Asymptotically Exact Relaxations for Robust LMI Problems based on Matrix-Valued Sum-of-Squares

C.W. Scherer* and C.W.J. Hol†
Delft Center for Systems and Control (DCSC)
Delft University of Technology
Mekelweg 2, 2628 CD Delft, The Netherlands
c.w.scherer@dcsc.tudelft.nl

Abstract

In this paper we consider the problem of characterizing whether a symmetric polynomial matrix is positive definite on a semi-algebraic set. Based on suitable sum-of-squares representations we can construct LMI relaxation for this decision problem. As key novel technical contributions it is possible to prove that these relaxations are exact. Our proof is based on a sum-of-squares representation of $r^2 - \|x\|^2$ with respect to affine functions with a priori constraints on the degree. This is a nontrivial extension of a rather deep result from [7] obtained by semi-definite duality arguments.

1 Introduction

The sum-of-squares (sos) approach has recently been identified as a crucial technique to address various non-convex optimization problems in control. In particular it is well-understood how to construct LMI relaxations of scalar polynomially constraint polynomial optimization problems [2, 9, 8, 4]. The underlying mathematical questions can be stated as follows: Given scalar polynomial functions $F(x)$ and $g_1(x), \ldots, g_m(x)$, decide whether $F(x)$ is positive on the semi-algebraic set $G = \{x \in \mathbb{R}^n : \; g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$. However, in particular in robust controller analysis and synthesis problems, the function $F(x)$ turns out to be matrix-valued, and one is rather interested in the question of whether $F(x)$ is positive definite on the set G. Even if $F(x)$ itself and the constraint functions $g_i(x)$ are affine, the resulting decision problem turns out to be NP-complete (in a sense as precisely specified in [1]). This motivates to construct efficient relaxations for this decision problem that are based on semi-definite programming.

*The research of this author is supported by the Technology Foundation STW, applied science division of NWO and the technology programme of the Ministry of Economic Affairs.
†Sponsored by Philips CFT.

As an example consider the worst-norm minimization problem

$$t_{\rm opt} = \inf_{y\in\mathbb{R}^m} \sup_{x\in G} \|L(x,y)\| \tag{1}$$

with some mapping $L : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^{p\times q}$ such that $L(x, y)$ is affine in the design variable y (for all $x \in \mathbb{R}^n$) and $L(x, y)$ is rational in the uncertainty variable x (for all $y \in \mathbb{R}^m$). It turns out that surprisingly many specific questions can be translated into this problem formulation [3, 1, 11]. Clearly $t_{\rm opt}$ equals the infimal γ for which there exists $y \in \mathbb{R}^m$ such that

$$M_\gamma(x,y) := \begin{pmatrix} \gamma I & L(x,y)^T \\ L(x,y) & \gamma I \end{pmatrix} \succ 0 \ \text{ for all } \ x \in G. \tag{2}$$

The full block S-procedure [6, 10] allows to construct various relaxations which compute upper bounds on $t_{\rm opt}$. If G is a compact polytope it has been proposed only recently how one can actually determine a whole family of such relaxations whose optimal values are guaranteed to converge to $t_{\rm opt}$ [11]. This procedure suffers from the disadvantage that it requires an explicit description of extreme points (generators) of G. As one of the key contributions of this paper, we show how one can work with the implicit (inequality) description of G while retaining the possibility to construct asymptotically exact relaxation families.

As a viable alternative one might replace (2) by the following scalar constraint, at the expense of increasing the dimension of the uncertainty variable, and apply scalar-valued relaxation techniques such as in [8]:

$$v^T M_\gamma(x,y) v \succ 0 \ \text{ for all } \ x \in G, \ 1 \le v^T v \le 2.$$

In order to guarantee exactness of the approximating relaxation family, however, it is required to use as well higher-order polynomials in the new variable vector v. The results in this paper might be interpreted as allowing a relaxation with an a priori constraint on the degree in v. How the two approaches actually compare with respect to their computational complexity for specific practical problems is a very interesting topic for future research.

The paper is structured as follows. If $g_1(x), \ldots, g_m(x)$ are affine, we show in section 2 that $r^2 - \|x\|^2$ can be represented as $s_0(x) + s_1(x)g_1(x) + \cdots + s_m(x)g_m(x)$ with *quadratic* sos polynomials $s_i(x)$, $i = 0, 1, \ldots, m$. This extends a rather deep result from [7] since no explicit bounds on the degrees of the sos polynomials have been obtained so far in the literature. Although tricky, our proof is elementary and based on SDP-duality arguments. This fact will be used in Section 3 to prove our second main result, namely that any matrix-valued polynomial $F(x)$ which is positive definite on G can be represented as $S_0(x) + S_1(x)g_1(x) + \cdots + S_m(x)g_m(x)$ with sos matrices $S_i(x)$, $i = 0, 1, \ldots, m$. Although well-known for scalar polynomials, the extension to matrix-valued polynomials is new, with a proof following arguments in [12]. Finally we will reveal how this result allows to construct novel relaxations for robust LMI problems in Section 4 for implicitly described uncertainty sets G. We will distinguish two different approaches to construct asymptotically exact relaxations of (1), one being based on a direct sos representation of $M_\gamma(x, y)$ and an alternative one based on the S-procedure.

Notation. $\langle A, B\rangle := \text{Trace}(A^T B)$ denotes the standard matrix inner product and the symbols $\succ$, $\succeq$, $\preceq$, $\prec$ are used for the Loewner partial ordering on the set of symmetric matrices. e_j and e denote the standard unit vectors and the all-ones vector in $\mathbb{R}^n$. We denote the standard simplex by $\Sigma := \{x \in \mathbb{R}^n : \ x \geq 0, \ e^T x \leq 1\}$. For any multi-index $\alpha \in \mathbb{N}_0$ we use the abbreviations $|\alpha| = e^T\alpha$ and $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$.

2 SOS representations of $r^2 - \|x\|^2$ with degree bounds

Let us assume in this section that the functions $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, are affine and that the polytope

$$G = \{x \in \mathbb{R}^n : \ g_i(x) \geq 0, \ i = 1, \ldots, m\} \tag{3}$$

is compact. The following result will be instrumental in the next section.

Theorem 1 *There exist $r \in \mathbb{R}$ and sos polynomials $s_0, s_1, \ldots, s_m$ of degree two such that*

$$r^2 - \|x\|^2 = s_0(x) + s_1(x)g_1(x) + \cdots + s_m(x)g_m(x). \tag{4}$$

This is in fact a consequence of rather deep results from real algebraic geometry [7]. To the best of our knowledge so far nothing has been known about a priori bounds on the degrees of $s_0, s_1, \ldots, s_m$. We can provide a full proof for this representation with degree bounds based on elementary (though tricky) SDP-duality arguments.

Proof. Let us first assume $g_i(x) = x_i$, $i = 1, \ldots, n$, and $g_{n+1}(x) = 1 - e^T x$. Let us abbreviate $z(x) = \mathrm{col}(1, x)$. Then any quadratic function can be represented as $z(x)^T Y z(x)$ with some symmetric matrix Y, and this function is sos iff $Y \succeq 0$. In particular, with $r^2 - \|x\|^2 = z(x)^T R z(x)$ for $R = \mathrm{diag}(r^2, -1, \cdots, -1)$, our goal is to prove that there exist $r > 0$ and $Y_i \succ 0$, $i = 0, \ldots, n$, with

$$\begin{aligned} z(x)^T R z(x) = r^2 - \|x\|^2 = z(x)^T Y_0 z(x)(1 - e^T x) + \sum_{i=1}^n z(x)^T Y_i z(x) x_i = \\ = \langle z(x)z(x)^T, Y_0\rangle + \sum_{i=1}^n \langle x_i z(x) z(x)^T, Y_i - Y_0\rangle. \end{aligned} \tag{5}$$

(Note that this is (4) with $s_0(x) = 0$). If $w_1(x) = 1$, and if $w_j(x)$, $j = 2, \ldots, k$ is the list of all pairwise different monomials in n variables of degree at least one and at most three (in any ordering) we can determine the representations

$$z(x)z(x)^T = \sum_{j=1}^k Q_j^0 w_j(x), \ \ x_i z(x) z(x)^T = \sum_{j=1}^k Q_j^i w_j(x)$$

to infer $z(x)^T R z(x) = \langle z(x)z(x)^T, R\rangle = \sum_{j=1}^k \langle Q_j^0, R\rangle w_j(x)$. Our problem reduces to proving the existence of positive definite $Y_0, Y_1, \ldots, Y_n$ with

$$\langle Q_j^0, R\rangle = \langle Q_j^0, Y_0\rangle + \sum_{i=1}^n \langle Q_j^i, Y_i - Y_0\rangle, \quad j = 1, \ldots, k. \tag{6}$$

There exist positive definite solutions of (6) iff the infimum of all t with $Y_i + tI \succeq 0$, $i = 0, \ldots, n$, and with (6) is negative. Since this programm is strictly feasible (choose

$Y_i = R$, $i = 0, \ldots, n$, and t large), we can dualize without gap. With multipliers $\Gamma_i \succeq 0$, $i = 0, \ldots, n$, and γ_j, $j = 1, \ldots, k$, the Lagrangian reads as

$$t + \sum_{j=1}^{k} \gamma_j \left(\langle Q_j^0, Y_0 \rangle + \sum_{i=1}^{n} \langle Q_j^i, Y_i - Y_0 \rangle - \langle Q_j^0, R \rangle \right) - \sum_{i=0}^{n} \langle \Gamma_i, Y_i + tI \rangle =$$

$$= t(1 - \langle \sum_{i=0}^{n} \Gamma_i, I \rangle) + \langle \sum_{j=1}^{k} \gamma_j [Q_j^0 - \sum_{i=1}^{n} Q_j^i] - \Gamma_0, Y_0 \rangle + \sum_{i=1}^{n} \langle \sum_{j=1}^{k} \gamma_j Q_j^i - \Gamma_i, Y_i \rangle - \langle \sum_{j=1}^{k} \gamma_j Q_j^0, R \rangle.$$

Therefore there exist positive definite solutions of (6) iff

$$\max_{\Gamma_i \succeq 0,\, 1 - \langle \sum_{i=0}^{n} \Gamma_i, I \rangle = 0,\, \sum_{j=1}^{k} \gamma_j Q_j^0 - \sum_{i=1}^{n} \sum_{j=1}^{k} \gamma_j Q_j^i - \Gamma_0 = 0,\, \sum_{j=1}^{k} \gamma_j Q_j^i - \Gamma_i = 0,\, i=1,\ldots,n} - \langle \sum_{j=1}^{k} \gamma_j Q_j^0, R \rangle < 0.$$

With the abbreviations $M_i^n(\gamma) := \sum_{j=1}^{k} \gamma_j Q_j^i$, $i = 0, \ldots, n$, this is equivalent to

$$M_i^n(\gamma) \succeq 0,\ i = 1, \ldots, n,\ M_0^n(\gamma) - \sum_{i=1}^{n} M_i^n(\gamma) \succeq 0,\ \langle M_0^n(\gamma), I \rangle = 1 \quad \Rightarrow$$

$$\Rightarrow \quad \langle M_0^n(\gamma), R \rangle > 0. \quad (7)$$

We have indicated the number of variables n in this definition since it is essential to exploit the structural relation of $M_i^{n-1}(\gamma)$ and $M_i^n(\gamma)$ that is easily identified as follows: If

$$M_n^n(\gamma) = \begin{pmatrix} c_0 & \cdots & c_{n-1} & c_n \\ d_0 & \cdots & d_{n-1} & d_n \end{pmatrix} \in \mathbb{R}^{(n+1) \times (n+1)}$$

for suitable $\mathbb{R}^n$-vectors c_i and scalars d_i then

$$M_i^n(\gamma) = \begin{pmatrix} M_i^{n-1}(\gamma) & c_i \\ c_i^T & d_i \end{pmatrix} \in \mathbb{R}^{(n+1) \times (n+1)} \quad \text{for} \quad i = 0, \ldots, n-1. \quad (8)$$

With the all-ones vector e and the standard unit vectors e_i let us define the length $(n+1)$-vectors

$$f_i = e - e_{i+1}, \quad g_i = e - e_1 - e_{i+1} = f_0 - e_{i+1}, \quad i = 0, \ldots, n$$

as well as the scalar

$$s^n := e^T \left(M_0^n(\gamma) - \sum_{i=1}^{n} M_i^n(\gamma) \right) e + 0.5 \left(\sum_{i=1}^{n} f_i^T M_i^n(\gamma) f_i + \sum_{i=1}^{n} g_i^T M_i^n(\gamma) g_i \right).$$

As an essential ingredient of our proof we exploit that $s^n \geq 0$ if the hypothesis in (7) is satisfied. Moreover the specific vectors in the definition of s^n are chosen to be able to obtain the following recursion. Indeed with the corresponding vectors e, e_i, f_i, g_i of length n we have, due to (8),

$$s^n = e^T M_0^{n-1}(\gamma) e + 2c_0^T e + d_0 - \sum_{i=1}^{n-1} (e^T M_i^{n-1}(\gamma) e + 2c_i^T e + d_i) - \sum_{i=0}^{n} (c_i^T e + d_i) +$$

$$+0.5 \left(\sum_{i=1}^{n-1} (f_i^T M_i^{n-1}(\gamma) f_i + 2c_i^T f_i + d_i) + (g_i^T M_i^{n-1}(\gamma) g_i + 2c_i^T g_i + d_i) + \sum_{i=0}^{n-1} c_i^T e + \sum_{i=1}^{n-1} c_i^T f_0 \right)$$

which simplifies to

$$s^n = s^{n-1} + 2c_0^T e + d_0 - \sum_{i=1}^{n-1}(3c_i^T e + 2d_i) - (c_0^T e + d_0) - (c_n^T e + d_n) +$$
$$+ \sum_{i=1}^{n-1}(c_i^T f_i + c_i^T g_i + d_i) + 0.5\left(\sum_{i=1}^{n-1} c_i^T e + c_i^T f_0\right) + 0.5c_0^T e.$$

Since $c_n^T e = d_0 + \cdots + d_{n-1}$ (just because $M_n^n(\gamma)$ is symmetric) we infer

$$s^n = s^{n-1} + 1.5c_0^T e - \sum_{i=1}^{n-1} 2d_i - (d_0 + d_n) + \sum_{i=1}^{n-1} c_i^T (f_i + g_i + 0.5f_0 - 2.5e).$$

With $f_i + g_i + 0.5f_0 = (e - e_{i+1}) + (e - e_1 - e_{i+1}) + 0.5(e - e_1) = 2.5e - 2e_{i+1} - 1.5e_1$ and $c_0^T e - c_1^T e_1 - \cdots - c_{n-1}^T e_1 = c_0^T e_1$ (again since $M_n^n(\gamma)$ is symmetric) we obtain

$$s^n = s^{n-1} - d_0 - d_n + 1.5c_0^T e_1 - \sum_{i=1}^{n-1} 2d_i - 2\sum_{i=1}^{n-1} c_i^T e_{i+1}.$$

Let us finally exploit

$$d_i = e_{n+1}^T M_i^n(\gamma) e_{n+1},\ i = 0,\ldots,n,\quad c_i^T e_{i+1} = e_{i+1}^T M_n^n(\gamma) e_{i+1},\quad i = 1,\ldots,n-1,$$
$$2c_0^T e_1 = e_1^T M_0^{n-1}(\gamma) e_1 + e_{n+1}^T M_0^n(\gamma) e_{n+1} - (e_1 - e_{n+1})^T M_0^n(\gamma)(e_1 - e_{n+1})$$

to end up with

$$s^n = s^{n-1} + 0.75e_1^T M_0^{n-1}(\gamma) e_1 - 0.25e_{n+1}^T M_0^n(\gamma) e_{n+1} - 0.75(e_1 - e_{n+1})^T M_0^n(\gamma)(e_1 - e_{n+1}) -$$
$$- e_{n+1}^T M_n^n(\gamma) e_{n+1} - 2\sum_{i=1}^{n-1} e_{n+1}^T M_i^n(\gamma) e_{n+1} - 2\sum_{i=1}^{n-1} e_{i+1}^T M_n^n(\gamma) e_{i+1}.$$

Due to the choice $w_1(x) = 1$ we have $e_1^T M_0^\nu(\gamma) e_1 = \gamma_1$ for all $\nu = 1, 2, \ldots, n$. Hence for $n = 1$ we conclude

$$M_1^1(\gamma) = \begin{pmatrix} c_0 & c_1 \\ d_0 & d_1 \end{pmatrix} = \begin{pmatrix} c_0 & d_0 \\ d_0 & d_1 \end{pmatrix} \quad \text{and} \quad M_0^1(\gamma) = \begin{pmatrix} \gamma_1 & c_0 \\ c_0^T & d_0 \end{pmatrix}$$

and thus

$$s^1 = (\gamma_1 - c_0) + 2(c_0 - d_0) + (d_0 - d_1) + 0.5c_0 =$$
$$= 1.75e_1^T M_0^1(\gamma) e_1 - 0.25e_2^T M_0^1(\gamma) e_2 - e_2^T M_1^1(\gamma) e_2 - 0.75(e_1 - e_2)^T M_0^1(\gamma)(e_1 - e_2).$$

Let us now assume that the hypothesis in (7) is satisfied. Just due to (8) we can conclude $M_i^\nu(\gamma) \succeq 0$ for $i = 0, \ldots, \nu$ and $\nu = n-1, n-2, \ldots, 1$ (by induction). This implies

$$s^1 \le 1.75e_1^T M_0^1(\gamma) e_1 - 0.25e_2^T M_0^1(\gamma) e_2,$$
$$0 \le s^\nu \le s^{\nu-1} + 0.75e_1^T M_0^{\nu-1}(\gamma) e_1 - 0.25e_{\nu+1}^T M_0^\nu(\gamma) e_{\nu+1},\quad \nu = 2, \ldots, n.$$

Since $e_1^T M_0^{\nu-1}(\gamma)e_1 = e_1^T M_0^n(\gamma)e_1$ and $e_{\nu+1}^T M_0^{\nu}(\gamma)e_{\nu+1} = e_{\nu+1}^T M_0^n(\gamma)e_{\nu+1}$ for $\nu = 2, \ldots, n$, we conclude by induction that

$$(1 + 0.75n)e_1^T M_0^n(\gamma)e_1 - 0.25\sum_{\nu=1}^{n} e_{\nu+1}^T M_0^n(\gamma)e_{\nu+1} \geq s^n \geq 0$$

and consequently, for $r = \sqrt{5 + 3n}$,

$$\langle R, M_0^n(\gamma)\rangle = \langle e_1e_1^T, M_0^n(\gamma)\rangle + \underbrace{\langle (4 + 3n)e_1e_1^T - \sum_{\nu=2}^{n+1} e_{\nu+1}e_{\nu+1}^T, M_0^n(\gamma)\rangle}_{\geq 0} \geq 0.$$

If $\langle R, M_0^n(\gamma)\rangle = 0$ we infer $0 = \langle e_1e_1^T, M_0^n(\gamma)\rangle = e_1^T M_0^n(\gamma)e_1$. This implies $e_1^T M_0^n(\gamma)e_{\nu+1} = 0$ and thus $e_1^T M_\nu^n(\gamma)e_1 = 0$ and thus $e_1^T M_\nu^n(\gamma)e_{\nu+1} = 0$ and thus $e_{\nu+1}^T M_0^n(\gamma)e_{\nu+1} = 0$ and thus $M_0^n(\gamma)e_{\nu+1} = 0$ for $\nu = 2, \ldots, n$. Therefore $M_0^n(\gamma) = 0$ in contradiction to $\langle M_0^n(\gamma), I\rangle = 1$. This proves the implication (7).

Therefore there do indeed exist $r > 0$ and $Y_i \succ 0$, $i = 0, \ldots, n$, with (5). Let us now turn to general affine functions $g_i(x)$, $i = 1, \ldots, m$. If $G \subset \Sigma$ then by LP duality (Farkas) there exist $v_{ij} \geq 0$, $i = 0, \ldots, n$, $j = 1, \ldots, m$, such that

$$1 - e^T x = \sum_{j=1}^{m} v_{0j}g_j(x), \quad x_i = \sum_{j=1}^{m} v_{ij}g_j(x), \quad i = 1, \ldots, n$$

which leads to

$$r^2 - \|x\|^2 = \sum_{i=0}^{n} z(x)^T Y_i z(x) \sum_{j=1}^{m} v_{ij}g_j(x) = \sum_{j=1}^{m} \left[z(x)^T \left(\sum_{i=0}^{n} Y_i v_{ij} \right) z(x) \right] g_j(x),$$

and we obtain the desired representation.

Finally if $G \not\subset \Sigma$ it is simple to find $\alpha > 0$ and a vector c such that $(G - c)/\alpha = \{y : g_j(\alpha y + c) \geq 0,\ j = 1, \ldots, m\} \subset \Sigma$ (due to compactness). Then there exists β and sos polynomials s_j with $\beta^2 - \|y\|^2 = s_1(y)g_1(\alpha y + c) + \cdots + s_m(y)g_m(\alpha y + c)$ which implies

$$\beta^2 - \|(x - c)/\alpha\|^2 = s_1((x - c)/\alpha)g_1(x) + \cdots + s_m((x - c)/\alpha)g_m(x).$$

It remains to observe that one can easily find scalars r, $t > 0$ such that $r^2 - \|x\|^2 - t(\beta^2 - \|(x - c)/\alpha\|^2) =: s_0(x)$ is non-negative and hence sos. This implies

$$\begin{aligned} r^2 - \|x\|^2 = s_0(x) + t(\beta^2 - \|(x - c)/\alpha\|^2) = \\ = s_0(x) + [ts_1((x - c)/\alpha)]g_1(x) + \cdots + [ts_m((x - c)/\alpha)]g_m(x) \end{aligned}$$

which is the desired representation in the general case.

■

3 SOS representations of polynomial matrices

We follow Schweighofer [12] and assume that the following constraint qualification holds for the description of the semi-algebraic set G: there exists $r > 0$ and sos polynomials

s_j, $j = 0, \ldots, m$ with (4). Theorem 1 just says that this property is true if G is compact and described with affine functions. Moreover we say that the polynomial matrix $S(x)$ is a sum-of-squares if there exists a polynomial matrix $T(x)$ with $S(x) = T(x)^T T(x)$. A full discussion of this concept together with the relation to LMI's is given in the accompanying paper [5].

Theorem 2 *If $F : \mathbb{R}^n \to \mathbb{R}^{k \times k}$ takes symmetric values and satisfies $F(x) \succ 0$ for all $x \in G$, then there there exist sos polynomial matrices $S_0(x), S_1(x), \ldots, S_m(x)$ such that*

$$F(x) = S_0(x) + S_1(x)g_1(x) + \cdots + S_m(x)g_m(x). \tag{9}$$

The remainder of this section is devoted to the proof of Theorem 2 which will be based on a polynomial-matrix version of Pólya's theorem in combination with a nice penalty technique suggested by Schweighofer [12]. The construction is simplified if assuming that a shifted and scaled version of G is contained in some ball B that itself is contained in the standard simple Σ. More specifically we consider the ball $B := \{x \in \mathbb{R}^n : \ \|x/\rho - e\| \leq 1\}$. It is not difficult to verify that the largest ρ for which $B \subset \Sigma$ is given by $\rho = 1/(n + \sqrt{n})$ which is fixed from now on. By compactness there clearly exists a scaling factor $\alpha > 0$ and a translation vector c such that

$$\hat{G} := (G - c)/\alpha = \{y \in \mathbb{R}^n : \ g_j(\alpha y + c) \geq 0, \ j = 1, \ldots, m\} \subset B.$$

As a rather immediate consequence of the constraint qualification it is even possible to choose α, c such that there exist quadratic sos polynomials $\hat{s}_0(y), \ldots, \hat{s}_m(y)$ with

$$1 - \|y/\rho - e\| = \hat{s}_0(y) + \hat{s}_1(y)g_1(\alpha y + c) + \cdots + \hat{s}_m(y)g_m(\alpha y + c). \tag{10}$$

Let us now define

$$\hat{F}(y) := F(\alpha y + c) \ \text{ and } \ \hat{g}_j(y) := \sigma_j g_j(\alpha y + c)$$

where $\sigma_j > 0$ is chosen such that $\hat{g}_j(y) \leq 1$ for all $y \in \Sigma$ (which is possible by compactness of Σ). Note that, still, $\hat{G} = \{y \in \mathbb{R}^n : \ \hat{g}_j(y) \geq 0, \ j = 1, \ldots, m\}$. Moreover, it is clear that $\hat{F}(y) \succ 0$ for all $y \in \hat{G}$. Our main effort will be devoted to proving that there exist sos matrices $\hat{S}_j(y)$, $j = 0, \ldots, m$, with

$$\hat{F}(y) = \hat{S}_0(y) + \hat{S}_1(y)\hat{g}_1(y) + \cdots + \hat{S}_m(y)\hat{g}_m(y).$$

It is then clear that we arrive at (9) by substitution of $y = (x - c)/\alpha$ and with $S_j(x) := \sigma_j \hat{S}_j((x - c)/\alpha)$, $j = 0, \ldots, m$.

Lemma 3 *There exists scalar sos polynomials $s_1(y), \ldots, s_m(y)$ with*

$$P(y) := \hat{F}(y) - [s_1(y)\hat{g}_1(y) + \cdots + s_m(y)\hat{g}_m(y)]I \succ 0 \ \textit{ for all } \ y \in \Sigma.$$

Proof. By compactness of Σ there exist $f_0 > 0$ and $\rho_0 > 0$ such that

$$\hat{F}(y) \succ f_0 I \ \text{ on } \ \{y \in \Sigma : \ \hat{g}_j(y) \geq -\rho_0, \ j = 1, \ldots, m\}.$$

Moreover we can fix $f_1 > 0$ with $\hat{F}(y) \succ f_1 I$ for $y \in \Sigma$. Similarly as in [12] we infer from $(z-1)^{2k} z \leq 1/(2k+1)$ for $z \in [0,1]$ that

$$\sum_{\{j:\ \hat{g}_j(y)\geq 0\}} (1-\hat{g}_j(y))^{2k}\hat{g}_j(y) \leq \frac{m}{2k+1} \quad \text{for all} \quad y \in \Sigma. \tag{11}$$

It is not difficult to fix a non-negative integer k and $\xi > 0$ such that

$$f_0 - \frac{m\xi}{2k+1} > 0 \quad \text{and} \quad f_1 - \frac{m\xi}{2k+1} + \xi(1+\rho_0)^{2k}\rho_0 > 0. \tag{12}$$

Since $y \in \hat{G}$ implies $\hat{g}_j(y) \geq 0$ for all $j = 1, \ldots, m$ we can combine (11) and the first relation in (12) to infer

$$\hat{F}(y) - \xi \sum_{j=1}^{m} (1-\hat{g}_j(y))^{2k}\hat{g}_j(y) I \succ 0 \quad \text{for all} \quad y \in \hat{G}.$$

It is even possible to show that this inequality holds for all $y \in \Sigma$ (which finishes the proof by setting $s_j(y) := \xi(1-\hat{g}_j(y))^{2k}$ which are clearly sos). Indeed consider

$$\begin{aligned}\hat{F}(y) - \xi \left(\sum_{\{j:\, \hat{g}_j(y)\geq 0\}} + \sum_{\{j:\, 0>\hat{g}_j(y)>-\rho_0\}} + \sum_{\{j:\, -\rho_0\geq \hat{g}_j(y)\}} \right) (1-\hat{g}_j(y))^{2k}\hat{g}_j(y) I \succeq \\ \succeq \hat{F}(y) - \frac{m\xi}{2k+1} I - \xi \sum_{\{j:\, -\rho_0\geq \hat{g}_j(y)\}} (1-\hat{g}_j(y))^{2k}\hat{g}_j(y) I.\end{aligned}$$

If $y \in \Sigma$ satisfies $\hat{g}_j(y) > -\rho_0$ for all $j = 1, \ldots, m$ then positivity follows from the first relation in (12). Otherwise there is at least one j_0 with $-\rho_0 \geq \hat{g}_{j_0}(y)$ which implies $(1-\hat{g}_{j_0}(y))^{2k} \geq (1+\rho_0)^{2k}$ and hence

$$-\xi \sum_{\{j:\, -\rho_0\geq \hat{g}_j(y)\}} (1-\hat{g}_j(y))^{2k}\hat{g}_j(y) \geq \xi(1+\rho_0)^{2k}[-\hat{g}_{j_0}(y)] \geq \xi(1+\rho_0)^{2k}\rho_0.$$

Then positivity follows from $\hat{F}(y) > f_1 I$ and the second inequality in (12).

■

Lemma 4 *Suppose that the polynomial matrix $P(y)$ is positive definite on Σ. Then there exists a non-negative integer M and positive definite matrices P_β such that P has the representation*

$$P(y) = \sum_{\beta\in\mathbb{N}_0^{n+1},\, |\beta|=M} P_\beta\, y_1^{\beta_1} \cdots y_n^{\beta_n} (1-e^T y)^{\beta_{n+1}}.$$

Proof. Suppose P has degree d and decompose P into $P_0 + \cdots + P_d$ with homogenous $P_0, \ldots, P_d$ of degrees $0, \ldots, d$ respectively. With $z \in \mathbb{R}$ define the homogenous d-degree polynomial

$$H(y,z) := \sum_{j=0}^{d} P_j(y)(e^T y + z)^{d-j}.$$

Since the restriction of $H(y,z)$ to the simplex $\{(y,z) \in \mathbb{R}^n \times \mathbb{R} : \; y \geq 0, \; z \geq 0, \; e^T y + z = 1\}$ equals $P(y)$ and is hence positive definite, we can apply the matrix-valued version of Pólya's theorem [11] to conclude that there exists a non-negative integer N such that

$$(e^T y + z)^N H(y,z) = \sum_{\beta \in \mathbb{N}_0^{n+1}, \, |\beta| = N+d} P_\beta \, y_1^{\beta_1} \cdots y_n^{\beta_n} z^{\beta_{n+1}} \quad \text{with} \quad P_\beta \succ 0.$$

The result follows by plugging in $z = 1 - e^T y$.

■

The inclusion $B \subset \Sigma$ implies that y_j and $1 - e^T y$ have sos representations in terms of 1 and $1 - \|y/\rho - e\|^2$. It is actually not difficult to construct such representations with degree bounds explicitly.

Lemma 5 *For nonzero $\beta \in \mathbb{N}_0^{n+1}$ there exist sos polynomials $s_\beta(x)$ and $t_\beta(x)$ with degrees $2|\beta|$ and $2|\beta| - 2$ such that*

$$y_1^{\beta_1} \cdots y_n^{\beta_n} (1 - e^T y)^{\beta_{n+1}} = s_\beta(y) + t_\beta(y)(1 - \|y/\rho - e\|^2). \tag{13}$$

Proof. With $t_0(y) := \rho\sqrt{n}/2 \geq 0$ is is elementary to verify that $s_0(y) := 1 - e^T y - t_0(y)(1 - \|y/\rho - e\|^2)$ is nonnegative. This leads to the sos representation

$$1 - e^T y = s_0(y) + t_0(y)(1 - \|y/\rho - e\|^2)$$

with s_0, t_0 of degrees 2 and 0 respectively. Similarly we have

$$y_i = s_i(y) + t_i(y)(1 - \|y/\rho - e\|^2) \quad \text{for} \quad s_i(y) = (\rho/2)\|y/\rho - (e - e_i)\|^2, \; t_i(y) = \rho/2.$$

This finishes the proof for $|\beta| = 1$, and it is simple to recursively construct (13) from this representation for $|\beta| > 1$.

■

Let us now just combine the previous two results.

Lemma 6 *There exist sos matrices $S(y)$, $T(y)$ with $P(y) = S(y) + T(y)(1 - \|y/\rho - e\|^2)$.*

Proof. Indeed the relation holds for the sos matrices

$$S(y) = \sum_{\beta \in \mathbb{N}_0^{n+1}, \, |\beta| = M} P_\beta \, s_\beta(y) \quad \text{and} \quad T(y) = \sum_{\beta \in \mathbb{N}_0^{n+1}, \, |\beta| = M} P_\beta \, t_\beta(y).$$

■

If we combine this with (10) and Lemma 3 we arrive at

$$\hat{F}(y) - [s_1(y)\hat{g}_1(y) + \cdots + s_m(y)\hat{g}_m(y)]I = [S(y) + \hat{s}_0(y)T(y)] + \sum_{j=1}^{m} [\hat{s}_j(y)T(y)]\hat{g}_j(y).$$

This is the desired representation result with $\hat{S}_0(y) = S(y) + \hat{s}_0(y)T(y)$ and $\hat{S}_j(y) = s_j(y) + T(y)\hat{s}_j(y)$, $j = 1, \ldots, m$.

4 Robust LMI problem

Let us now illustrate how we can apply Theorem 2 in order to construct relaxations of Problem 1. For the direct approach we assume that $L(x, y)$ is polynomial in x. By Theorem 2 applied to (2), the optimal value t_{opt} equals the infimal γ such that there exist y and sos matrices $S_0(x), S_1(x), \ldots, S_m(x)$ with

$$M_\gamma(x, y) = S_0(x) + S_1(x)g_1(x) + \cdots + S_m(x)g_m(x).$$

If we impose a priori constraints on the degrees of the sos matrices, this problem can be readily translated into a standard SDP [5], whose optimal value will be a lower bound of t_{opt}. If we let the degree bound grow, the relaxations' optimal values converge to t_{opt} by Theorem 2.

For the second approach we only require $L(x, y)$ to be rational in x without pole at zero. Then one can determine a linear fractional representation

$$L(x, y) = \Delta(x) \star \begin{pmatrix} A & B \\ C(y) & D(y) \end{pmatrix} := D(y) + C(y)\Delta(x)(I - A\Delta(x))^{-1}B$$

where the matrices A, B are fixed and $C(y)$, $D(y)$, $\Delta(x)$ depend affinely and linearly on y, x respectively. Since G is compact we can apply the full block S-procedure [6, 10] for each fixed y. Therefore $I - A\Delta(x)$ is nonsingular for all $x \in G$ and (2) holds iff there exists a so-called multiplier P such that

$$\begin{pmatrix} \Delta(x) \\ I \end{pmatrix}^T P \begin{pmatrix} \Delta(x) \\ I \end{pmatrix} \succ 0 \quad \text{for all} \quad x \in G \tag{14}$$

and

$$\begin{pmatrix} I & 0 \\ A & B \end{pmatrix}^T P \begin{pmatrix} I & 0 \\ A & B \end{pmatrix} + \begin{pmatrix} 0 & I \\ C(y) & D(y) \end{pmatrix}^T \begin{pmatrix} -I\gamma & 0 \\ 0 & I/\gamma \end{pmatrix} \begin{pmatrix} 0 & I \\ C(y) & D(y) \end{pmatrix} \prec 0. \tag{15}$$

Clearly (15) can be rendered affine in P, γ, y (Schur complement). Therefore the only difficulty in computations arises from describing all P with (14), and exactly at this point we apply Theorem 2. Indeed P satisfies (14) iff there exist sos matrices $S_0(x), S_1(x), \ldots, S_m(x)$ (generally depending upon P) such that

$$\begin{pmatrix} \Delta(x) \\ I \end{pmatrix}^T P \begin{pmatrix} \Delta(x) \\ I \end{pmatrix} = S_0(x) + S_1(x)g_1(x) + \cdots + S_m(x)g_m(x). \tag{16}$$

With degree bounds on $S_i(x)$, $i = 0, 1, \ldots, m$, the set of all such P admits an LMI representation [5], such that we can determine the infimal γ with (15) and (16) by solving a linear SDP. Again this forms a relaxation of (1) for computing a lower bound on t_{opt}, and for growing degrees of the sos matrices this bound is guaranteed to converge to t_{opt}. We conclude by stressing that this relaxation procedure is the complete analogue to [11], with the essential difference that we can work with an implicit description of the set G.

References

[1] A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. SIAM-MPS Series in Optimizaton. SIAM Publications, Philadelphia, 2001.

[2] G. Chesi, A. Garulli, A. Tesi, and A. Vicino. An LMI-based approach for characterizing the solution set of polynomial systems. In *Proc. 39th IEEE Conf. Decision and Control*, Sydney, Australia, 2000.

[3] L. El Ghaoui, F. Oustry, and H. Lebret. Robust solutions to uncertain semidefinite programs. *SIAM J. Optim.*, 9(1):33–52, 1999.

[4] D. Henrion, M. Sebek, and V. Kucera. Positive polynomials and robust stabilization with fixed-order controllers. *IEEE Transactions on Automatic Control*, 48:1178–1186, 2003.

[5] C.W.J. Hol and C.W. Scherer. Sum of squares relaxations for polynomial semidefinite programming. In *Proceedings MTNS 2004*, Leuven, Belgium, 2004.

[6] T. Iwasaki and G. Shibata. LPV system analysis via quadratic separator for uncertain implicit systems. *IEEE Trans. Aut. Contr.*, 46(8):1195–1208, 2001.

[7] T. Jacobi and A. Prestel. Distinguished representations of strictly positive polynomials. *J. Reine Angew. Math.*, 532:223–235, 2001.

[8] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.*, 11(3):796–817, 2001.

[9] P. Parrilo. *Structure semidefinite programs and semialgebraic geometry methods in robustness and optimization.* PhD thesis, California Institute of Technology, 2000.

[10] C.W. Scherer. LPV control and full block multipliers. *Automatica*, 37:361–375, 2001.

[11] C.W. Scherer. Higher-order relaxations for robust LMI problems with verifications for exactness. In *Proc. 42nd IEEE Conf. Decision and Control*, pages 4652–4657, Maui, Hawaii, 2003.

[12] M. Schweighofer. Optimization of polynomials on compact semialgebraic sets. *preprint*, 2003.