J. William Helton · Jiawang Nie

Semidefinite Representation of Convex Sets

Abstract. Let $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \cdots, g_m(x) \geq 0\}$ be a semialgebraic set defined by multivariate polynomials $g_i(x)$. Assume S is convex, compact and has nonempty interior. Let $S_i = \{x \in \mathbb{R}^n : g_i(x) \geq 0\}$, and ∂S (resp. ∂S_i) be the boundary of S (resp. S_i). This paper, as does the subject of semidefinite programming (SDP), concerns Linear Matrix Inequalities (LMIs). The set S is said to have an LMI representation if it equals the set of solutions to some LMI and it is known that some convex S may not be LMI representable [9]. A question arising from [15], see [9,16], is: given a subset S of $\mathbb{R}^n$, does there exist an LMI representable set $\hat{S}$ in some higher dimensional space $\mathbb{R}^{n+N}$ whose projection down onto $\mathbb{R}^n$ equals S. Such S is called semidefinite representable or SDP representable. This paper addresses the SDP representability problem.

The following are the main contributions of this paper: **(i)** Assume $g_i(x)$ are all concave on S. If the positive definite Lagrange Hessian (PDLH) condition holds, i.e., the Hessian of the Lagrange function for optimization problem of minimizing any nonzero linear function $\ell^T x$ on S is positive definite at the minimizer, then S is SDP representable. **(ii)** If each $g_i(x)$ is either sos-concave ($-\nabla^2 g_i(x) = W(x)^T W(x)$ for some possibly nonsquare matrix polynomial $W(x)$) or strictly quasi-concave on S, then S is SDP representable. **(iii)** If each S_i is either sos-convex or poscurv-convex (S_i is compact convex, whose boundary has positive curvature and is nonsingular, i.e. $\nabla g_i(x) \neq 0$ on $\partial S_i \cap S$), then S is SDP representable. This also holds for S_i for which $\partial S_i \cap S$ extends smoothly to the boundary of a poscurv-convex set containing S. **(iv)** We give the complexity of Schmüdgen and Putinar's matrix Positivstellensatz, which are critical to the proofs of (i)-(iii).

Key words. Convex sets, semialgebraic geometry, semidefinite programming (SDP),linear matrix inequality (LMI), sum of squares (SOS), modified Hessian, moments, convex polynomials positive curvature, Schmüdgen and Putinar's matrix Positivstellensatz, positive definite Lagrange Hessian (PDLH) condition, extendable poscurv-convex, positive second fundamental form, poscurv-convex, sos-concave (sos-convex)

1. Introduction

One of the main advances in optimization which has had a profound effect on control theory and nonconvex optimization as well as many other disciplines is semidefinite programming (SDP) [16,28]. This gives effective numerical algorithms for solving problems presented in terms of Linear Matrix Inequalities (LMIs). Arising from this is the very basic issue of which problems can be presented with LMIs and this paper addresses one of the most classical aspects of this problem.

We say a set S have an *LMI representation* or is *LMI representable* if

$$S = \{x \in \mathbb{R}^n : A_0 + \sum_{i=1}^n A_i x_i \succeq 0\} \tag{1.1}$$

for some symmetric matrices A_i. Here the notation $X \succeq 0 \, (\succ 0)$ means the matrix X is positive semidefinite (definite). If S has an interior point, A_0 can be assumed to be positive definite without

J. William Helton: Department of Mathematics, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093. e-mail: `helton@math.ucsd.edu`

Jiawang Nie: Department of Mathematics, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093. e-mail: `njw@math.ucsd.edu`

loss of generality. Obvious necessary conditions for S to be LMI representable are that S must be convex and S must also be a basic closed semialgebraic set

$$S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \cdots, g_m(x) \geq 0\}$$

where $g_i(x)$ are multivariate polynomials. We shall always assume S has an interior point. For example, any convex quadratic constraint $\{x \in \mathbb{R}^n : a + b^T x - x^T C^T C x \geq 0\}$ can be represented by the LMI

$$\left\{ x \in \mathbb{R}^n : \begin{bmatrix} I_n & Cx \\ (Cx)^T & a + b^T x \end{bmatrix} \succeq 0 \right\}$$

where I_n is the $n \times n$ identity matrix. Here B^T denotes the transpose of matrix B. A basic question (asked in [19]) is: which convex sets can be represented by LMIs? It turns out that some convex sets are not LMI representable. Helton and Vinnikov [9] proved that a strong condition called *rigid convexity* is necessary for a set to have an LMI representation (as well as sufficient in case of dimension two). For instance, the convex set

$$T = \left\{ x \in \mathbb{R}^2 : 1 - (x_1^4 + x_2^4) \geq 0 \right\} \tag{1.2}$$

does not admit an LMI representation [9], since it is not rigidly convex.

However, the set T is the projection onto x-space of the set

$$\hat{S} := \left\{ (x, w) \in \mathbb{R}^2 \times \mathbb{R}^2 : \begin{bmatrix} 1 + w_1 & w_2 \\ w_2 & 1 - w_1 \end{bmatrix} \succeq 0, \begin{bmatrix} 1 & x_1 \\ x_1 & w_1 \end{bmatrix} \succeq 0, \begin{bmatrix} 1 & x_2 \\ x_2 & w_2 \end{bmatrix} \succeq 0 \right\}$$

in $\mathbb{R}^4$ which is represented by an LMI. This motivates

Question: *Which convex sets S are the projection of a set $\hat{S}$ having an LMI representation; in other words, do there exist symmetric matrices F_i, G_j such that S equals $\{x : (x, y) \in \hat{S}\}$ where*

$$\hat{S} = \left\{ (x, y) \in \mathbb{R}^{(n+N)} : F_0 + \sum_{i=1}^n F_i x_i + \sum_{j=1}^N G_j y_j \succeq 0 \right\}. \tag{1.3}$$

Such sets S are called *semidefinite representable* or *SDP representable*. Ben-Tal and Nemirovskii ([1]), Nesterov and Nemirovskii ([15]), and Nemirovskii ([16]) gave collections of examples of SDP representable sets. Thereby leading to the question which sets are SDP representable? In §4.3.1 of his excellent 2006 survey [16] Nemirovsky said " this question seems to be completely open". Obviously, to be SDP representable, S must be convex and semialgebraic. What are the sufficient conditions that guarantee S is SDP representable? This paper addresses this kind of question. Sometimes we refer to a semidefinite representation as a *lifted LMI representation* of the convex set S and to the LMI in (1.3) as the *lifted LMI for S*.

A construction of the SDP representation for convex sets was proposed by Lasserre [13] and also in the dimension two case by Parrilo, for example in [18], and could be viewed using the following idea. Let $\mathcal{M}$ denote the space of Borel measures on S and let $\hat{S}$ denote the convex subset of all nonnegative mass one measures. The Krein Millman Theorem [4] says that $\hat{S}$ projects down onto S via

$$P(\mu) := \int_S x \, d\mu(x) \qquad \mu \in \hat{S}.$$

Unfortunately $\hat{S}$ is infinite dimensional, so unsuitable as an SDP representation. The Lasserre and Parrilo proposal, which will be sketched later, is to cut down $\hat{S}$ by looking at it as the set of all positive mass one linear functionals on the polynomials of some fixed degree N. Moment and sum of squares (SOS) techniques show that this gives an LMI, denoted by $\mathcal{L}_N$, for each degree N, and that the projection onto $x-$ space of the set $\hat{S}_N := \{(x, y) : \mathcal{L}_N(x, y) \geq 0\}$ contains S for all N. The open question remaining is whether there exists an integer N large enough to produce the equality.

The validity of this general type of construction has been supported by very nice recent findings on the SDP representation of convex sets. Parrilo [18] proved this gives a lifted LMI representation

in the two dimensional case when the boundary of S is a single rational planar curve of genus zero. Lasserre [13] proved this construction can give arbitrarily accurate approximations when N goes to infinity.

This article gives sufficient conditions (presented as the hypotheses of Theorems 1 through 4) on a convex set S guaranteeing that it is SDP representable. The first condition we present (Theorem 1) is the PDLH condition, and we prove validatity of Lasserre-Parrilo moment type constructions when PDLH holds. After that come three theorems, each having weaker hypotheses than the preceding one and each concluding that S is SDP representable. The last of them (Theorem 4) is a bit weaker than saying that a strictly convex basic semialgebraic set S has an SDP representation, provided its boundary is nonsingular (the gradients of defining polynomials for S do not vanish).

More specifically, Theorems 3 and 4 are based on geometric properties of the boundary of a convex set. Any convex set has boundary, which if a smooth manifold, has nonnegative curvature and conversely, having positive curvature everywhere is slightly stronger than strict convexity. Strict convexity and positively curved boundary are not the same as is illustrated by the set in example (1.2) which while strictly convex has zero curvature at $(-1,0)$, $(1,0)$, $(0,-1)$ and $(0,1)$, although it has positive curvature everywhere else. A good illustration of the gap between our necessary and sufficient conditions for SDP representability is Theorem 3 when specialized to a convex set S defined by a single polynomial $g(x)$. It implies that if the boundary ∂S is nonsingular ($\nabla g(x) \neq 0$ on ∂S) and if ∂S has positive curvature at all points, then S has an SDP representation. Thus for a single defining function $g(x)$ the necessary vs. sufficient gap lies only in the gradient being zero or the curvature being zero somewhere on ∂S. Theorems 3 and 4 also give generalizations of this for more than one defining function.

A subsequent paper, based on this one, [7, Section 3], extends these results and shows that if each component of the boundary of convex and bounded S is positively curved and nonsingular, then S is SDP representable.

We should emphasize that while our description stresses a clean characterization of existence of LMI lifts, we shall soon introduce a class of sets we call SOS- convex and constructions for them which might be practical on modest size problems. Now we turn to a formal presentation of results.

Let $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \cdots, g_m(x) \geq 0\}$ be a basic closed semialgebraic set; here the $g_i(x)$ are in the ring $\mathbb{R}[x]$ of multivariate polynomials with real coefficients and are called the *defining polynomials* for S. Assume S is convex, compact and has nonempty interior. Let $S_i = \{x \in \mathbb{R}^n : g_i(x) \geq 0\}$ and $Z(g_i) = \{x \in \mathbb{R}^n : g_i(x) = 0\}$ be the zero set of g_i. Denote by ∂S and ∂S_i the boundaries of S and S_i respectively. Note that ∂S_i might be contained in $Z(g_i)$ properly.

First, consider the case that all the defining polynomials $g_i(x)$ are concave on S. The *positive definite Lagrange Hessian (PDLH)* condition requires that for any nonzero vector $\ell \in \mathbb{R}^n$, the Hessian of the Lagrange function corresponding to the optimization problem of minimizing $\ell^T x$ over S is positive definite at each minimizer, i.e., $-\sum_{i=1}^{m} \lambda_i \nabla^2 g_i(u)$ is positive definite for every minimizer u and the corresponding Lagrange multipliers $\lambda_i \geq 0$. Obviously, if every $g_i(x)$ has negative definite Hessian on the boundary, then the PDLH condition holds.

Theorem 1. *Suppose $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \cdots, g_m(x) \geq 0\}$ is compact convex and has nonempty interior. Assume $g_i(x)$ are concave on S. If the PDLH condition holds, then S is SDP representable.*

Remark: In Theorem 1, where the $g_i(x)$ are concave on S, the matrix $-\sum_{i=1}^{m} \lambda_i \nabla^2 g_i(u)$ must be positive semidefinite. The PDLH condition requires it is positive definite, i.e., its determinant is nonzero, which defines a Zariski open set. So the PDLH condition is a generic condition subject to the property that $g_i(x)$ are concave on S.

The SDP representation of S in Theorem 1 can be constructed explicitly, which will be shown in Section 2. The lifted LMI (2.2) or (2.4) (archimedean condition is then required) represents S exactly under the PDLH condition. However, when the PDLH condition fails, the constructed LMIs in Section 2 might not represent S correctly. This leads to our next result.

Second, consider the case that all the defining polynomials $g_i(x)$ are *quasi-concave* on S. This means the super level set $S_i(\alpha) = \{x \in S : g_i(x) \geq \alpha\}$ is convex for every $\alpha \in g_i(S)$. So the level set

$Z(g_i - \alpha) = \{x \in \mathbb{R}^n : g_i(x) = \alpha\}$ when smooth has *nonnegative curvature* in S, i.e., for all $x \in S$,

$$-v^T \nabla^2 g_i(x) v \geq 0, \quad \forall\, v \in \nabla g_i(x)^\perp = \{v \in \mathbb{R}^n : v^T \nabla g_i(x) = 0\}.$$

We say $g_i(x)$ is *strictly quasi-concave* on S if every $Z(g_i - \alpha)$ has *positive curvature* in S, i.e.,

$$-v^T \nabla^2 g_i(x) v > 0, \quad \forall\, 0 \neq v \in \nabla g_i(x)^\perp, \, \forall\, x \in S.$$

By Exercise 3.44(a) in Boyd and Vandenberghe [3], the above is equivalent to the *modified Hessian*

$$-\nabla^2 g_i(x) + M \nabla g_i(x) \nabla g_i(x)^T \succ 0$$

for some constant $M > 0$. It will be shown (Lemma 11) that the constant M can be chosen uniformly for all $x \in S$ if $g_i(x)$ is strictly quasi-concave on S.

A polynomial $p(x)$ is said to be a *sum of squares (SOS)* if $p(x) = w(x)^T w(x)$ for some column vector polynomial $w(x)$. The necessary condition for $p(x)$ to be SOS is that it is nonnegative on the whole space $\mathbb{R}^n$, but the converse might not be true. We refer to [22] for a survey on SOS polynomials. A symmetric matrix polynomial $P(x) \in \mathbb{R}[x]^{n \times n}$ is SOS if there exists a possibly nonsquare matrix polynomial $W(x)$ with n columns such that $P(x) = W(x)^T W(x)$. The defining polynomial $g_i(x)$ is called *sos-concave* if the negative Hessian $-\nabla^2 g_i(x) = -\left(\frac{\partial^2 g_i}{\partial x_k \partial x_\ell}\right)$ is SOS. Similarly, $g_i(x)$ is called *sos-convex* if the Hessian $\nabla^2 g_i(x)$ is SOS.

Theorem 2. *Suppose $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \cdots, g_m(x) \geq 0\}$ is compact convex and has nonempty interior. If each $g_i(x)$ is either sos-concave or strictly quasi-concave on S, then S is SDP representable.*

Remark: The case where a piece of the boundary is a linear subspace is included in Theorem 2, since if some $g_i(x)$ is a linear polynomial, then its Hessian is identically zero and hence $g_i(x)$ is sos-concave.

It is possible that the defining functions of a convex set can be neither concave nor quasi-concave, because the defining functions of a convex set can behave badly in the interior. However, they have nice properties near the boundary which are helpful for us to establish the semidefinite representability. This leads to the following results.

Third, consider the case that S is convex but the defining polynomials $g_i(x)$ are not quasi-concave on S. This is because the super level sets of $g_i(x)$ might not be all convex. We call S_i *poscurv-convex* if S_i is compact convex, its boundary ∂S_i equals $Z(g_i)$, and ∂S_i is nonsingular (the gradient does not vanish) and has positive curvature at each point on it, which means that

$$-v^T \nabla^2 g_i(x) v > 0, \quad \forall\, 0 \neq v \in \nabla g_i(x)^\perp, \, \forall\, x \in \partial S_i. \tag{1.4}$$

Note that the definition of poscurv-convex sets also applies to $g_i(x)$ which are *smooth functions* (not necessarily polynomials). S_i is called *sos-convex* if g_i is a polynomial and g_i is sos-concave.

Theorem 3. *Suppose $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \cdots, g_m(x) \geq 0\}$ is compact convex and has nonempty interior. If each S_i is either sos-convex or poscurv-convex, then S is SDP representable.*

We turn now to more general sets S_i. We say S_i is *extendable poscurv-convex with respect to* S if $g_i(x) > 0$ whenever $x \in S$ lies in the interior of S_i and there exists a poscurv-convex set $T_i = \{x : f_i(x) \geq 0\} \supseteq S$ such that $\partial T_i \cap S = \partial S_i \cap S$. Here $f_i(x)$ is a smooth function (not necessarily a polynomial) such that T_i is compact convex, $\partial T_i = Z(f_i)$ and ∂T_i is nonsingular. In other words, $\partial S_i \cap S$ can be extended to become a part of the boundary of a poscurv-convex set defined by a smooth function.

Theorem 4. *Suppose $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \cdots, g_m(x) \geq 0\}$ is compact convex and has nonempty interior. If each S_i is either sos-convex or extendable poscurv-convex with respect to S, then S is SDP representable.*

Theorem 3 is a special case of Theorem 4, since every poscurv-convex set is of course extendable poscurv-convex. However, it turns out, see the follow-up paper [7] to this one, that extendable poscurv-convexity of S_i with respect to S does not require much more than the boundary $\partial S_i \cap \partial S$ has positive curvature. In [7] this is combined with Theorem 4 to obtain a stronger result: if for every i either S_i is sos-convex or $\partial S_i \cap \partial S$ is positively curved and nonsingular, then S is SDP representable.

The proofs for the above theorems are based on a variety of techniques, and they produce new results which might be of interest independent of SDP representation. First, the proofs introduce a natural technique of writing a polynomial as a sum of squares by twice integrating its Hessian, which is very suited to handling sos-concavity. Second, we give degree bounds for polynomials appearing in Schmüdgen's and Putinar's matrix Positivstellensatz, see the Appendix §6. These two techniques allow us to obtain bounds on the degrees of polynomials which appear in SOS representations. Third, it is possible that the set $S_i = \{x \in \mathbb{R}^n : g_i(x) \geq 0\}$ is strictly convex but that the polynomial $g_i(x)$ is neither concave nor quasi-concave. In §4 we show under modest hypotheses that there is a new set of (local) defining polynomials p_i for the S_i which are strictly concave. This allows us to prove Theorem 4 by using the new defining polynomials p_i together with the original polynomials g_i.

Now we say a few words about constructions. When all the defining polynomials $g_i(x)$ are sos-concave, an explicit SDP representation for S is given by (3.1) in Section 3. The forthcoming article [8] illustrates this in several examples, and also shows how to incorporate sparsity. If all $g_i(x)$ are concave on S and every $g_i(x)$ is either sos-concave or strictly concave (having negative definite Hessian) on the boundary ∂S where it vanishes, an explicit SDP representation for S is given by (2.2) or (2.4) in Section 2 when the relaxation order N is big enough. For the time being, we do not have an estimate of how large N is sufficient. In Theorem 1, the SDP representation can be constructed in the same way as in Section 2. In Theorems 2, 3 and 4, we have only shown the existence of SDP representations for S. We would expect it would be very difficult to use the proof there constructively.

The following notations will be used. For $x \in \mathbb{R}^n$, $\|x\| = \sqrt{\sum_{i=1}^n x_i^2}$. $\mathbb{N}$ denotes the set of non-negative integers. For $\alpha \in \mathbb{N}^n$, $|\alpha| := \alpha_1 + \cdots + \alpha_n$, $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. A d-form is a homogenous polynomial of degree d. Given a set $K \subset \mathbb{R}^n$, $C^k(K)$ denotes the set of k-times continuously differentiable functions in an open set containing K, and $C^\infty(K)$ denotes the set of infinitely times differentiable (smooth) functions in an open set containing K. Given $\alpha \in \mathbb{N}^n$ and $f(x) \in C^k(K)$, $D^\alpha f(x) := \frac{\partial^{|\alpha|} f(x)}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}$. For a symmetric matrix A, $\lambda_{\min}(A)$ denotes the smallest eigenvalue of A, and $\|A\|_2$ denotes the standard 2-norm of A.

The paper is organized as follows. Section 2 gives the constructions of lifted LMIs when g_i are concave on S, and states some theorems about the sharpness of these lifted LMIs, whose proofs will be given in Section 5. Then Section 3 turns to Lasserre and Parrilo moment type of constructions of lifted LMIs in [13, 18], and we give a sufficient condition that guarantees these constructed LMIs are the SDP representations of S. Section 4 discusses how to find concave defining functions for poscurv-convex sets used to prove Theorem 4. Section 5 gives proofs of the theorems in Section 2 and in the Introduction. Section 6 is an appendix bounding the degrees of polynomials arising in Schmüdgen's and Putinar's matrix Positivstellensatz. Section 7 summarizes conclusions of the paper.

2. The SDP representations when $g_i(x)$ are concave on S

In this section we assume $g_i(x)$ are concave on S. Two kinds of SDP representations will be constructed. The construction of these SDP representations which we use can be found in Lasserre [13]. We review the construction here, which facilitates understanding the proof of their sharpness which we give in Section 5.

For any integer N, define the monomial vector

$$[x^N] = \begin{bmatrix} 1 & x_1 & \cdots & x_n & x_1^2 & x_1 x_2 & \cdots & x_n^N \end{bmatrix}^T.$$

Then $[x^N][x^N]^T$ is a square matrix and we write

$$[x^N][x^N]^T = \sum_{0 \leq |\alpha| \leq 2N} A_\alpha x^\alpha$$

for some symmetric 0/1-matrices A_α. When $n = 1$, the A_α are Hankel matrices, and when $n > 1$, the A_α are generalized Hankel matrices. Suppose μ is a nonnegative measure on $\mathbb{R}^n$ with total mass equal to one. Integrating the above identity gives us

$$M_N(y) = \int_{\mathbb{R}^n} [x^N][x^N]^T d\mu(x) = \sum_{0 \leq |\alpha| \leq 2N} A_\alpha y_\alpha$$

where $y_\alpha = \int_{\mathbb{R}^n} x^\alpha d\mu(x)$ are the *moments of* μ. The matrix $M_N(y)$ is also called the *moment matrix of order N*.

2.1. SDP representation I

Now we give the first construction of the lifted LMI which only uses finitely many moments. Let $\mu(\cdot)$ be any nonnegative measure such that $\mu(\mathbb{R}^n) = 1$. For any $\nu \in \{0,1\}^m$, define new polynomials $g^\nu(x) := g_1^{\nu_1}(x) \cdots g_m^{\nu_m}(x)$. Let $d_\nu = \lceil \deg(g_1^{\nu_1} \cdots g_m^{\nu_m})/2 \rceil$. For an fixed integer $N \geq d_\nu$, define the *localizing moment matrix* $M_{N-d_\nu}(g^\nu y)$ by

$$M_{N-d_\nu}(g^\nu y) = \int_{\mathbb{R}^n} g^\nu(x)[x^{N-d_\nu}][x^{N-d_\nu}]^T d\mu(x) = \sum_{0 \leq |\alpha| \leq 2N} A_\alpha^\nu y_\alpha$$

where $y_\alpha = \int_{\mathbb{R}^n} x^\alpha d\mu(x)$ are the moments and symmetric A_α^ν are the coefficient matrices such that

$$g^\nu(x)[x^{N-d_\nu}][x^{N-d_\nu}]^T = \sum_{0 \leq |\alpha| \leq 2N} A_\alpha^\nu x^\alpha.$$

For any integer $N \geq \max_\nu d_\nu = \lceil \deg(g_1 \cdots g_m)/2 \rceil$, if $\mathrm{supp}(\mu) \subseteq S$, then

$$\forall \nu \in \{0,1\}^m, \quad M_{N-d_\nu}(g^\nu y) \succeq 0, \quad y_0 = 1. \tag{2.1}$$

Let e_i denote the i-th unit vector in $\mathbb{R}^n$ whose only nonzero entry is one and occurs at index i. If we set $y_0 = 1$ and $y_{e_i} = x_i$ in (2.1), then it becomes the LMI

$$\forall \nu \in \{0,1\}^m, \quad A_0^\nu + \sum_{1 \leq i \leq n} A_{e_i}^\nu x_i + \sum_{1 < |\alpha| \leq 2N} A_\alpha^\nu y_\alpha \succeq 0. \tag{2.2}$$

We mention that the LMI (2.2) is essential the same as the LMI (2.11) in [13].

Let $\hat{S}_N$ denote the set of all vectors (x, y) satisfying (2.2). Note that $\hat{S}_N \subset \mathbb{R}^n \times \mathbb{R}^{\binom{n+2N}{n}-n-1}$. For each N, define the projection mapping

$$\rho_N : \mathbb{R}^n \times \mathbb{R}^{\binom{n+2N}{n}-n-1} \rightarrow \mathbb{R}^n$$
$$(x, y) \mapsto x.$$

The set S is contained in the projection $\rho_N(\hat{S}_N)$ of $\hat{S}_N$ onto x space, because for any $x \in S$ the vector $y = (y_\alpha)$ given by $y_\alpha = x^\alpha$ makes (x, y) satisfies the LMI (2.2). And obviously, the bigger N is, the smaller the projection $\rho_N(\hat{S}_N)$ is. So, for any $N \geq \max_\nu d_\nu$, we have the following chain

$$\rho_N(\hat{S}_N) \supseteq \rho_{N+1}(\hat{S}_{N+1}) \supseteq \cdots \supseteq S.$$

A natural question is whether there exists a finite integer N such that $\rho_N(\hat{S}_N) = S$.

One typical approach to this question is to use linear functionals to separate points in $\hat{S}_N$ from the convex set S. Specifically, given a unit length vector $\ell \in \mathbb{R}^n$, let ℓ^* be the minimum value of the linear function $\ell^T x$ over the set S, let $u \in S$ denote the minimizer, which must exist and be on the boundary ∂S. Since S has nonempty interior, the *Slater's condition* holds, and hence the *first order optimality condition* is satisfied. So there exist Lagrange multipliers $\lambda_1 \geq 0, \cdots, \lambda_m \geq 0$ such that $\ell = \sum_i \lambda_i \nabla g_i(u)$. Suppose each defining polynomial $g_i(x)$ is concave on S. Then

$$f_\ell(x) := \ell^T x - \ell^* - \sum_i \lambda_i g_i(x) \tag{2.3}$$

is a convex function such that $f_\ell(u) = 0$ and $\nabla f_\ell(u) = 0$. Thus

$$f_\ell(x) \geq f_\ell(u) + \nabla f_\ell(u)^T(x-u) = 0, \quad \forall x \in S.$$

In other words, f_ℓ is nonnegative on S and so we could wish $f_\ell(x)$ to have Schmüdgen's representation

$$f_\ell(x) = \sum_{\nu \in \{0,1\}^m} \sigma_\nu(x) g^\nu(x)$$

for some particular SOS polynomials $\sigma_\nu(x)$. Notice that this representation is not implied by Schmüdgen's Positivstellensatz [27] because $f_\ell(x)$ has a zero point u on S.

Indeed, validating the lifted LMI $\hat{S}_N$ for some finite integer N amounts to proving that for all ℓ the polynomial $f_\ell(x)$ has Schmüdgen's representation with uniform (in ℓ) degree bounds on SOS polynomials $\sigma_\nu(x)$; this we will see in Section 5. This is equivalent to proving that a property on S called *Schmüdgen's Bounded Degree Representation (S-BDR)* of affine polynomials holds for S, which means that there exists $N > 0$ such that for *almost every* pair $(a, b) \in \mathbb{R}^n \times \mathbb{R}$

$$a^T x + b > 0 \text{ on } S \quad \Rightarrow \quad a^T x + b = \sum_{\nu \in \{0,1\}^m} \sigma_\nu(x) g^\nu(x)$$

for some SOS polynomials $\sigma_\nu(x)$ with degree bounds $\deg(\sigma_\nu) + \deg(g^\nu) \leq 2N$. When S is a compact set (not necessarily convex), Lasserre [13] showed that if the S-BDR property holds then the convex hull $\text{conv}(S)$ of S equals $\rho_N(\hat{S}_N)$ for N big enough.

S-BDR is a very nice restatement that the lift (2.1) and (2.2) produces an SDP representation. It reduces the main issue to finding concrete and practical conditions assuring the exactness of the lifted LMI (2.2), which is what we do in this paper. Actually, when every polynomial $g_i(x)$ is concave on S and strictly concave on $\partial S_i \cap \partial S$, we can prove a stronger property called *Schmüdgen's Bounded Degree Nonnegative Representation (S-BDNR)* of affine polynomials holds for S, which means that there exists $N > 0$ such that for *every* pair $(a, b) \in \mathbb{R}^n \times \mathbb{R}$

$$a^T x + b \geq 0 \text{ on } S \quad \Rightarrow \quad a^T x + b = \sum_{\nu \in \{0,1\}^m} \sigma_\nu(x) g^\nu(x)$$

for some SOS polynomials $\sigma_\nu(x)$ with degree bounds $\deg(\sigma_\nu) + \deg(g^\nu) \leq 2N$. As we can see, S-BDNR is a stronger property than S-BDR. When S-BDNR property holds, S-BDR also holds and hence $S = \rho_N(\hat{S}_N)$ for N big enough by Theorem 2 in Lasserre [13]. To illustrate this we now state our theorem for concave functions and S-BDNR, while its proof will not be given until Section 5.

Theorem 5. *Suppose $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \cdots, g_m(x) \geq 0\}$ is compact convex and has nonempty interior. Assume the $g_i(x)$ are concave on S. For each i, if either $-\nabla^2 g_i(x)$ is SOS or $-\nabla^2 g_i(u) \succ 0$ for all $u \in \partial S_i \cap \partial S$, then the S-BDNR property holds for S and there exists $N > 0$ such that $S = \rho_N(\hat{S}_N)$.*

2.2. SDP representation II

In LMI (2.2), the size of LMI is unfortunately exponential in m. It is huge when m is big. This is because we have used all the possible products $g^\nu(x) = g_1^{\nu_1}(x) \cdots g_m^{\nu_m}(x)$ for all index vector $\nu \in \{0,1\}^m$. If we use only linear products, we can get a similar LMI

$$\forall 0 \leq k \leq m, \quad A_0^{(k)} + \sum_{1 \leq i \leq n} A_{e_i}^{(k)} x_i + \sum_{1 < |\alpha| \leq 2N} A_\alpha^{(k)} y_\alpha \succeq 0 \tag{2.4}$$

where symmetric matrices $A_\alpha^{(k)} = A_\alpha^{e_k}$ in LMI (2.2) (e_0 is the zero index vector). We mention that the LMI (2.2) is essential the same as the LMI (2.12) in [13].

Similar to LMI (2.2), let $\tilde{S}_N$ be the set of all vectors (x, y) satisfying (2.4) and $\tilde{\rho}_N$ be the projection mapping into x-space. Then the following chain relation again holds

$$\tilde{\rho}_N(\tilde{S}_N) \supseteq \tilde{\rho}_{N+1}(\tilde{S}_{N+1}) \supseteq \cdots \supseteq S.$$

The natural question is whether $\tilde{\rho}_N(\tilde{S}_N) = S$ for some finite integer N. This can be shown true under the so called *archimedean condition*: There exist SOS polynomials $s_0(x), s_1(x), \cdots, s_m(x)$ and a number $R > 0$ big enough such that

$$R - \sum_{i=1}^{n} x_i^2 = s_0(x) + s_1(x)g_1(x) + \cdots + s_m(x)g_m(x).$$

Note that the archimedean condition implies S is compact. But the converse might not be true. However, a compact S can be forced to satisfy the archimedean condition by adding a "redundant" constraint like $R - \sum_{i=1}^{m} x_i^2 \geq 0$ for sufficiently large R.

Similar to the lifted LMI (2.2), validating the exactness of the lifted LMI $\hat{S}_N$ amounts to proving that for every ℓ the polynomial $f_\ell(x)$ has Putinar's representation

$$f_\ell(x) = \sigma_0(x) + \sigma_1(x)g_1(x) + \cdots + \sigma_m(x)g_m(x)$$

with uniform degree bounds on SOS polynomials $\sigma_i(x)$. This is equivalent to proving that the so-called *Putinar-Prestel's Bounded Degree Representation (PP-BDR)* property [13] holds for S, that is, there exists $N > 0$ such that for *almost every* $(a, b) \in \mathbb{R}^n \times \mathbb{R}$

$$a^T x + b > 0 \text{ on } S \qquad \Rightarrow \qquad a^T x + b = \sigma_0(x) + \sigma_1(x)g_1(x) + \cdots + \sigma_m(x)g_m(x)$$

for some SOS polynomials $\sigma_i(x)$ with degrees $\deg(\sigma_i) + \deg(g^\nu) \leq 2N$. When S is a compact set (not necessarily convex), Lasserre [13] showed that if the PP-BDR property holds then the convex hull conv(S) of S equals $\tilde{\rho}_N(\tilde{S}_N)$ for N big enough.

As with S-BDR, the PP-BDR property is a nice reformulation of the exactness of the lifted LMI (2.4) and is usually not directly checkable. In this paper, under the archimedean condition, when every polynomial $g_i(x)$ is concave on S and strictly concave on $\partial S_i \cap \partial S$, we can prove a property called *Putinar-Prestel's Bounded Degree Nonnegative Representation (PP-BDNR)* of affine polynomials holds for S, which means that there exists $N > 0$ such that for *every* pair $(a, b) \in \mathbb{R}^n \times \mathbb{R}$

$$a^T x + b \geq 0 \text{ on } S \qquad \Rightarrow \qquad a^T x + b = \sigma_0(x) + \sigma_1(x)g_1(x) + \cdots + \sigma_m(x)g_m(x)$$

for some SOS polynomials $\sigma_i(x)$ with degrees $\deg(\sigma_i g_i) \leq 2N$ (we denote $g_0(x) = 1$). Obviously, PP-BDNR implies PP-BDR. When PP-BDNR property holds, PP-BDR also holds and hence $S = \tilde{\rho}_N(\tilde{S}_N)$ for N big enough by Theorem 2 in Lasserre [13]. The following theorem, whose proof will be given in Section 5, illustrates the this for concave functions and PP-BDNR.

Theorem 6. *Suppose* $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \cdots, g_m(x) \geq 0\}$ *is compact and has nonempty interior. Assume the* $g_i(x)$ *are concave on* S *and the archimedean condition holds. For each* i, *if either* $-\nabla^2 g_i(x)$ *is SOS or* $-\nabla^2 g_i(u) \succ 0$ *for all* $u \in \partial S_i \cap \partial S$, *then the PP-BDNR property holds and there exists* $N > 0$ *such that* $S = \tilde{\rho}_N(\tilde{S}_N)$.

It is possible that the defining polynomials $g_i(x)$ are not concave but the set S is still convex. In this case, does S have an SDP representation? After some modifications in LMI (2.2), the answer is affirmative in very general situations, which is our Theorem 2. However, our proof of Theorem 2 uses Theorem 5 or 6 as a stepping stone.

3. The SDP representation when $g_i(x)$ are sos-concave

Lasserre [13] and Parrilo [18] proposed recipes for an SDP representation. In this section we give a sufficient condition such that the LMI constructed in Lasserre [13] is a lifted LMI of S. We assume the polynomials $g_i(x)$ are concave (not necessarily strictly concave) in the whole space $\mathbb{R}^n$. Certainly the set $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \cdots, g_m(x) \geq 0\}$ is convex.

As was shown in Lasserre [13], the set S is contained in the projection of $\hat{S}$ defined by LMI

$$\left. \begin{aligned} M_{d_g}(y) &\succeq 0 \\ L_{g_1}(y), \cdots, L_{g_m}(y) &\geq 0 \\ y_0 &= 1 \end{aligned} \right\} \tag{3.1}$$

where $d_g = \max_i \lceil \deg(g_i)/2 \rceil$ and $L_{g_i}(y) = \sum_\alpha g_\alpha^{(i)} y_\alpha$ if we write $g_i(x) = \sum_\alpha g_\alpha^{(i)} x^\alpha$. The projection onto x-space of $\hat{S}$ is $\{x : \exists y \in \hat{S}, \ x_i = y_{e_i}, 1 \leq i \leq n\}$. It is natural to ask whether the LMI (3.1) is a lifted LMI for S, i.e., the projection of LMI (3.1) onto x-space equals S.

The standard approach to this question is to use separating linear functionals. As we did in Section 2, for each vector $\ell \in \mathbb{R}^n$, let ℓ^* be the minimum value of $\ell^T x$ over the set S, $u \in S$ be the minimizer, which must be on the boundary ∂S. If there is some point in the interior of S, then the Slater's condition holds and hence there exist Lagrange multipliers $\lambda_1 \geq 0, \cdots, \lambda_m \geq 0$ such that the optimality condition $\ell = \sum_i \lambda_i \nabla g_i(u)$ holds, and hence

$$f_\ell(x) := \ell^T x - \ell^* - \sum_i \lambda_i g_i(x) \tag{3.2}$$

is a convex and nonnegative polynomial in the whole space $\mathbb{R}^n$ such that $f_\ell(u) = 0$ and $\nabla f_\ell(u) = 0$ (see [13]). Under the assumption that the polynomial f_ℓ is SOS for *every* vector $\ell \in \mathbb{R}^n$, Lasserre [13] showed the LMI (3.1) is a lifted LMI for S. If f_ℓ is not a sum of squares for some particular $\ell \in \mathbb{R}^n$, then the LMI (3.1) might not be a lifted LMI for S.

Although it is very difficult to tell if a polynomial is nonnegative, it is more tractable to check if a polynomial is SOS, which can be done by solving an SDP feasibility problem, e.g., by softwares like *SOSTOOLS* [20] and *Gloptipoly* [10]. However, it is impossible to check if f_ℓ is SOS for uncountably many vectors $\ell \in \mathbb{R}^n$. Here we give a sufficient condition for the LMI (3.1) to be a lifted LMI of S, which can be checked numerically. Let us start the discussion with a lemma.

Lemma 7 *If a symmetric matrix polynomial $P(x) \in \mathbb{R}[x]^{r \times r}$ is SOS, i.e., $P(x) = W(x)^T W(x)$ for some possibly nonsquare matrix polynomial $W(x) \in \mathbb{R}[x]^{k \times r}$, then for any $u \in \mathbb{R}^n$ the double integral*

$$\int_0^1 \int_0^t P(u + s(x - u)) \, ds \, dt$$

is also a symmetric SOS matrix polynomial in $\mathbb{R}[x]^{r \times r}$. In particular, when $r = 1$, the above double integral of scalar SOS polynomials is also SOS.

Proof Let $2d = \deg(P)$. Let $\xi \in \mathbb{R}^r$ be a symbolic vector. Then $P(x)$ is SOS if and only if

$$\xi^T P(x) \xi = [\xi x^d]^T A^T A [\xi x^d]$$

for matrix A. Here $[\xi x^d]$ denotes the vector of monomials

$$\begin{bmatrix} \xi_1 [x^d]^T & \cdots & \xi_r [x^d]^T \end{bmatrix}^T.$$

Note that $\xi^T P(x) \xi$ has degree 2 in ξ. If $P(x)$ is SOS, we can assume the above A exists. In monomial vector $[\xi x^d]$ we replace x by $u + s(x - u)$. Each entry of $[\xi(u + s(x - d))^d]$ is a polynomial in x whose coefficients are polynomials in u and s. So there exists a matrix polynomial $C(u, s)$ such that $[\xi(u + s(x - d))^d] = C(u, s)[\xi x^d]$. Therefore we have

$$\int_0^1 \int_0^t \xi^T P(u + s(x - u)) \xi \, ds \, dt = [\xi x^d]^T B^T B [\xi x^d]$$

where B is a matrix such that

$$\int_0^1 \int_0^t C(u, s)^T A^T A C(u, s) \, ds \, dt = B^T B.$$

Therefore, the double integral of matrix polynomial in the lemma is SOS. $\square$

Remark: The above integral is the limit of a sequence of Riemann sums, which are all SOS with bounded degrees. So intuitively the integral must also be SOS.

Lemma 8 *Let $p(x)$ be polynomial such that $p(u) = 0$ and $\nabla p(u) = 0$ for some point $u \in \mathbb{R}^n$. If the Hessian $\nabla^2 p(x)$ is SOS, then $p(x)$ is SOS.*

Proof Let $q(t) = p(u + t(x - u))$ be a univariate polynomial in t. Then

$$q''(t) = (x - u)^T \nabla^2 p(u + t(x - u))(x - u).$$

So we have

$$p(x) = q(1) = (x - u)^T \left(\int_0^1 \int_0^t \nabla^2 p(u + s(x - u)) \, ds \, dt \right) (x - u).$$

Since $\nabla^2 p(x)$ is SOS, the middle double integral above should also be SOS, by Lemma 7. So $p(x)$ is also SOS. $\square$

The following theorem gives the sufficient condition which we are aiming at.

Theorem 9. *Assume S has nonempty interior. If every $g_i(x)$ is sos-concave, then (3.1) is a lifted LMI representation for S.*

Proof It is obviously that S is contained in the projection of the set $\hat{S}$. If they are not equal, there must exist some $\hat{y} \in \hat{S}$ such that $\hat{x} = (\hat{y}_{e_1}, \cdots, \hat{y}_{e_n}) \notin S$. Since S is closed, there exists a supporting hyperplane of S that excludes $\hat{x}$, i.e., there exists a unit length vector $\ell \in \mathbb{R}^n$ such that

$$\ell^T x - \ell^* \geq 0, \ \ \forall \, x \in S, \quad \ell^T u - \ell^* = 0, \ \ \exists \, u \in \partial S, \qquad \ell^T \hat{x} - \ell^* < 0.$$

Then u is a minimizer for the above. Since S has nonempty interior, the Slater's condition holds, and hence there must exist Lagrange multipliers $\lambda_1 \geq 0, \cdots, \lambda_m \geq 0$ such that $\ell = \sum_{i=1}^m \lambda_i \nabla g_i(u)$. Thus $f_\ell(x) = \ell^T x - \ell^* - \sum_i \lambda_i g_i(x)$ is convex polynomial in $\mathbb{R}^n$ such that $f_\ell(u) = 0$ and $\nabla f_\ell(u) = 0$. Note that $\nabla^2 f_\ell(x) = \sum_{i=1}^m \lambda_i (-\nabla^2 g_i(x))$ is SOS, since all $g_i(x)$ are sos-concave. Hence Lemma 8 implies that $f_\ell(x)$ is SOS, i.e., there exists a symmetric matrix $W \succeq 0$ such that the identity

$$\ell^T x - \ell^* = \sum_{i=1}^m \lambda_i g_i(x) + [x^{d_g}]^T W [x^{d_g}]$$

holds. In the above identity, replace each monomial x^α by $\hat{y}_\alpha$, then we get

$$\ell^T \hat{x} - \ell^* = \sum_{i=1}^m \lambda_i L_{g_i}(\hat{y}) + Trace\big(W \cdot M_{d_g}(\hat{y})\big) \geq 0,$$

which contradicts the previous assertion $\ell^T \hat{x} - \ell^* < 0$. $\square$

Remarks: (i) We do not need assume S is compact whenever every defining polynomial $g_i(x)$ is assumed to be sos-concave. (ii) Checking whether $g_i(x)$ is sos-concave can be done numerically. Obviously, $-\nabla^2 g_i(x)$ is SOS if and only if the polynomial $-\sum_{k,\ell=1}^n \frac{\partial^2 g_i(x)}{\partial x_k \partial x_\ell} \xi_k \xi_\ell$ in (x, ξ) is SOS. This can be checked numerically by solving an SDP feasibility problem, e.g., by softwares *SOSTOOLS* and *Gloptipoly.*

4. Concave defining functions for poscurv-convex sets

It is possible that the set $S_i = \{x \in \mathbb{R}^n : g_i(x) \geq 0\}$ is convex but that its defining polynomials $g_i(x)$ are neither concave nor quasi-concave. The goal of this section is to find a new set of defining polynomials for S. Under some modest hypotheses, we show that S can be defined by a set of polynomials having negative definite Hessians on S when S is defined by strictly quasi-concave polynomials, and that there is a new set of (local) defining polynomials p_i for the S_i which are strictly concave. These results, Proposition 10, Proposition 17 and Corollary 18, might be of interest independent of our SDP representation application. Proposition 10 and Proposition 17 allow us to prove Theorems 2 and 4 by using the p_i together with the given g_i.

We start with a short review of curvature and convexity. For each $u \in Z(g_i)$ with $\nabla g_i(u) \neq 0$, the hyperplane

$$H_i(u) = \{x \in \mathbb{R}^n : \nabla g_i(u)^T(x-u) = 0\}$$

is tangent to $Z(g_i)$ at u. Also the set S_i is strictly convex, in some neighborhood of u if and only if

$$H_i(u) \cap Z(g_i) \cap B(u, \epsilon) = \{u\} \tag{4.1}$$

for a small $\epsilon > 0$, which holds if and only if

$$g_i(x) = (x-u)^T \nabla^2 g_i(u)(x-u) + o(\|x-u\|^2) \ < 0, \qquad \forall \, u \neq x \in H_i(u) \cap B(u, \epsilon).$$

This implies the quadratic form associated with the negative Hessian

$$\Phi(v) := -v^T \nabla^2 g_i(u) v$$

is nonnegative for all v in the tangent space

$$\nabla g_i(u)^\perp = \{v \in \mathbb{R}^n : \nabla g_i(u)^T v = 0\}.$$

We follow common usage (c.f. [6])and call the quadratic function Φ on the tangent space $\nabla g_i(u)^\perp$ the *second fundamental form* of $Z(g_i)$. A surface has *positive curvature* at u if and only if the second fundamental form is strictly positive definite there. Obviously, when S_i is convex positive curvature of $Z(g_i)$ everywhere on ∂S_i implies strict convexity of S_i, but the converse is not necessarily true. Results like this which assume nonnegative curvature, but do not require S_i to be convex as a hypothesis are in [5].

Note that while in this paper we sometimes require $\nabla g_i(u) \neq 0$ at $u \in Z(g_i) \cap S$, the definition of positive curvature of $Z(g_i)$ itself does not. Indeed if the gradient $\nabla g_i(u)$ vanishes, then one can interpret $\nabla g_i(u)^\perp$ as the whole space and its negative Hessian $-\nabla^2 g_i(u)$ is required to be positive definite.

We will distinguish the cases that S is defined by strictly quasi-concave functions and S has positively curved boundary. A new set of defining polynomials for S will be discussed in two subsections separately.

4.1. Convex sets defined by quasi-concave functions

The following proposition gives S_i a new defining polynomial p_i whose Hessian is negative definite on S when $g_i(x)$ is strictly quasi-concave on S.

Proposition 10 *Assume $g_i(x)$ is strictly quasi-concave on S. Then there exists a polynomial $h_i(x)$ such that $h_i(x)$ is positive on S and the product $p_i(x) = g_i(x)h_i(x)$ is concave on S and has negative definite Hessian there.*

We give the proof of this proposition after introducing some lemmas. Without loss of generality, suppose $0 \leq g_i(x) \leq 1$ for all $x \in S$, since S is compact, because otherwise we can scale the coefficients of $g_i(x)$. The set S_i can be convex without $g_i(x)$ being concave. So the Hessian $-\nabla^2 g_i(x)$ might be indefinite. However, the Hessian $-\nabla^2 g_i(x)$ can have at most one negative eigenvalue for $x \in S$, and better yet the Hessian can be "relaxed" to yield the "modified Hessian" which is positive definite.

Lemma 11 *Assume $g_i(x)$ is strictly quasi-concave on S. Then we have*

*(a) There exists M sufficiently large such that the **modified Hessian***

$$-\nabla^2 g_i(x) + M \nabla g_i(x) \nabla g_i(x)^T$$

is positive definite for all $x \in S$.
(b) If $g_i(x)$ is concave, then the above modified Hessian is positive definite for any $M > 0$.

Proof (a). Let $U_i = \{x \in S : -\nabla^2 g_i(x) \succ 0\}$, which is an open set in S. Then $V_i = S - U_i$ is a compact set. Choose an arbitrary point $u \in V_i$ and let $\alpha = g_i(u) \geq 0$. Then $Z(g_i - \alpha) = \{x \in \mathbb{R}^n : g_i(x) = \alpha\}$ has positive curvature in S. Note that for every $u \in V_i$, the negative Hessian $-\nabla^2 g_i(u)$ is not positive definite and hence $\nabla g_i(u) \neq 0$, because otherwise $\nabla g_i(u)^\perp$ is the whole space $\mathbb{R}^n$ which implies $-\nabla^2 g_i(u)$ is positive definite.

Let $Q_i(u) = [\nabla g_i(u) \;\; \hat{Q}_i(u)] \in \mathbb{R}^{n \times n}$ be a nonsingular matrix such that $\nabla g_i(u)^T \hat{Q}_i(u) = 0$. Then

$$Q_i(u)^T \left(-\nabla^2 g_i(u) + M \nabla g_i(u) \nabla g_i(u)^T\right) Q_i(u)$$
$$= \begin{bmatrix} -\nabla g_i(u)^T \nabla^2 g_i(u) \nabla g_i(u) + M \|\nabla g_i(u)\|^2 & -\nabla g_i(u)^T \nabla^2 g_i(x) \hat{Q}_i(u) \\ -\hat{Q}_i(u)^T \nabla^2 g_i(u) \nabla g_i(u) & -\hat{Q}_i(u)^T \nabla^2 g_i(u) \hat{Q}_i(u) \end{bmatrix}.$$

For $u \in V_i$, $-\hat{Q}_i(u)^T \nabla^2 g_i(u) \hat{Q}_i(u) \succ 0$ and $\nabla g_i(u) \neq 0$ on V_i. Since V_i is compact, we can choose M big enough such that the modified Hessian is positive definite for all $u \in V_i$. When $u \in U_i$, the modified Hessian is obviously positive definite.

(b). If $g_i(x)$ is concave, then the modified Hessian is obviously positive semidefinite. We need show it is positive definite for any $M > 0$. Suppose for some $u \in S$ and a vector $\xi \in \mathbb{R}^n$

$$-\xi^T \nabla^2 g_i(u) \xi + M \xi^T \nabla g_i(u) \nabla g_i(u)^T \xi = 0.$$

Then it must hold

$$-\xi^T \nabla^2 g_i(u) \xi = 0, \qquad \nabla g_i(u)^T \xi = 0.$$

Since $-\nabla^2 g_i(u) \succ 0$ in the tangent space $\nabla g_i(u)^\perp$, we must have $\xi = 0$, which completes the proof. $\square$

Example 12 *(1) The following set is strictly convex*

$$\{x \in \mathbb{R}^2 : \underbrace{x_1 x_2 - 1}_{g_1(x)} \geq 0, \; \underbrace{1 - (x_1 - 1)^2 - (x_2 - 1)^2}_{g_2(x)} \geq 0\}.$$

$g_2(x)$ *is strictly concave, but* $g_1(x)$ *is not concave. However, for any* $M > \frac{1}{2}$, *the modified Hessian*

$$-\nabla^2 g_1(x) + M \nabla g_1(x) \nabla g_1(x)^T$$

is positive definite on S.

(2) The condition that g_i *is strictly quasi-concave in* S *in Lemma 11 can not be weakened to* S *is strictly convex. For a counterexample, consider the strictly convex set*

$$\{x \in \mathbb{R}^2 : \underbrace{x_2 - x_1^3}_{g_1(x)} \geq 0, \; \underbrace{x_2 + x_1^3}_{g_2(x)} \geq 0, \; \underbrace{1 - (x_1 - 1)^2 - (x_2 - 1)^2}_{g_3(x)} \geq 0\}.$$

For $i = 1, 2$, *no matter how big* M *is, the modified Hessian*

$$-\nabla^2 g_i(x) + M \nabla g_i(x) \nabla g_i(x)^T = \begin{bmatrix} \pm 6x_1 + 9M x_1^4 & \pm 3x_1^2 \\ \pm 3x_1^2 & M \end{bmatrix}$$

can not be positive semidefinite near the origin.

Lemma 13 *For an arbitrarily large number* $M > 0$, *there exists a univariate polynomial function* $\phi(t)$ *such that for all* $t \in [0, 1]$

$$\phi(t) > 0, \qquad \phi(t) + \phi'(t)t > 0, \qquad \frac{2\phi'(t) + \phi''(t)t}{\phi(t) + \phi'(t)t} \leq -M. \qquad (4.2)$$

Proof The smooth function $\psi(t) = \frac{1-e^{-(M+1)t}}{(M+1)t}$ satisfies the following

$$\psi(t) + t\psi'(t) = (t\psi(t))' = e^{-(M+1)t}$$
$$2\psi'(t) + t\psi''(t) = (\psi(t) + t\psi'(t))' = -(M+1)e^{-(M+1)t}.$$

So $\psi(t)$ satisfies (4.2). Let $\psi(t) = \sum_{k=0}^{\infty} a_k t^k$ be the power series expansion, and let $\psi_N(x) = \sum_{k=0}^{N} a_k t^k$ be the truncated summation. Note that ψ_N converges to ψ uniformly on $[0,1]$. For arbitrarily small $\varepsilon > 0$, we can choose N big enough such that for all $t \in [0,1]$

$$|\psi_N(t) - \psi(t)| < \varepsilon, \qquad |\psi_N'(t) - \psi'(t)| < \varepsilon, \qquad |\psi_N''(t) - \psi''(t)| < \varepsilon.$$

Then the polynomial $\phi(t) = \psi_N(t)$ satisfies (4.2) when N is big enough. $\square$

Proof of Proposition 10 Let $\phi(t)$ be a polynomial satisfying (4.2) and $h_i(x) = \phi(g_i(x))$, which is positive on S, since $g_i(S) \subseteq [0,1]$. Then a direct calculation shows for $p_i(x) = g_i(x)h_i(x)$

$$-\nabla^2(p_i(x)) = -\left(\phi(g_i) + \phi'(g_i)g_i\right)\nabla^2 g_i(x) + \left(2\phi'(g_i) + \phi''(g_i)g_i\right)\nabla g_i(x)\nabla g_i(x)^T$$
$$= \left(\phi(g_i) + \phi'(g_i)g_i\right)\left(-\nabla^2 g_i(x) - \frac{2\phi'(g_i) + \phi''(g_i)g_i}{\phi(g_i) + \phi'(g_i)g_i}\nabla g_i(x)\nabla g_i(x)^T\right).$$

If M in (4.2) is chosen big enough, by Lemma 11 and Lemma 13, the negative Hessian $-\nabla^2(p_i(x))$ must be positive definite for all $x \in S$. $\square$

Remark: From the proof, we can see that both Proposition (10) and Lemma 4.2 remain true if S is replaced by any compact set Ω which is not convex or even connected.

4.2. Convex sets with boundary having positive curvature

This subsection ultimately shows that such an extendable poscurv-convex set has a very well behaved defining function. Recall S_i is extendable poscurv-convex with respect to S if $g_i(x) > 0$ whenever $x \in S$ lies in the interior of S_i and there exists a poscurv-convex set $T_i \supseteq S$ such that $\partial T_i \cap S = \partial S_i \cap S$. First we give a result which says a poscurv-convex set can be defined by a strictly concave smooth function.

Proposition 14 *Suppose T_i is a poscurv-convex set with the origin in the interior. Then there is a function $G_i(x) \in C^2(\mathbb{R}^n) \cap C^\infty(\mathbb{R}^n - \{0\})$ such that $T_i = \{x \in \mathbb{R}^n : G_i(x) \geq 0\}$, $\partial T_i = \{x \in \mathbb{R}^n : G_i(x) = 0\}$, $\nabla G_i(x) \neq 0$ for all $x \in \partial T_i$, and $G_i(x)$ has negative definite Hessian on T_i.*

Proof Our strategy is to build a concave defining function for T_i and then to approximate it by a concave smooth function. This takes several steps. Since T_i is compact, for any $0 \neq x \in \mathbb{R}^n$, there exists a unique positive scalar $\alpha(x)$ such that $\frac{1}{\alpha(x)}x =: r(x)$ lies on the boundary ∂T_i. Define $\alpha(0) = 0$. Indeed α is the classical Minkowski function ([4]), and $T_i = \{x \in \mathbb{R}^n : \alpha(x) \leq 1\}$. The function $\alpha(x)$ is convex. Note we can write $x = \alpha(x)r(x)$ and $\alpha(x)$ is smooth at $x \neq 0$, because the boundary ∂T_i is smooth.

Let $\tilde{G}(x) = 1 - \alpha(x)^3$. Thus $\tilde{G}(x)$ is a concave function and is smooth everywhere except at 0. Moreover, the super level sets satisfy

$$\left\{x: \ \tilde{G}(x) \geq c\right\} = \left\{x: \ 1 \geq \alpha\left(\frac{x}{\sqrt[3]{1-c}}\right)\right\} = \left\{x: \ \frac{x}{\sqrt[3]{1-c}} \in T_i\right\} = \sqrt[3]{1-c}\,T_i$$

for all $0 \leq c < 1$. Since ∂T_i has positive curvature, $\sqrt[3]{1-c}\,\partial T_i$ also has positive curvature. In summary, the function $\tilde{G}$ is concave, strictly quasi-concave and smooth except at $x = 0$. However, we need a function that is twice continuously differentiable on T_i and has negative definite Hessian there. The following produces one.

Claim: $G_i(x) := (1 - \epsilon\|x\|^2)[1 - (\alpha(x))^3] \in C^2(\mathbb{R}^n) \cap C^\infty(\mathbb{R}^n - \{0\})$ *has negative definite Hessian on T_i when ϵ is small enough.*

Proof of the Claim. Let $\psi(t) := 1 - t^3$ and then $\tilde{G} := \psi \circ \alpha$. So at $x \neq 0$

$$\nabla \tilde{G} = \psi'(\alpha)\nabla\alpha, \qquad \nabla^2 \tilde{G} = \psi''(\alpha)\nabla\alpha\nabla\alpha^T + \psi'(\alpha)\nabla^2\alpha. \qquad (4.3)$$

Note that $\nabla\tilde{G}(x) \neq 0$ for all $x \in \partial T_i$, since ∂T_i is smooth. Now we use the above to prove at $x \neq 0$ the Hessian $\nabla^2\tilde{G}(x)$ is negative definite. Obviously for $0 \neq x \in T_i$

$$\nabla\alpha\nabla\alpha^T \succeq 0, \qquad \psi''(\alpha) < 0, \qquad \nabla^2\alpha \succeq 0, \qquad \psi'(\alpha) < 0$$

and (4.3) has the form of the modified Hessian. Thus part (b) of Lemma 11 implies $\nabla^2\tilde{G}(x)$ is negative definite at $0 \neq x \in T_i$. From $x = \alpha(x)r(x)$, we have

$$|\alpha(x)| = \frac{|x^T r(x)|}{\|r(x)\|^2} \leq \frac{\|x\|}{\|r(x)\|}.$$

For $x \neq 0$, $r(x)$ is on the boundary ∂T_i and hence $\|r(x)\| \geq \delta$ for some constant $\delta > 0$. Thus $\alpha(x) = \mathcal{O}(\|x\|)$ and then $\alpha(x)^3 = \mathcal{O}(\|x\|^3)$. So we can see $\alpha(x)^3$ is at least twice differentiable at the origin; its gradient and Hessian vanish there, and so do those of $\tilde{G}(x)$. The function $\tilde{G}$ has negative definite Hessian except at $x = 0$. Obviously $G_i(x) \in C^2(\mathbb{R}^n) \cap C^\infty(\mathbb{R}^n - \{0\})$ and $\nabla G(x) \neq 0$ for all $x \in \partial T_i$.

To achieve strict concavity at 0 take $G_i(x) := (1 - \epsilon\|x\|^2)\tilde{G}(x)$. Then

$$\nabla^2 G_i(x) = (1 - \epsilon\|x\|^2)\nabla^2\tilde{G}(x) - 2\epsilon\left(\nabla\tilde{G}(x)x^T + x\nabla\tilde{G}(x)^T\right) - 2\epsilon\tilde{G}(x)I_n.$$

At $x = 0$, $\nabla^2 G_i(0) = -2\epsilon$. Thus for $\epsilon > 0$ the Hessian of G_i at $x = 0$ is negative definite. We can take ϵ small enough to keep the Hessian of G_i negative definite on the compact set T_i away from $x = 0$ as well, which completes the proof of the claim.

Obviously, $x \in T_i$ if and only if $G_i(x) \geq 0$, and $x \in \partial T_i$ if and only if $G_i(x) = 0$. $\square$

Lemma 15 *Assume $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \cdots, g_m(x) \geq 0\}$ is compact convex and $S_i = \{x \in \mathbb{R}^n : g_i(x) \geq 0\}$ is extendable poscurv-convex with respect to S. Then we have*

(i) $\nabla g_i(x)$ does not vanish on the boundary $\partial S_i \cap S$, and hence $\partial S_i \cap S$ is smooth.
(ii) Let $G_i(x)$ be the defining function for T_i given by Proposition 14. Then

$$w(x) := \frac{G_i(x)}{g_i(x)} \in C^2(S) \cap C^\infty(\partial S), \qquad and \qquad w(x) > 0, \forall x \in S.$$

Proof (i) We show $\nabla g_i(x) \neq 0$ on the boundary $\partial S_i \cap S$. For a contradiction, suppose $\nabla g_i(u) = 0$ for some $u \in \partial S_i \cap S$. Since ∂S_i has positive curvature, we have $-\nabla^2 g_i(u) \succ 0$. By continuity, $-\nabla^2 g_i(x) \succ 0$ when $x \in B(u, \epsilon)$ for some $\epsilon > 0$. Since S is convex and has nonempty interior, there exists $v \in B(u, \epsilon)$ in the interior of S. Thus

$$g_i(v) < g_i(u) + \nabla g_i(u)^T(v - u) = g_i(u) = 0.$$

which contradicts $g_i(v) \geq 0$ since $v \in S$.

(ii) By assumption, let $T_i \supseteq S$ be a poscurv-convex set such that $\partial T_i \cap S = \partial S_i \cap S$; thus and ∂T_i is nonsingular and has positive curvature. Without loss of generality, assume the origin is in the interior of S. Then apply Proposition 14 to T_i and get a concave defining function for T_i such that $G_i(x) \in C^2(\mathbb{R}^n) \cap C^\infty(\partial S)$ and it has a negative definite Hessian on $T_i \supseteq S$. Similarly we can prove $\nabla G_i(x) \neq 0$ on the boundary ∂T_i.

Now we need to show $w(x) = \frac{G_i(x)}{g_i(x)}$ is positive on S and belongs to $C^2(S) \cap C^\infty(\partial S)$. Obviously it is smooth in the interior or exterior of S_i except at $x = 0$, and twice differentiable at $x = 0$. We need to show w is smooth on the boundary $\partial S_i \cap S$. Now fix a $u \in \partial S_i \cap S$. Since $\nabla g_i(u) \neq 0$, we can find a local coordinate transformation $x - u = t\nabla g_i(u) + By$ to new coordinates (t, y) in $\mathbb{R} \times \mathbb{R}^{n-1}$. Here B is a matrix such that $\nabla g_i(u)^T B = 0$ and $[\nabla g_i(u)\ B]$ is invertible. The point u corresponds to $(0, 0)$ in the new coordinate. Then apply the Taylor series expansion at point u and

get $G_i(x) = G_i(t,y) = \sum_{k=1}^{\infty} a_k(y)t^k$ and $g_i(x) = g_i(t,y) = \sum_{k=1}^{\infty} b_k(y)t^k$ for some smooth scalar functions $a_k(y), b_k(y)$. The fact $\nabla g_i(u) \neq 0$ and $\nabla G_i(u) \neq 0$ implies $a_1(0) \neq 0$ and $b_1(0) \neq 0$. Thus we can see

$$w(x) = \frac{a_1(y) + \sum_{k=2}^{\infty} a_k(y)t^{k-1}}{b_1(y) + \sum_{k=2}^{\infty} a_k(y)t^{k-1}}$$

is smooth at u. Note that $a_1(0)$ and $b_1(0)$ are directional derivatives in the direction of gradients. Since the boundary $\partial S_i \cap S$ is defined equivalently both by $G_i(x) = 0$ and $g_i(x) = 0$ near u, the functions $G_i(x)$ and $g_i(x)$ must have parallel gradients in the same direction at $u \in \partial S_i \cap S$. So $a_1(0)/b_1(0) > 0$ and hence $w(u) > 0$. Obviously $w(x) > 0$ for interior points x of S_i in S. $\square$

The above lemma shows the product $g_i(x)w(x)$ has negative definite Hessian on S. Unfortunately, $w(x)$ might not be a polynomial. However, we can use polynomials to approximate $w(x)$ and its derivatives. Thus we need an improved version of *Stone-Weierstrass Approximation Theorem* which shows the density of polynomials in the space $C^k(\Omega)$ for a bounded open set Ω. Define the norm in $C^k(\Omega)$ as

$$\|f\|_{C^k(\Omega)} := \max_{x \in \Omega} \max_{\alpha \in \mathbb{N}^n, 0 \leq |\alpha| \leq k} \{|D^\alpha f(x)|\}.$$

Proposition 16 *Suppose $f \in C^k(\Omega)$ is supported in a bounded open set Ω in R^n. For any $\epsilon > 0$, there exists a polynomial h such that $\|f - h\|_{C^k(\Omega)} < \epsilon$.*

The basic idea for proving this theorem is that $C^\infty(\Omega)$ is dense in $C^k(\Omega)$, which contains f, and then polynomials are dense in $C^\infty(\Omega)$. The proof is straightforward, for example, it is an exercise in Hirsch [11, Chapter 2]. Thus we omit the proof here.

Proposition 17 *Assume $S_i = \{x \in \mathbb{R}^n : g_i(x) \geq 0\}$ is extendable poscurv-convex with respect to S. Then there exists a polynomial $h_i(x)$ positive on S such that the product $p_i(x) := g_i(x)h_i(x)$ has negative definite Hessian on S.*

Proof Let $T_i \supseteq S$ be a compact convex set such that ∂T_i is nonsingular and $\partial T_i \cap S = \partial S_i \cap S$. Then apply Proposition 14 to T_i and get a concave defining function $G_i(x)$ for T_i with negative definite Hessian on $T_i \supseteq S$. Lemma 15 shows $w(x) = \frac{G_i(x)}{g_i(x)} \in C^2(S)$ is positive on S. So $w(x) \in C^2(U)$ for some bounded open set U containing S. Extend $w(x)$ to the whole space $\mathbb{R}^n$ such that $w(x) = 0$ for all $x \notin U$. Let $w_\epsilon(x)$ be the mollified function

$$w_\epsilon(x) = \int \frac{1}{\epsilon^n} \eta\left(\frac{x-y}{\epsilon}\right) w(y) dy$$

where $\eta(x)$ is the *standard mollifier* function

$$\eta(x) = \begin{cases} ae^{\frac{1}{\|x\|^2 - 1}} & \text{if } \|x\| < 1 \\ 0 & \text{if } \|x\| \geq 1 \end{cases}.$$

Here the constant a is chosen to make $\int_{\mathbb{R}^n} \eta(x)dx = 1$. The function $w_\epsilon(x)$ is a smooth function supported in a bounded open set $U' \supseteq U \supseteq S$. Also $w(x)$ and $w_\epsilon(x)$ are both twice differentiable on S, and $\|w_\epsilon(x) - w(x)\|_{C^2(S)}$ can be made arbitrarily small by sending $\epsilon \to 0$.

Note that $G_i = g_i w$ is a concave function such that $-\nabla^2 G_i(x) \succ 0$ for all $x \in S$. Obviously

$$\nabla^2 G_i(x) = w(x)\nabla^2 g_i(x) + \nabla g_i(x)\nabla w(x)^T + \nabla w(x)^T \nabla g_i(x) + g_i(x)\nabla^2 w(x).$$

By Proposition 16, for arbitrary $\tau > 0$, there exists a polynomial $h_i(x)$ such that

$$\|w_\epsilon(x) - h_i(x)\|_{C^2(\Omega)} < \tau.$$

If ϵ and $\tau > 0$ are small enough, then $h_i(x)$ is positive on S and the product $p_i(x) = g_i(x)h_i(x)$ has negative definite Hessian on S. $\square$

A simpler result on new defining polynomials which requires less terminology to understand is:

Corollary 18 *Given a polynomial $g(x)$, if $T = \{x \in \mathbb{R}^n : g(x) \geq 0\}$ is a poscurv-convex set with nonempty interior, then there is a polynomial $p(x)$ strictly concave on T satisfying $p(x) = 0, \nabla p(x) \neq 0$ for $x \in \partial T$ and $p(x) > 0$ for x inside of T.*

Proof Obviously, T is extendable poscurv-convex with respect to itself. By Proposition 17, there exists a polynomial $h(x)$ positive on T such that the product polynomial $p(x) = g(x)h(x)$ has negative definite Hessian on T. If $x \in \partial T$, then $p(x) = 0$. If x is in the interior of T, then $p(x) > 0$. By an argument similar to that for part (i) of Lemma 15, $\nabla p(x)$ does not vanish on the boundary $\partial T = \{x \in \mathbb{R}^n : g(x) = 0\}$. $\square$

5. Proofs

As we have seen, the projections of the sets $\hat{S}_N$ defined by LMI (2.2) contain the convex set S, for all integers $N \geq \max_\nu d_\nu$. We need to prove that the projection actually equals S for some sufficiently large N. The basic idea of the proof of this sharpness is to apply the Convex Set Separating Theorem to produce a linear functional which is nonnegative on S and negative on a given point outside S. We need to prove Schmüdgen's or Putinar's representations for such linear functionals with uniform degree bounds. The uniform degree bounds will be proved in §6, but will be used in this section. They are Theorems 27 and 29.

5.1. Proofs of Theorem 5 and Theorem 6

Given a unit length vector $\ell \in \mathbb{R}^n$, consider the optimization problem

$$\ell^* := \min_{x \in \mathbb{R}^n} \quad \ell^T x$$
$$s.t. \quad g_1(x) \geq 0, \cdots, g_m(x) \geq 0.$$

Let $u = u(\ell) \in S$ denote the minimizer, which must exist due to the compactness of S. Note that u must be on the boundary ∂S.

Suppose $g_i(x)$ are concave on S and S has non-empty interior, i.e., there exists $\xi \in S$ such that $g_1(\xi) > 0, \cdots, g_m(\xi) > 0$. So the Slater's condition holds, which implies that there exist nonnegative Lagrange multipliers $\lambda_1, \cdots, \lambda_m \geq 0$ such that

$$\ell = \sum_{i=1}^m \lambda_i \nabla g_i(u), \quad \lambda_i g_i(u) = 0, \forall i = 1, \cdots, m.$$

So $f_\ell(x) = \ell^T x - \ell^* - \sum_{i=1}^m \lambda_i \nabla g_i(x)$ is a convex function on S such that $f_\ell(u) = 0$ and $\nabla f_\ell(u) = 0$. Hence for all $x \in S$ we have

$$f_\ell(x) \geq f_\ell(u) + \nabla f_\ell(u)^T (x - u) = 0.$$

So $f_\ell(x)$ is nonnegative on S. We hope to find a Schmüdgen's or Putinar's representation of f_ℓ in terms of polynomials $g_1(x), \cdots, g_m(x)$. But we want the representation to have a uniform degree bound. Indeed, validating the lifted construction in §2 amounts to proving that there is a N, such that for all $\|\ell\| = 1$ the polynomials in the resulting representation of $f_\ell(x)$ have degree at most $2N$.

Lemma 19 *Use the above notations. Suppose S has non empty interior and its defining polynomials $g_i(x)$ are concave on S. Suppose either $-\nabla^2 g_i(x)$ is SOS or $-\nabla^2 g_i(u) \succ 0$ for all $u \in \partial S_i \cap \partial S$. Then for every unit length vector ℓ we have the representation*

$$f_\ell(x) = \sum_{i=1}^m \lambda_i \ (x - u)^T F_i(u, x)(x - u)$$

where u is the minimizer, $\lambda_i \geq 0$ are the Lagrange multipliers, and $F_i(u, x)$ is SOS in x or such that

$$\delta I_n \preceq F_i(u, x) \preceq M I_n, \quad \forall x \in S \tag{5.1}$$

for some positive constants $M > \delta > 0$ which are independent of ℓ.

Proof Since $f_\ell(u) = 0$ and $\nabla f_\ell(u) = 0$, we get

$$f_\ell(x) = \sum_{i=1}^m \lambda_i\,(x-u)\underbrace{\left(\int_0^1\int_0^t -\nabla^2 g_i(u+s(x-u))\,ds dt\right)}_{F_i(u,x)}(x-u).$$

Let $J(u) = \{1 \le i \le m : g_i(u) = 0\}$ be the index set of active constraints. For $i \notin J(u)$, the Lagrange multiplier $\lambda_i = 0$, so we can choose $F_i(u,x)$ to be the zero matrix which is of course SOS. Note that for all $i \in J(u)$, $u \in \partial S_i \cap \partial S$.

If $-\nabla^2 g_i(x)$ is SOS, then $F_i(u,x)$ is also SOS in x by Lemma 7. If $-\nabla^2 g_i(x)$ is not SOS but positive definite on the boundary $\partial S_i \cap \partial S$, then $F_i(u,x)$ must be positive definite on S. To see this point, we first show that $F_i(u,x)$ is positive semidefinite. For any $u \in \partial S_i \cap \partial S$, $x \in S$, the line segment $\{u+s(x-u) : 0 \le s \le 1\}$ is contained in S and g_i are concave, so

$$F_i(u,x) = \int_0^1\int_0^t -\nabla^2 g_i(u+s(x-u))\,ds dt \succeq 0.$$

Second, we show $F_i(u,x)$ is positive definite. Suppose for some vector ξ,

$$\xi^T F_i(u,x)\xi = \int_0^1\int_0^t \xi^T\left(-\nabla^2 g_i(u+s(x-u))\right)\xi ds dt = 0.$$

By the concavity of g_i, we must have

$$\xi^T\left(-\nabla^2 g_i(u+s(x-u))\right)\xi = 0, \quad \forall s \in [0,1].$$

Choose $s = 0$ in the above, then $-\nabla^2 g_i(u) \succ 0$ implies $\xi = 0$. Hence $F_i(u,x) \succ 0$ for all $x \in S$ and $u \in \partial S_i \cap S$. Now we need show $F_i(u,x)$ satisfies the inequality (5.1). Obviously, by definition, as a function of u and x, $F_i(u,x)$ is continuous in $u \in \partial S_i \cap \partial S$ and $x \in S$. And $F_i(u,x)$ is positive definite for all $u \in \partial S_i \cap S$ and $x \in S$. Since the minimum eigenvalue is a continuous function of the matrix, the existence of constants $M > \delta > 0$ independent of ℓ is due to the compactness of S. $\square$

Theorem 20. *Assume polynomials $g_i(x)$ satisfy the hypotheses of Theorem 5. Then there exists a finite integer N such that for every vector ℓ with $\|\ell\| = 1$*

$$f_\ell(x) = \sum_{\nu\in\{0,1\}^m} \sigma_\nu(x)g_1^{\nu_1}(x)\cdots g_m^{\nu_m}(x)$$

where $\sigma_\nu(x)$ are SOS polynomials with degree

$$\deg(\sigma_\nu g_1^{\nu_1}\cdots g_m^{\nu_m}) \le 2N.$$

Furthermore, if the archimedean condition on the g_i holds, then $f_\ell(x)$ has the representation

$$f_\ell(x) = \sigma_0(x) + \sigma_1(x)g_1(x) + \cdots + \sigma_m(x)g_m(x)$$

with degree bounds such that $\deg(\sigma_i g_i) \le 2N$.

Proof Let N be the maximum integer such that $\max_i \deg(g_i) \le 2(N-1) \le \Omega(\frac{M}{\delta})$ where function $\Omega(\cdot)$ is given by Theorem 27, and M, δ are given by (5.1). Fix an arbitrary vector ℓ and let u be the minimizer of $\ell^T x$ on ∂S. By Lemma 19

$$f_\ell(x) = \sum_{i=1}^m \lambda_i(x-u)^T F_i(u,x)(x-u)$$

holds for matrix polynomials $F_i(u,x)$ which are either SOS in x or such that

$$\delta I_n \preceq F_i(u,x) \preceq M I_n$$

with some positive constants $M > \delta > 0$ which are independent of ℓ. Let $K = \{1 \le i \le m : F_i(u,x) \text{ is SOS }\}$. If $i \in K$, then $(x-u)^T F_i(u,x)(x-u)$ is an SOS polynomial of degree at most $\deg(g_i)$. If $i \notin K$, by Theorem 27, there exist SOS matrix polynomials $G_\nu^{(i)}(x)$ such that

$$F_i(u,x) = \sum_{\nu \in \{0,1\}^m} G_\nu^{(i)}(x) g_1^{\nu_1}(x) \cdots g_m^{\nu_m}(x)$$

with degree bounds $\deg(G_\nu^{(i)} g_1^{\nu_1} \cdots g_m^{\nu_m}) \le 2N - 2$. Now let

$$\sigma_0(x) = \sum_{i \in K} \lambda_i (x-u)^T F_i(u,x)(x-u) + \sum_{i \notin K} \lambda_i (x-u)^T G_0^{(i)}(x)(x-u)$$

$$\sigma_\nu(x) = \sum_{i \notin K} \lambda_i (x-u)^T G_\nu^{(i)}(x)(x-u), \qquad \text{if } \nu \ne 0.$$

which must also be SOS polynomials such that $\deg(\sigma_\nu g_1^{\nu_1} \cdots g_m^{\nu_m}) \le 2N$. So we have the Schmüdgen representation for f_ℓ with uniform (in ℓ) degree bounds.

Similarly, Putinar's representation for f_ℓ with uniform degree bounds follows from Theorem 29. $\square$

Now we are able to complete the proofs of Theorem 5 and 6. The basic idea for the proof is as follows. Theorem 20 essentially guarantees that the so-called S-BDNR and PP-BDNR (under archimedean condition) properties mentioned in Section 2 hold for S, which implies the S-BDR and PP-BDR properties also hold for S, and thus the results in [13] can be applied to validate the exactness of the constructed LMIs for Theorems 5 and 6.

Proof of Theorem 5 First, we prove the S-BDNR property holds for S for integer N claimed by Theorem 20. Let $a^T x + b$ be nonnegative on S and u be a minimizer of $a^T x + b$ on S. Since S has nonempty interior, the Slater's condition holds, that is, there exist Lagrange multipliers $\lambda_1, \cdots, \lambda_m \ge 0$ such that

$$a = \sum_{i=1}^m \lambda_i \nabla g_i(u), \quad \lambda_i g_i(u) = 0, \quad \forall i = 1, \cdots, m.$$

Applying Theorem 20 for $\ell = a$ and $\ell^* = a^T u$, we get the representation

$$a^T x - a^T u - \sum_i \lambda_i g_i(x) = \sum_{\nu \in \{0,1\}^m} \sigma_\nu'(x) g_1^{\nu_1}(x) \cdots g_m^{\nu_m}(x)$$

for some SOS polynomials σ_ν' with degree bounds

$$\deg(\sigma_\nu' g_1^{\nu_1} \cdots g_m^{\nu_m}) \le 2N.$$

Or equivalently we have the identity (note that $a^T u + b \ge 0$)

$$a^T x + b = \sum_{\nu \in \{0,1\}^m} \sigma_\nu(x) g_1^{\nu_1}(x) \cdots g_m^{\nu_m}(x)$$

for some SOS polynomials σ_ν with the same degree bounds as σ_ν'. Thus the S-BDNR property holds for S, and so does S-BDR.

Second, we prove that the LMI (2.2) $\hat{S}$ constructed in §2 is a lifted LMI of S for integer N claimed by Theorem 20. Since the S-BDNR property implies the S-BDR property for S, Theorem 2 in Lasserre [13] can be applied to validate the exactness of the lifted LMI (2.2). For the convenience of readers, we give the direct proof here, because the proof is short and the approach will be used in proving Theorems 2 and 4 (these theorems can not be shown by only proving the S-BDR or PP-BDR property, since their lifted LMIs are not purely based on Schmüdgen's or Putinar's representation).

Obviously, the set S is contained in the projection down of each LMI $\hat{S}_N$ defined by (2.2). We show they are equal. Otherwise, in pursuit of a contradiction, suppose there exists a vector $(\hat{x}, \hat{y})$ in

$\hat{S}$ such that $\hat{x}$ is not in the convex set S. By the Hahn-Banach Separation Theorem, there must exist a vector ℓ of unit norm such that

$$\ell^T \hat{x} < \ell^* := \min_{x \in S} \ell^T x. \tag{5.2}$$

So $\ell^T x - \ell^*$ is nonnegative on S. By the S-BDNR property, we have the identity

$$\ell^T x - \ell^* = \sum_{\nu \in \{0,1\}^m} \sigma_\nu(x) g_1^{\nu_1}(x) \cdots g_m^{\nu_m}(x) \tag{5.3}$$

for some SOS polynomials σ_ν with degree bounds $\deg(\sigma_\nu g_1^{\nu_1} \cdots g_m^{\nu_m}) \leq 2N$. Since σ_ν is SOS, we can write $\sigma_\nu(x) = [x^{d-d_\nu}]^T W_\nu [x^{d-d_\nu}]$ for some symmetric matrix $W_\nu \succeq 0$. Now in identity (5.3), replace each monomial x^α by $\hat{y}_\alpha$, we get

$$\ell^T \hat{x} - \ell^* = \sum_{\nu \in \{0,1\}^m} Trace\Big(W_\nu \cdot \big(\sum_{0 \leq |\alpha| \leq 2N} A_\alpha^\nu \hat{y}_\alpha \big) \Big) \geq 0,$$

which contradicts (5.2). $\square$

Proof of Theorem 6   We can prove Theorem 6 in a very similar manner to the proof above of Theorem 5. Here we only show the distinctive parts.

First, by the archimedean condition and Theorem 20, we can prove the PP-BDNR property holds for S for integer N claimed by Theorem 20, that is, for any affine polynomial $a^T x + b$ nonnegative on S, we have the identity (we denote $g_0(x) = 1$)

$$a^T x + b = \sum_{k=0}^{m} \sigma_k(x) g_k(x)$$

for some SOS polynomials σ_i with degree bounds $\deg(\sigma_i g_i) \leq 2N$. This implies the PP-BDR property also holds for S.

Second, since PP-BDR property holds for S, Theorem 2 in Lasserre [13] (same argument as above) can be applied to validate the exactness of the lifted LMI (2.4). Here we directly give the proof by contradiction. Follow the same contradiction approach we have done in the proof of Theorem 5. Suppose there exists $(\hat{x}, \hat{y})$ in $\hat{S}$ such that $\hat{x} \notin S$. Then there must exist $\ell \in \mathbb{R}^n$ of unit norm and ℓ^* such that

$$\ell^T \hat{x} < \ell^* := \min_{x \in S} \ell^T x.$$

So $\ell^T x - \ell^*$ is nonnegative on S. Then the PP-BDNR property implies the identity

$$\ell^T x - \ell^* = \sum_{k=0}^{m} \sigma_k(x) g_k(x)$$

for some SOS polynomials σ_i with degree bounds $\deg(\sigma_i g_i) \leq 2N$. Since σ_i is SOS, we can write $\sigma_i(x) = [x^{d-d_\nu}]^T W_i [x^{d-d_\nu}]$ for some symmetric matrix $W_i \succeq 0$. By substituting $\hat{y}_\alpha$ for each x^α in the above identity, we get

$$\ell^T \hat{x} - \ell^* = \sum_{k=0}^{m} Trace\Big(W_i \cdot \big(\sum_{0 \leq |\alpha| \leq 2N} A_\alpha^{(k)} \hat{y}_\alpha \big) \Big) \geq 0,$$

which results in the contradiction $0 > \ell^T \hat{x} - \ell^* \geq 0$. $\square$

5.2. Proof of Theorem 1

In this subsection, we assume S is convex, compact and has nonempty interior, and $g_i(x)$ are concave on S. Then Slater's condition holds and the Lagrange multipliers exist for the linear objective $\ell^T x$. For unit length vectors $\ell \in \mathbb{R}^n$, let $f_\ell(x), \ell^*, u, \lambda_i$ be defined as before.

Lemma 21 *Assume the PDLH condition holds, then there exist constants $M > \delta > 0$ such that*

$$\delta I_n \preceq \underbrace{\int_0^1 \int_0^t \left(-\sum_{i=1}^m \lambda_i \nabla^2 g_i(u + s(x - u)) \right) ds dt}_{L(u,x)} \preceq M I_n$$

for every unit length vector ℓ.

Proof Let $\xi \in S$ be a fixed interior point. Note that

$$f_\ell(x) = \ell^T x - \ell^* - \sum_{i=1}^m \lambda_i g_i(x) \geq 0, \quad \forall\, x \in S.$$

Choose $x = \xi$ in the above, then we have

$$0 \leq \lambda_i \leq \frac{\ell^T x - \ell^*}{g_i(\xi)} = \frac{\ell^T x - \ell^T u}{g_i(\xi)} \leq \frac{D}{g_i(\xi)}$$

where D is the diameter of S. So $\max_i \lambda_i \leq \frac{D}{\min_i g_i(\xi)}$. Thus λ_i are uniformly bounded and hence the existence of M is obvious. Since $g_i(x)$ are concave on S, we have that $L(u, x)$ must be positive semidefinite on S. We need to show δ exists. Otherwise, in pursuit of a contradiction, suppose we have a sequence $\{\ell^{(k)}\}$, $\{u^{(k)}\}$, $\{x^{(k)}\}$, and $\{\lambda^{(k)}\}$ such that $\lambda_{\min}(L(u^{(k)}, x^{(k)})) \to 0$. Since $\{\ell^{(k)}\}, \{u^{(k)}\}, \{x^{(k)}\}, \{\lambda^{(k)}\}$ are all bounded, without loss of generality, we can assume

$$\ell^{(k)} \to \hat{\ell}, \qquad u^{(k)} \to \hat{u}, \qquad x^{(k)} \to \hat{x}, \qquad \lambda^{(k)} \to \hat{\lambda}.$$

The limit $\hat{\ell}$ also has unit length, $\hat{u}$ is the minimizer of $\hat{\ell}^T x$ on S and $\hat{\lambda}_i$ are the corresponding Lagrange multipliers. That the limit $L(u^{(k)}, x^{(k)})$ is singular implies there exists $0 \neq \zeta \in \mathbb{R}^n$ such that

$$\int_0^1 \int_0^t \zeta^T \left(-\sum_{i=1}^m \hat{\lambda}_i \nabla^2 g_i(\hat{u} + s(\hat{x} - \hat{u})) \right) \zeta\, ds dt = 0.$$

Then we must get

$$\zeta^T \left(-\sum_{i=1}^m \hat{\lambda}_i \nabla^2 g_i(\hat{u} + s(\hat{x} - \hat{u})) \right) \zeta = 0, \quad \forall\, 0 \leq s \leq 1.$$

Choose $s = 0$ in the above. But the PDLH condition implies $\zeta = 0$, which is a contradiction. $\square$

Theorem 22. *Assume the hypotheses of Theorem 1. Then there exists a finite integer N such that for every unit length vector ℓ with $\|\ell\| = 1$*

$$f_\ell(x) = \sum_{\nu \in \{0,1\}^m} \sigma_\nu(x) g_1^{\nu_1}(x) \cdots g_m^{\nu_m}(x)$$

where $\sigma_\nu(x)$ are SOS polynomials with degree

$$\deg(\sigma_\nu g_1^{\nu_1} \cdots g_m^{\nu_m}) \leq 2N.$$

Furthermore, if the archimedean condition holds for polynomials $g_1, \cdots, g_m$, then $f_\ell(x)$ has the representation

$$f_\ell(x) = \sigma_0(x) + \sigma_1(x) g_1(x) + \cdots + \sigma_m(x) g_m(x)$$

with degree bounds such that $\deg(\sigma_i g_i) \leq 2N$.

Proof Let N be the maximum integer such that $2(N-1) \leq \Omega(\frac{M}{\delta})$ where the function $\Omega(\cdot)$ is given by Theorem 27, and M, δ are given by the preceding lemma. Fix an arbitrary unit length vector ℓ. Since $f_\ell(u) = 0$ and $\nabla f_\ell(u) = 0$, we get

$$f_\ell(x) = (x - u)^T \underbrace{\int_0^1 \int_0^t \left(-\sum_{i=1}^m \lambda_i \nabla^2 g_i(u + s(x - u))\right) dsdt}_{L(u,x)} (x - u).$$

with $\delta I_n \preceq L(u, x) \preceq M I_n$. By Theorem 27, there exist SOS matrix polynomials $G_\nu^{(i)}(x)$ such that

$$L(u, x) = \sum_{\nu \in \{0,1\}^m} G_\nu(x) g_1^{\nu_1}(x) \cdots g_m^{\nu_m}(x)$$

with degree bounds $\deg(G_\nu g_1^{\nu_1} \cdots g_m^{\nu_m}) \leq 2N - 2$. Now let $\sigma_\nu(x) = (x - u)^T G_\nu(x)(x - u)$, which must also be SOS polynomials such that $\deg(\sigma_\nu g_1^{\nu_1} \cdots g_m^{\nu_m}) \leq 2N$, then the first part of the theorem holds.

The second part of the theorem can proved by applying Theorem 29 in a similar way. $\square$

Proof of Theorem 1 We claim that LMI (2.2) is an SDP representation of S when N is sufficiently large. The proof is very similar to what we have done for proving Theorem 5.

First, as we have done in proving Theorem 5, Theorem 22 can be applied to show the S-BDNR property, which implies the S-BDR property holds for S. So, Theorem 2 in Lasserre [13] validates the exactness of the lifted LMI (2.2). Here we give the direct proof by contradiction, which is very similar to what we have done for Theorem 5. Here we only give the distinctive parts. Suppose there exists $(\hat{x}, \hat{y})$ in $\hat{S}$ such that $\hat{x} \notin S$. Then there must exist $\ell \in \mathbb{R}^n$ of unit norm and ℓ^* such that

$$\ell^T \hat{x} < \ell^* := \min_{x \in S} \ell^T x.$$

So $\ell^T x - \ell^*$ is nonnegative on S. The S-BDNR property implies that there exists an $N > 0$ such that

$$\ell^T x - \ell^* = \sum_{\nu \in \{0,1\}^m} \sigma_\nu(x) g_1^{\nu_1}(x) \cdots g_m^{\nu_m}(x).$$

for some SOS polynomials σ_ν with degree bounds $\deg(\sigma_\nu(x) g_1^{\nu_1}(x) \cdots g_m^{\nu_m}(x)) \leq 2N$. So we can write $\sigma_\nu(x) = [x^{d-d_\nu}]^T W_\nu [x^{d-d_\nu}]$ for some symmetric matrix $W_\nu \succeq 0$. In the above identity, similar to what we have done in the proof of Theorem 5, replacing each monomial x^α by $\hat{y}_\alpha$ results in the contradiction $0 > \ell^T \hat{x} - \ell^* \geq 0$.

Furthermore, if the archimedean condition holds for polynomials $g_1, \cdots, g_m$, then we can similarly prove that the LMI (2.4) is the lifted LMI for S when N is big enough, as we have done for proving Theorem 6. $\square$

5.3. Proof of Theorem 2

In this subsection, we no longer assume the defining polynomials $g_i(x)$ are concave on S but only that they are quasi-concave. The set S is still assumed to be convex, compact and have nonempty interior. The key point of our proof is to find a different set of concave polynomials defining the same convex set S.

Now we return to the proof of Theorem 2. Assume the hypotheses of Theorem 2 hold. If $-\nabla^2 g_i(x)$ is SOS in x, let $p_i(x) = g_i(x)$ which is obviously concave. If g_i is strictly quasi-concave on S, Proposition 10 implies that we can find new defining polynomials p_i that have negative definite Hessian on S. So in some open set U containing S, we have

$$\underbrace{\{x \in \mathbb{R}^n : p_1(x) \geq 0, \cdots, p_m(x) \geq 0\}}_{P} \cap U = S.$$

We should mention that the set P might not coincide with S, since it might be possible that for some point v far away from S such that $p_i(v) \geq 0$ for all $1 \leq i \leq m$. Let $y_\alpha = \int x^\alpha d\mu(x)$ be the α-th moment. Write $p_i(x) = \sum_\alpha p_\alpha^{(i)} x^\alpha$. If $\mathrm{supp}(\mu) \subseteq S$, then

$$\sum_\alpha p_\alpha^{(1)} y_\alpha \geq 0, \quad \cdots, \quad \sum_\alpha p_\alpha^{(m)} y_\alpha \geq 0.$$

Therefore, the set S is contained in the projection of the set $\widehat{S}_N$ of solutions to the following refined LMI

$$\left. \begin{array}{c} \forall \nu \in \{0,1\}^m, \quad A_0^\nu + \displaystyle\sum_{0 < |\alpha| \leq 2N} A_\alpha^\nu y_\alpha \succeq 0 \\ L_{p_1}(y) \geq 0, \cdots, L_{p_m}(y) \geq 0 \\ x_1 = y_{e_1}, \cdots, x_n = y_{e_n}, y_0 = 1 \end{array} \right\}. \tag{5.4}$$

Here $L_{p_i}(y) = \sum_\alpha p_\alpha^{(i)} y_\alpha$ and symmetric matrices A_α^ν are the same as those in LMI (2.2). Note the A_α^ν are determined by the g_i. So (5.4) uses both p_i and g_i.

Our goal is to prove (5.4) is a lifted LMI for S for sufficiently large N, thereby validating Theorem 2. For this purpose, we need a lemma similar to Lemma 19. For arbitrary unit length vector ℓ, consider optimization

$$\ell^* := \min_{x \in U} \quad \ell^T x$$
$$s.t. \quad p_1(x) \geq 0, \cdots, p_m(x) \geq 0,$$

which is the same as to minimize $\ell^T x$ on S. Let $u = u(\ell)$ be the minimizer, whose existence is guaranteed by the compactness of S. Note that S has an interior point $\xi \in S$, i.e., $g_i(\xi) > 0$ for all $i = 1, \cdots, m$. By our construction, $p_i(x) = h_i(x)g_i(x)$ for some polynomials $h_i(x)$ which are positive on S. So $\xi \in S$ is also an interior point for the new defining polynomials $p_1(x), \cdots, p_m(x)$, and hence the Slater's condition holds for the constraints $p_1(x) \geq 0, \cdots, p_m(x) \geq 0$. Therefore there exist Lagrange multipliers $\lambda = [\lambda_1 \cdots \lambda_m] \geq 0$ such that the function

$$\tilde{f}_\ell(x) := \ell^T x - \ell^* - \sum_{i=1}^m \lambda_i p_i(x)$$

is a nonnegative convex function on S such that $\tilde{f}_\ell(u) = 0$, $\nabla \tilde{f}_\ell(u) = 0$. Note that $-\nabla^2 p_i(x)$ is either SOS or positive definite on S.

Lemma 23 *Let $p_i(x), \tilde{f}_\ell(x), \ell^*, \lambda_i, u$ be defined as above. Then we have the representation*

$$\tilde{f}_\ell(x) = \sum_{i=1}^m \lambda_i (x-u)^T F_i(u,x)(x-u)$$

where the symmetric matrix polynomial $F_i(u,x)$ is either SOS or such that

$$\delta I_n \preceq F_i(u,x) \preceq M I_n, \quad \forall x \in S$$

for some positive constants $M > \delta > 0$ which are independent of ℓ.

Proof Since $\tilde{f}_\ell(u) = 0$, $\nabla \tilde{f}_\ell(u) = 0$, we have

$$\tilde{f}_\ell(u) = \sum_{i=1}^m \lambda_i (x-u)^T \underbrace{\left(-\int_0^1 \int_0^t \nabla^2 p_i(x + s(x-u) ds\, dt \right)}_{F_i(u,x)} (x-u).$$

If $p_i(x)$ is sos-concave, $F_i(u,x)$ is SOS in x by Lemma 7. If $p_i(x)$ is strictly concave on S, then we can prove $F_i(u,x)$ is positive definite on S. Apply the same argument in the proof for Lemma 19. $\square$

Theorem 24. *Assume polynomials $g_i(x)$ satisfy the hypotheses of Theorem 2. Then there exists a finite integer N such that for every vector ℓ with $\|\ell\| = 1$*

$$\tilde{f}_\ell(x) = \sum_{\nu \in \{0,1\}^m} \sigma_\nu(x) g_1^{\nu_1}(x) \cdots g_m^{\nu_m}(x)$$

where $\sigma_\nu(x)$ are sums of squares of polynomials with degree

$$\deg(\sigma_\nu g_1^{\nu_1} \cdots g_m^{\nu_m}) \leq 2N.$$

Furthermore, if the archimedean condition on the g_i holds, then $\tilde{f}_\ell(x)$ has the representation

$$\tilde{f}_\ell(x) = \sigma_0(x) + \sigma_1(x) g_1(x) + \cdots + \sigma_m(x) g_m(x)$$

with degree bounds such that $\deg(\sigma_i g_i(x)) \leq 2N$.

Proof The proof is almost the same as for Theorem 20. Just follow the argument for proving Theorem 20. The only differences are replacing $f_\ell(x)$ by $\tilde{f}_\ell(x)$ and then applying Lemma 23 instead of Lemma 19. $\square$

Proof of Theorem 2 If $-\nabla^2 g_i(x)$ is SOS in x, let $p_i(x) = g_i(x)$ which is obviously concave. If g_i is strictly quasi-concave on S, let $p_i(x)$ be the new defining polynomials for S_i given by Proposition 10, which have negative definite Hessian on S. For some small open set U containing S, the convex set S is equivalently defined as

$$S = \{x \in U : p_1(x) \geq 0, \cdots, p_m(x) \geq 0\}.$$

As we have seen earlier, S is contained in the projection of LMI (5.4). We claim that this projection is sharp for N given by Theorem 24. The proof is very similar to the proof for Theorem 5.

Otherwise, seeking a contradiction, suppose there exists a vector $(\hat{x}, \hat{y})$ in $\hat{S}_N$ such that $\hat{x} \notin S$. By the Hahn-Banach Separation Theorem, there must exist a vector ℓ of unit length such that

$$\ell^T \hat{x} < \ell^* := \min_{x \in S} \ell^T x. \tag{5.5}$$

Let $u \in S$ be the minimizer of $\ell^T x$ on S, which must be on the boundary ∂S. Note that $p_1(x), \cdots, p_m(x)$ are concave polynomials, and S has nonempty interior. Since $p_i(x) = h_i(x) g_i(x)$ for $h_i(x)$ positive on S, the new equivalent constraints $p_1(x) \geq 0, \cdots, p_m(x) \geq 0$ also have nonempty interior. Thus the Slater's condition holds and hence there exist Lagrange multipliers $\lambda_1, \cdots, \lambda_m \geq 0$ such that

$$\ell = \sum_{i=1}^m \lambda_i \nabla p_i(u), \quad \lambda_i p_i(u) = 0, \ \forall\, i = 1, \cdots, m.$$

By Lemma 23 and Theorem 24, we get

$$\tilde{f}_\ell(x) = \ell^T x - \ell^* - \sum_{i=1}^m \lambda_i p_i(x) = \sum_{\nu \in \{0,1\}^m} \sigma_\nu(x) g_1^{\nu_1}(x) \cdots g_m^{\nu_m}(x)$$

for some SOS polynomials σ_ν with $\deg(\sigma_\nu g_1^{\nu_1} \cdots g_m^{\nu_m}) \leq 2N$. So we have the identity

$$\ell^T x - \ell^* = \sum_{i=1}^m \lambda_i p_i(x) + \sum_{\nu \in \{0,1\}^m} \sigma_\nu(x) g_1^{\nu_1}(x) \cdots g_m^{\nu_m}(x).$$

We can write $\sigma_\nu(x) = [x^{d-d_\nu}]^T W_\nu [x^{d-d_\nu}]$ for some symmetric matrix $W_\nu \succeq 0$. In the above identity, replacing each monomial x^α by $\hat{y}_\alpha$, we get

$$\ell^T \hat{x} - \ell^* = \sum_{i=1}^m \lambda_i L_{p_i}(\hat{y}) + \sum_{\nu \in \{0,1\}^m} Trace\Big(W_\nu \cdot \big(\sum_{0 \leq |\alpha| \leq 2N} A_\alpha^\nu \hat{y}_\alpha\big)\Big) \geq 0,$$

which contradicts (5.5). $\square$

Remark: In LMI (5.4), we use all the products $g^\nu(x) = g_1^{\nu_1}(x) \cdots g_m^{\nu_m}(x)$ for all $\nu \in \{0,1\}^m$ which results an exponential size of LMI. As we did in the end of Section 2, the set S is also the projection of the following set

$$
\left.
\begin{array}{c}
\forall\, 0 \leq k \leq m, \quad A_0^{(k)} + \displaystyle\sum_{0 < |\alpha| \leq 2N} A_\alpha^{(k)} y_\alpha \succeq 0 \\
L_{p_1}(y) \geq 0, \cdots, L_{p_m}(y) \geq 0 \\
x_1 = y_{e_1}, \cdots, x_n = y_{e_n}, y_0 = 1
\end{array}
\right\}
\tag{5.6}
$$

where symmetric matrices $A_\alpha^{(k)}$ are defined in LMI (2.4). If the archimedean condition holds, we can similarly prove (5.6) is a lifted LMI for S when N is sufficiently large, as we did in the above proof.

5.4. Proof of Theorems 3 and 4

In the remarks after Theorem 4, we mentioned that Theorems 3 can be implied by Theorem 4. So we only need to prove Theorem 4.

Lemma 25 *Let S be as in Theorem 4. Then there exists an open set containing S and polynomials $p_1(x), \cdots, p_m(x)$ which either are sos-concave or have negative definite Hessian on U such that*

$$
S = \{x \in U : p_1(x) \geq 0, \cdots, p_m(x) \geq 0\}.
$$

Proof If $S_i = \{x \in \mathbb{R}^n : g_i(x)\}$ is sos-convex, choose $p_i(x) = g_i(x)$. If S_i is extendable poscurv-convex with respect to S, by Proposition 17, there exists a polynomial $h_i(x)$ positive on S such that the product $g_i(x)h_i(x)$ has negative definite Hessian on S, then choose $p_i(x) = g_i(x)h_i(x)$. Since S is compact, we can choose an open set $U \supset S$ small enough to make the lemma true. $\square$ *Proof of*

Theorem 4 The proof is almost the same as the one for Theorem 2. We follow the approach there, and only list the distinctive parts here. Let U and polynomials $p_i(x)$ be given by Lemma 25. Then define LMI (5.4) using both p_i and g_i, and S is contained in the projection of this LMI. Then we claim (5.4) is a lifted LMI for S for N given by Theorem 24.

Similarly, we prove this by contradiction. Suppose there exists a vector $(\hat{x}, \hat{y})$ in $\hat{S}_N$ such that $\hat{x} \notin S$. By the Hahn-Banach Separation Theorem, there must exist a vector ℓ of unit length and ℓ^* such that

$$
\ell^T \hat{x} - \ell^* < 0, \qquad \ell^T x - \ell^* \geq 0, \, \forall\, x \in S.
$$

Let u be a minimizer of $\ell^T x$ over S. Then Slater's condition implies the first order optimality condition holds at u for the set of defining polynomials $\{x \in U : p_1(x) \geq 0, \cdots, p_m(x) \geq 0\}$. Let $\lambda_i \geq 0$ be the corresponding Lagrange multipliers and

$$
\tilde{f}_\ell(x) := \ell^T x - \ell^* - \sum_{i=1}^m \lambda_i p_i(x).
$$

Similarly, by Theorem 24, we can get the identity

$$
\ell^T x - \ell^* = \sum_{i=1}^m \lambda_i p_i(x) + \sum_{\nu \in \{0,1\}^m} \sigma_\nu(x) g_1^{\nu_1}(x) \cdots g_m^{\nu_m}(x)
$$

for some SOS polynomials σ_ν with degree bounds $\deg(\sigma_\nu g_1^{\nu_1} \cdots g_m^{\nu_m}) \leq 2N$. We can also write $\sigma_\nu(x) = [x^{d-d_\nu}]^T W_\nu [x^{d-d_\nu}]$ for some symmetric matrix $W_\nu \succeq 0$. Similar to what we have done in the proof for Theorem 2, a contradiction to $\ell^T \hat{x} - \ell^* \geq 0$ can be obtained by replacing each monomial x^α by $\hat{y}_\alpha$ in the above identity. $\square$

6. Appendix: The complexity of the matrix Positivstellensatz

Throughout this section, we only need assume $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \cdots, g_m(x) \geq 0\}$ is compact. We do not need either $g_i(x)$ is concave or S is convex. Without loss of generality, assume $S \subset (-1, 1)^n$, otherwise do some coordinate transformation.

Suppose we have a symmetric matrix polynomial $F(x) \in \mathbb{R}[x]^{r \times r}$ which is positive definite on S. Our goal is to give a Positivstellensatz representation of $F(x)$ in terms of defining polynomials $g_1(x), \cdots, g_m(x)$ with bounds on the degrees of the representing polynomials. When $r = 1$, that is, $F(x)$ are scalar polynomials, Schmüdgen's Positivstellensatz [27] says that $F(x)$ has the representation

$$F(x) = \sum_{\nu \in \{0,1\}^m} \sigma_\nu(x) g_1^{\nu_1}(x) \cdots g_m^{\nu_m}(x)$$

for some SOS polynomials σ_ν. Furthermore, if the archimedean condition holds, Putinar's Positivstellensatz [21] says that $F(x)$ has the representation

$$F(x) = \sigma_0(x) + \sigma_1(x) g_1(x) + \cdots + \sigma_m(x) g_m(x)$$

for some SOS polynomials σ_i.

These representation results can be generalized to the case $r > 1$. Schmüdgen's matrix Positivstellensatz says that there exist symmetric SOS matrix polynomials $G_\nu(x) \in \mathbb{R}[x]^{r \times r}$ such that

$$F(x) = \sum_{\nu \in \{0,1\}^m} g_1^{\nu_1}(x) \cdots g_m^{\nu_m}(x) G_\nu(x).$$

Under the archimedean condition, Putinar's matrix Positivstellensatz says that there exist symmetric SOS matrix polynomials $G_i(x) \in \mathbb{R}[x]^{r \times r}$ such that

$$F(x) = G_0(x) + g_1(x) G_1(x) + \cdots + g_m(x) G_m(x).$$

We refer to [23] for these representations of positive definite matrix polynomials. The goal of this section is to give degree bounds for $G_\nu(x)$ in these representations.

6.1. Schmüdgen's matrix Positivstellensatz

For a scalar polynomial $f(x) = \sum_\alpha f_\alpha x^\alpha$, its norm $\|f\|$ is defined to be

$$\|f\| = \max_\alpha \left\{ |f_\alpha| \frac{\alpha_1! \cdots \alpha_n!}{|\alpha|!} \right\}. \tag{6.1}$$

For a matrix polynomial $F(x) = \sum_\alpha F_\alpha x^\alpha$, its norm is defined to be

$$\|F\| = \max_\alpha \left\{ \|F_\alpha\|_2 \frac{\alpha_1! \cdots \alpha_n!}{|\alpha|!} \right\} = \max_\alpha \left\{ \frac{\|\partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n} F(0)\|_2}{|\alpha|!} \right\}. \tag{6.2}$$

Here $\|A\|_2$ denotes the maximum singular value of matrix A.

Lemma 26 *Suppose polynomials $g_i(x)$ are scaled such that for some $\varepsilon > 0$*

$$S = \{x \in (-1 + \varepsilon, 1 - \varepsilon)^n : \quad g_1(x) \geq 0, \cdots, g_m(x) \geq 0, \ \sum_{i=1}^m g_i(x) < 2n\varepsilon\}.$$

Define new polynomials

$$p_1 = 1 - \varepsilon + x_1, \cdots, p_n = 1 - \varepsilon + x_n$$
$$p_{n+1} = 1 - \varepsilon - x_1, \cdots, p_{2n} = 1 - \varepsilon - x_n$$
$$p_{2n+1} = g_1, \cdots, p_{2n+m} = g_m, \quad p_{2n+m+1} = 2n\varepsilon - (g_1 + \cdots + g_m).$$

Then there exists an integer $c > 0$ depending only on the polynomials $g_1, \cdots, g_m$ such that for every symmetric matrix polynomial $F(x) \succ 0$ on S can be written

$$F = \sum_{|\alpha| \leq N} p_1^{\alpha_1} \cdots p_{2n+m+1}^{\alpha_{2n+m+1}} F_\alpha \tag{6.3}$$

with constant symmetric matrices $F_\alpha \succ 0$ and $N \leq \Theta\left(\frac{\|F\|}{F^*}\right)$. Here

$$d = \deg(F(x)), \quad F^* := \min_{x \in S} \lambda_{\min}(F(x)), \quad \Theta(s) := cd^2 \left(1 + \left(d^2 n^d s\right)^c\right).$$

Proof The proof is very similar to the one of Lemma 9 by Schweighofer [24] and uses a famous theorem of Polyá. The only difference is the scalar polynomials $f(x)$ in [24] are replaced by matrix polynomials $F(x)$. We follow the approach for proving Lemma 9 in [24]. Without loss of generality, assume $\|F\| = 1$. Introduce new variables $y = (y_1, \cdots, y_{2n+m+1})$. Define the homomorphism

$$\varphi : \mathbb{R}[y] \to \mathbb{R}[x] : \ y_i \mapsto p_i.$$

Then $\varphi(y_1 + \cdots + y_{2n+m+1}) = 2n$ and hence $y_1 + \cdots + y_{2n+m+1} - 2n \in \ker(\varphi)$. By Hilbert's basis theorem, there exist polynomials $r_1, \cdots, r_t$ so that

$$\ker(\varphi) = \langle y_1 + \cdots + y_{2n+m+1} - 2n, r_1, \cdots, r_t \rangle.$$

Then define new sets

$$\Delta := \{y \in \mathbb{R}_+^{2n+m+1} : \ y_1 + \cdots + y_{2n+m+1} = 2n\},$$
$$Z := \{y \in \Delta : \ r_1(y) = \cdots = r_t(y) = 0\}.$$

The properties listed below hold, which are essentially Claim 1 and Claim 2 of [24].

(P1) The linear map

$$\ell : \mathbb{R}^{2n+m+1} \to \mathbb{R}^n : \ (y_1, \cdots, y_{2n+m+1}) \mapsto \frac{1}{2}(y_1 - y_{n+1}, \cdots, y_n - y_{2n})$$

induces a bijection $\ell|_Z : Z \to S$.

(P2) There exists an integer $d_0 \geq 1$ and a d_0-form $R_0 \in \ker \varphi$ such that $R_0 \geq 0$ on Δ and $Z = \{y \in \Delta : \ R_0(y) = 0\}$.

By Lojasiewicz's inequality (Corollary 2.6.7 in [2]), there exist integers $c_0, c_1 \geq 1$ such that

$$(\mathbf{E1}) \qquad\qquad \mathrm{dist}(y, Z)^{c_0} \leq c_1 R_0(y), \quad \forall \, y \in \Delta.$$

Define new constants

$$c_2 := 2^{c_0+1} c_1 \sqrt{2n}, \qquad c_3 := c_2 (2n)^{d_0} \|R_0\|, \qquad c_4 := (2n)^{d_0}.$$

Then choose $c > 0$ big enough so that

$$d_0^2 (1 + c_4 a + c_3 a^{c_0+1}) \leq c(1 + a^c), \quad \forall \, a \in [0, \infty).$$

Now we write $F(x) = F_0(x) + \cdots + F_d(x)$ with F_k being matrix k-forms (homogeneous matrix polynomials of degree k). Set $d_1 := \max\{d, d_0\}$ and

$$P(y) := \sum_{k=0}^{d} P_k(y) \left(\frac{y_1 + \cdots + y_{2n+m+1}}{2n}\right)^{d_1-k}, \qquad P_k(y) := F_k \left(\frac{1}{2}(y_1 - n_{n+1}), \cdots, \frac{1}{2}(y_n - y_{2n})\right).$$

Then P is a d_1-form such that

$$\varphi(P) = F, \qquad\qquad P(y) = F(\ell(y)) \ \forall y \in \Delta.$$

So we can see (P1) implies

(**E2**) $\min\left\{\lambda_{\min}\Big(P(y)\Big) : y \in Z\right\} = \min\left\{\lambda_{\min}(F(x)) : x \in S\right\} = F^* > 0.$

Next, define the d_1-form R as follows

$$R(y) := R_0(y) \cdot \left(\frac{y_1 + \cdots + y_{2n+m+1}}{2n}\right)^{d_1 - d_0}.$$

By equations (17) and (18) of [24], we know $\|R\| \leq \frac{1}{(2n)^{d_1-d_0}}\|R_0\|$ and

(**E3**) $R(y) = R_0(y), \quad \forall\, y \in \Delta.$

Now we claim the property listed below holds

(P3) For all $y, y' \in \Delta$, it holds

$$\left|\lambda_{\min}\Big(P(y)\Big) - \lambda_{\min}\Big(P(y')\Big)\right| \leq \|P(y) - P(y')\|_2 \leq \sqrt{n}d^2n^{d-1}\|y - y'\|.$$

The first inequality of (P3) can be obtained by noting the fact that for any two symmetric matrices A, B it holds

$$\lambda_{\min}(A) = \min_{\|\xi\|=1} \xi^T A\xi = \min_{\|\xi\|=1}\left(\xi^T B\xi + \xi^T(A - B)\xi\right) \leq \lambda_{\min}(B) + \|A - B\|_2,$$
$$\lambda_{\min}(B) = \min_{\|\xi\|=1} \xi^T B\xi = \min_{\|\xi\|=1}\left(\xi^T A\xi + \xi^T(B - A)\xi\right) \leq \lambda_{\min}(A) + \|A - B\|_2.$$

The second inequality of (P3) is a consequence of Claim 3 in [24], because

$$\|P(y) - P(y')\|_2 = \sup_{\|\xi\|=1} |\xi^T P(y)\xi - \xi^T P(y')\xi|$$

and Claim 3 in [24] can be applied to the scalar polynomials $\xi^T P(y)\xi$.

For those $y, y' \in \Delta$ satisfying $\lambda_{\min}\Big(P(y)\Big) \leq \frac{1}{2}F^*$ and $\lambda_{\min}\Big(P(y')\Big) \geq F^*$, by (P3), we have

$$\|y - y'\| \geq \frac{F^*}{2\sqrt{n}d^2n^{d-1}} \geq \frac{F^*}{2d^2n^d}.$$

Therefore, Property (E2) implies that, for all $y \in \Delta$ with $\lambda_{\min}\Big(P(y)\Big) \leq \frac{1}{2}F^*$, it holds

$$\text{dist}(y, Z) \geq \frac{F^*}{2d^2n^d},$$

and hence Properties (E2) and (E3) imply, for all $y \in \Delta$ with $\lambda_{\min}\Big(P(y)\Big) \leq \frac{1}{2}F^*$, it holds

$$\left(\frac{F^*}{2d^2n^d}\right)^{c_0} \leq c_1 R(y).$$

In (P3), if we choose y' to be a minimizer of $\lambda_{\min}\Big(P(y)\Big)$ on Z, then for all $y \in \Delta$ we have

$$\left|\lambda_{\min}\Big(P(y)\Big) - F^*\right| \leq \text{diameter}(\Delta)\sqrt{n}d^2n^{d-1} \leq 2\sqrt{2n}d^2n^d$$

which obviously implies

(**E4**) $\lambda_{\min}\Big(P(y)\Big) \geq F^* - 2\sqrt{2n}d^2n^d, \qquad \forall\, y \in \Delta.$

Let $\lambda := c_2 d^2 n^d \left(\frac{d^2 n^d}{F^*} \right)^{c_0}$ and define a new set

$$\Delta_1 := \{ y \in \mathbb{R}_+^{2n+m+1} : y_1 + \cdots + y_{2n+m+1} = 1 \}.$$

Now we claim that

(**E5**) $$\lambda_{\min}\Big(P(y) \Big) + \lambda R(y) \geq \frac{1}{2} F^*, \forall\, y \in \Delta$$

where r is the dimension of the matrix polynomial $F(x)$. Now we prove (E5). Obviously, (E5) holds for those $y \in \Delta$ with $\lambda_{\min}(P(y)) \geq \frac{1}{2} F^*$. We only need to verify (E5) for those $y \in \Delta$ with $\lambda_{\min}(P(y)) \leq \frac{1}{2} F^*$. The choice of λ shows

$$\lambda R(y) \geq \frac{c_2}{c_1 2^{c_0}} d^2 n^d$$

and hence (E4) implies

$$\lambda_{\min}\Big(P(y) \Big) + \lambda R(y) \geq F^* - 2\sqrt{2n} d^2 n^d + \frac{c_2}{c_1 2^{c_0}} d^2 n^d = F^* \geq \frac{1}{2} F^*.$$

Therefore we obtain that (by concavity of function $\lambda_{\min}(\,\cdot\,)$)

$$\lambda_{\min}\Big(P(y) + \lambda R(y) I_r \Big) \geq \lambda_{\min}\Big(P(y) \Big) + \lambda R(y) \geq \frac{1}{2} F^*, \quad \forall\, y \in \Delta$$

which by homogeneity implies

$$\lambda_{\min}\Big(P(y) + \lambda R(y) I_r \Big) \geq \frac{F^*}{2(2n)^{d_1}}, \quad \forall\, y \in \Delta_1.$$

Then Theorem 3 from Scherer and Hol [23] guarantees that the product

$$Q(y) := \Big(P(y) + \lambda R(y) \cdot I_r \Big) \cdot \left(\frac{y_1 + \cdots + y_{2n+m+1}}{2n} \right)^N$$

has positive definite matrix coefficients for all

$$N > \frac{d_1(d_1-1)\|P + \lambda R \cdot I_r\|}{2 \frac{F^*}{2(2n)^{d_1}}} - d_1 = d_1(d_1-1)(2n)^{d_1} \frac{\|P + \lambda R \cdot I_r\|}{F^*} - d_1.$$

If N is chosen to be the smallest integer in the above, then

$$\deg(Q) \leq c d^2 \left(1 + \left(\frac{d^2 n^d}{F^*} \right)^c \right),$$

as is shown at the end of the proof of Lemma 9 in [24]. Since $F(x) = \varphi(P(x)) = \varphi(Q(x))$, we have proved $F(x)$ can be represented like (6.3) with $F_\alpha \succ 0$. $\square$

Now we arrive at our theorem giving degree bounds.

Theorem 27. *If matrix polynomial $F(x) \succeq \delta I \succ 0$ for all x in a compact subset S of $\mathbb{R}^n$, then*

$$F(x) = \sum_{\nu \in \{0,1\}^m} g_1^{\nu_1} \cdots g_m^{\nu_m} G_\nu(x)$$

where $G_\nu(x)$ are SOS matrix polynomials such that

$$\deg(g_1^{\nu_1} \cdots g_m^{\nu_m} G_\nu) \leq \Omega \left(\frac{\|F\|}{\delta} \right) := \kappa \cdot \Theta \left(\frac{\|F\|}{\delta} \right)$$

where κ is a constant depending only on the polynomials $g_i(x)$.

Proof Again take $S \subset (-1+\varepsilon, 1-\varepsilon)^n$. Then the polynomials $p_1, \cdots, p_{2n}, p_{2n+m+1}$ in the preceding lemma are positive on S. By Schmüdgen's Positivstellensatz, for every $i \in \{1, \cdots, 2n, 2n+m+1\}$ we have

$$p_i(x) = \sum_{\nu \in \{0,1\}^m} \sigma_\nu^{(i)} g_1^{\nu_1} \cdots g_m^{\nu_m} \tag{6.4}$$

with $\sigma_\nu^{(i)}(x)$ being SOS polynomials. Let

$$\kappa = \max \left\{ \deg(\sigma_\nu^{(i)} g_1^{\nu_1} \cdots g_m^{\nu_m}) : i = 1, \cdots, 2n, 2n+m+1, \; \nu \in \{0,1\}^m \right\}.$$

Now into identity (6.3) we plug the representation (6.4) for $p_1(x), \cdots, p_{2n}(x), p_{2n+m+1}(x)$, then we obtain the conclusion in the theorem. $\square$

6.2. Putinar's matrix Positivstellensatz

Lemma 28 *Suppose $g_i(x)$ are scaled such that $g_i(x) \leq 1$ on $[-1,1]^n$. Then there exist constants $c_0, c_1, c_2 > 0$ with the property:*

For all symmetric matrix polynomials $F(x) \in \mathbb{R}[x]^{r \times r}$ of degree d such that $F(x) \succeq \delta I_r$ for all $x \in S$, if we set

$$L := d^2 n^{d-1} \frac{\|F\|}{\delta}, \qquad \lambda := c_1 d^2 n^{d-1} \|F\| L^{c_2}$$

and if $k \in \mathbb{N}$ satisfies

$$2k+1 \geq c_0(1 + L^{c_0}),$$

then the inequality

$$F(x) - \lambda \sum_{i=1}^m (g_i(x) - 1)^{2k} g_i(x) I_r \succeq \frac{\delta}{2} I_r$$

holds on $[-1,1]^n$ and hence on the unit ball $B(0,1) = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$.

Proof Apply Lemma 13 in [17] to the polynomial $f^{(\xi)}(x) := \xi^T F(x) \xi$ where $\xi \in \mathbb{R}^r$ is a unit length vector. Note that the minimum value of $f^{(\xi)}(x)$ is at least δ and $\|f^{(\xi)}(x)\|$ is at most $\|F\|$. For L, λ, k given in the lemma, we have for all $x \in [-1,1]^n$

$$\xi^T F(x) \xi - \lambda \sum_{i=1}^m (g_i(x) - 1)^{2k} g_i(x) \geq \frac{\delta}{2}, \quad \forall \xi \in \mathbb{R}^r, \|\xi\| = 1$$

which implies

$$F(x) - \lambda \sum_{i=1}^m (g_i(x) - 1)^{2k} g_i(x) I_r \succeq \frac{\delta}{2} I_r$$

for all $x \in [-1,1]^n$. $\square$

Theorem 29. *Assume the archimedean condition holds for the g_i. If $F(x)$ is a symmetric matrix polynomial of degree d such that $F(x) \succeq \delta I \succ 0$ for all $x \in S \subset \mathbb{R}^n$, then*

$$F(x) = G_0(x) + g_1(x) G_1(x) + \cdots + g_m(x) G_m(x) \tag{6.5}$$

where $G_i(x)$ are SOS matrix polynomials such that

$$\deg(G_0), \; \deg(g_1 G_1), \cdots, \deg(g_m G_m) \; \leq \; c \left(d^2 n^d \frac{\|F\|}{\delta} \right)^c$$

for some constant c depending only on the polynomials $g_i(x)$.

Proof The proof is almost the same as for Theorem 6 in [17]. By the archimedean condition, we can assume

$$R - \sum_{i=1}^{n} x_i^2 = s_0(x) + s_1(x)g_1(x) + \cdots + s_m(x)g_m(x) \tag{6.6}$$

for SOS polynomials $s_i(x)$. Without loss of generality, we can assume $R = 1$, because otherwise we can apply some coordinate transformation. Let

$$d_1 = \max_i(\deg(s_i g_i)), \qquad d_2 = 1 + \max_i(\deg(g_i)).$$

First, apply Lemma 28 to find constants L, λ, k such that

$$\tilde{F}(x) := F(x) - \lambda \sum_{i=1}^{m} (g_i(x) - 1)^{2k} g_i(x) I_r \succeq \frac{\delta}{2} I_r, \quad x \in [-1, 1]^n.$$

By (33) in the proof of Theorem 6 in [17], we have for every $\|\xi\| = 1$

$$\|\xi^T \tilde{F}(x)\xi\| \leq \|\xi^T F(x)\xi\| + \lambda d_2^{2k+1}$$

which implies

$$\|\tilde{F}\| \leq \|F\| + \lambda d_2^{2k+1}.$$

By (34) in the proof of Theorem 6 in [17], we get

$$\deg(\tilde{F}(x)) \leq \max\{d, (2k+1)d_2, 1\} := d_h.$$

Then we apply Theorem 27 to $\tilde{F}(x)$ on the unit ball $B(0,1)$. So there exists some constant $c_3 > 0$ such that

$$\tilde{F}(x) = H_0(x) + (1 - \sum_{i=1}^{n} x_i^2)H_1(x) \tag{6.7}$$

for some SOS matrix polynomials $H_i(x)$ with degree

$$\deg(H_0) \leq k_h, \quad 2 + \deg(H_1) \leq k_h \qquad where \quad k_h := c_3 d_h^2 \left(1 + d_h^2 n^{d_h} \frac{2\|\tilde{F}\|}{\delta}\right)^{c_3}.$$

By (36) in the proof of Theorem 6 in [17], we know for some constant $c_4 > 0$

$$k_h \leq c_4 \exp(c_4 d_h) \left(\frac{2\|\tilde{F}\|}{\delta}\right)^{c_3}.$$

By (37) in the proof of Theorem 6 in [17], we have for some constant $c_5 > 0$

$$\frac{2\|\tilde{F}\|}{\delta} \leq \exp(L^{c_5}).$$

Combine (6.6) and (6.7) to get

$$\tilde{F}(x) = \underbrace{H_0(x) + s_0(x)H_1(x)}_{G_0(x)} + g_1(x)\underbrace{s_1(x)H_1(x)}_{G_1(x)} + \cdots + g_m(x)\underbrace{s_m(x)H_1(x)}_{G_m(x)}.$$

Now we can estimate $\deg(g_i G_i)$ by following the proof for Theorem 6 in [17]. The techniques are exactly same. Finally we can obtain the degree bound in (6.5) for some constant $c > 0$. $\square$

7. Conclusions

This paper studies the SDP representation of convex sets. Obviously, for a set S to be SDP representable, necessary conditions are that S must be convex and semialgebraic. It is not known if these conditions are also sufficient, but the main contribution of this paper is to give some additional conditions which are sufficient. Given $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \cdots, g_m(x) \geq 0\}$ which is convex, compact and has nonempty interior, we have proved S is SDP representable in either of the following cases: (i) All $g_i(x)$ are concave on S, and the PDLH condition holds; (ii) Each S_i is either sos-convex or extendable poscurv-convex with respect to S.

The key to our proofs is to find and prove a well-behaved Schmüdgen or Putinar's representation for the affine polynomial $\ell^T x - \ell^*$ nonnegative on S, that is, to find conditions for the S-BDR property or the PP-BDR property to hold for affine polynomials. When $\ell^T x - \ell^*$ is nonnegative on S and equals zero at u in S, we can not directly apply Schmüdgen or Putinar's Positivstellensatz to prove the existence of the representation. However, we should mention that it is possible to prove the existence of Schmüdgen or Putinar's representations for $\ell^T x - \ell^*$ by applying the representation of nonnegative polynomials in Marshall [14] and Scheiderer [25, 26]. But the degrees of these representations depend on the choice of ℓ and we can not get a uniform degree bound from these papers. So we were motivated to use Hessians of defining polynomials to get the degree bound independent of ℓ. One interesting future work is to get the SDP representability of S by using methods in Marshall [14] and Scheiderer [25, 26].

The main result of this paper is that if the boundary of every S_i is either sos-convex or extendable poscurv-convex with respect to S, then the compact convex set S is semidefinite representable. We point out that the condition of extendable poscurv-convexity does not require much more than that the boundary $\partial S_i \cap \partial S$ is nonsingular and has positive curvature. In the follow-up paper [7] to this one, the authors have proved a stronger result: if for every i either S_i is sos-convex or $\partial S_i \cap \partial S$ is positively curved and nonsingular, then the compact convex set S is SDP representable; this is based on constructions of the lifted LMIs and theorems we have proved in this paper. Since a necessary condition for a set S to be convex is that its boundary ∂S can have only nonnegative curvature (under some nonsingularity assumption on ∂S), we can see that the sufficient conditions of semidefinite representability given in this paper are not far away from the necessary conditions.

References

1. A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications.* MPS-SIAM Series on Optimization, SIAM, Philadelphia, 2001.
2. J. Bochnak, M. Coste and M-F. Roy. *Real Algebraic Geometry*, Springer, 1998.
3. S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004.
4. J.B. Conway. *A course in Functional Analysis*, Springer Graduate Texts in Math Series, 1985.
5. V. G. Dmitriev. Submanifolds with everywhere semidefinite second fundamental form. *Journal of Mathematical Sciences*, Vol. 10, NO. 3, pp. 448-450, September, 1978.
6. M. Hazewinkel, *Encyclopedia of Mathematics*, Springer, http://eom.springer.de/f/f042200.htm
7. J.W. Helton and J. Nie. Sufficient and Necessary Conditions for Semidefinite Representability of Convex Hulls and Sets. Submitted to *SIAM Journal on Optimization*, 2007. http://arxiv.org/abs/0709.4017.
8. J.W. Helton and J. Nie. Structured Semidefinite Representation of Some Convex Sets. To appear in *Proceedings of 47th IEEE Conference on Decision and Control (CDC)*, 2008. http://arXiv.org/abs/0802.1766.
9. W. Helton and V. Vinnikov. *Linear matrix inequality representation of sets. Comm. Pure Appl. Math.* **60** (2007), No. 5, pp. 654-674.
10. D. Henrion and J. Lasserre. GloptiPoly: Global optimization over polynomials with Matlab and SeDuMi. *ACM Trans. Math. Soft.*, 29:165-194, 2003.
11. M.W. Hirsch. *Differential Topology*, Graduate Texts in Mathematics **33**, Springer-Verlag, 1997.

12. J. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.*, **11** (2001), No. 3, 796–817.
13. J. Lasserre. Convex sets with lifted semidefinite representation. Preprint, LAAS, 2006.
14. M. Marshall. Representation of non-negative polynomials, degree bounds and applications to optimization. To appear in *Canad. J. Math.*. `http://math.usask.ca/~marshall/`.
15. Y. Nesterov and A. Nemirovskii. Interior-point polynomial algorithms in convex programming. SIAM Studies in Applied Mathematics, 13. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994.
16. A. Nemirovskii. Advances in convex optimization: conic programming. *Plenary Lecture, International Congress of Mathematicians (ICM)*, Madrid, Spain, 2006.
17. J. Nie and M. Schweighofer. On the complexity of putinar's positivstellensatz. *Journal of Complexity*, 23(2007) 135-150.
18. P.Parrilo. Exact semidefinite representation for genus zero curves. Talk at the Banff workshop "Positive Polynomials and Optimization", Banff, Canada, October 8-12, 2006.
19. P. Parrilo and B. Sturmfels. Minimizing polynomial functions, Proceedings of the DIMACS Workshop on Algorithmic and Quantitative Aspects of Real Algebraic Geometry in Mathematics and Computer Science (March 2001), (eds. S. Basu and L. Gonzalez-Vega), American Mathematical Society, 2003, pp. 83–100.
20. S. Prajna, A. Papachristodoulou, P. Seiler and P. Parrilo. *SOSTOOLS User's Guide*. Website: `http://www.mit.edu/~parrilo/SOSTOOLS/`.
21. M. Putinar. Positive polynomials on compact semi-algebraic sets, *Ind. Univ. Math. J.* **42** (1993) 203–206.
22. B. Reznick. Some concrete aspects of Hilbert's 17th problem. In *Contemp. Math.*, volume 253, pages 251-272. American Mathematical Society, 2000.
23. C. Scherer and C. Hol. Matrix Sum-of-Squares Relaxations for Robust Semi-Definite Programs. *Math. Prog.* 107 (2006), No. 1-2, Ser. B, 189–211.
24. M. Schweighofer. On the complexity of Schmüdgen's Positivstellensatz. *Journal of Complexity* 20, No. 4, 529-543 (2004).
25. C. Scheiderer. Sums of squares on real algebraic curves. *Math. Z.* 245, 725-760 (2003).
26. C. Scheiderer. Distinguished representations of non-negative polynomials. *J. Algebra* 289, 558-573 (2005).
27. K. Schmüdgen. *The K-moment problem for compact semialgebraic sets*, Math. Ann. **289** (1991), 203–206.
28. H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors. *Handbook of semidefinite programming.* Kluwer's Publisher, 2000.