A. Parrilo. SOSTOOLS and its control applications. In *Positive polynomials in control*, volume 312 of *Lecture Notes in Control and Inform. Sci.*, pages 273–292. Springer, Berlin, 2005. http://www.cds.caltech.edu/sostools/.

[PW98] V. Powers and T. Wörmann. An algorithm for sums of squares of real polynomials. *J. Pure Appl. Algebra*, 127(1):99–104, 1998.

[RFP10] B. Recht, M. Fazel, and P.A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Rev.*, 52(3):471–501, 2010.

[WKK+09] H. Waki, S. Kim, M. Kojima, M. Muramatsu, and H. Sugimoto. Algorithm 883: sparsePOP—a sparse semidefinite programming relaxation of polynomial optimization problems. *ACM Trans. Math. Software*, 35(2):Art. 15, 13, 2009.

[WSV00] H. Wolkowicz, R. Saigal, and L. Vandenberghe. *Handbook of Semidefinite Programming*. Kluwer, 2000.