ON THE EXACTNESS OF LASSERRE RELAXATIONS FOR COMPACT CONVEX BASIC CLOSED SEMIALGEBRAIC SETS

TOM-LUKAS KRIEL AND MARKUS SCHWEIGHOFER

ABSTRACT. Consider a finite system of non-strict real polynomial inequalities and suppose its solution set $S \subseteq \mathbb{R}^n$ is convex, has nonempty interior and is compact. Suppose that the system satisfies the Archimedean condition, which is slightly stronger than the compactness of S. Suppose that each defining polynomial satisfies a second order strict quasiconcavity condition where it vanishes on S (which is very natural because of the convexity of S) or its Hessian has a certain matrix sums of squares certificate for negative-semidefiniteness on S (fulfilled trivially by linear polynomials). Then we show that the system possesses an exact Lasserre relaxation.

In their seminal work of 2009, Helton and Nie showed under the same conditions that S is the projection of a spectrahedron, i.e., it has a semidefinite representation. The semidefinite representation used by Helton and Nie arises from glueing together Lasserre relaxations of many small pieces obtained in a non-constructive way. By refining and varying their approach, we show that we can simply take a Lasserre relaxation of the original system itself. Such a result was provided by Helton and Nie with much more machinery only under very technical conditions and after changing the description of S.

1. INTRODUCTION

Throughout the article, $\mathbb{N}$ and $\mathbb{N}_0$ denote the set of positive and nonnegative integers, respectively. We fix $n \in \mathbb{N}_0$ and denote by $\underline{X} := (X_1, \dots, X_n)$ a tuple of n variables. We denote by $\mathbb{R}[\underline{X}] := \mathbb{R}[X_1, \dots, X_n]$ the polynomial ring in these variables over $\mathbb{R}$. For $\alpha \in \mathbb{N}_0^n$, we denote $|\alpha| := \alpha_1 + \ldots + \alpha_n$ and $\underline{X}^\alpha := X_1^{\alpha_1} \cdots X_n^{\alpha_n}$. For $p = \sum_\alpha a_\alpha \underline{X}^\alpha \in \mathbb{R}[\underline{X}]$ with all $a_\alpha \in \mathbb{R}$, the degree of p is defined as $\deg p := \max\{|\alpha| \mid a_\alpha \neq 0\}$ if $p \neq 0$ and $\deg p := -\infty$ if $p = 0$. For each $d \in \mathbb{R}$, we consider the real vector space

$$\mathbb{R}[\underline{X}]_d := \{p \in \mathbb{R}[\underline{X}] \mid \deg p \leq d\}$$

of all polynomials of degree at most d. We admit here real numbers d for technical reasons but note that $\mathbb{R}[\underline{X}]_d = \mathbb{R}[\underline{X}]_{\lfloor d \rfloor}$ for all $d \in \mathbb{R}$ and $\mathbb{R}[\underline{X}]_d = \{0\}$ for all $d < 0$. Occasionally, we will need the real polynomial ring in one variable as an auxiliary tool, and we will denote it by $\mathbb{R}[T]$. We will denote the $n \times n$ identity matrix by I_n.

Date: January 29, 2018.

2010 *Mathematics Subject Classification.* Primary 14P10, 52A20; Secondary 13J30, 52A41, 90C22, 90C26.

Key words and phrases. moment relaxation, Lasserre relaxation, basic closed semialgebraic set, sum of squares, polynomial optimization, semidefinite programming, linear matrix inequality, spectrahedron, semidefinitely representable set.

For a tuple $\underline{g} := (g_1, \dots, g_m) \in \mathbb{R}[\underline{X}]^m$ of m polynomials, the set

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

is called a *basic closed semialgebraic set* [PD, Def. 2.1.1]. Boolean combinations of such sets are called *semialgebraic sets* [PD, Def. 2.1.4]. The finiteness theorem from real algebraic geometry says that every closed semialgebraic set is a finite union of basic closed ones [PD, Thm. 2.4.1]. In general, it is hard to answer questions about the geometry $S(\underline{g})$ from its description $\underline{g}$. This is of course due to the nonlinear monomials $\underline{X}^\alpha$ with $|\alpha| \geq 2$ that might appear in $\underline{g}$. An extremely naive idea would be to replace each such nonlinear monomial $\underline{X}^\alpha$ in $\underline{g}$ by a new variable Y_α. This would lead to a system of m *linear* inequalities whose solution set is a (closed convex) polyhedron in a higher-dimensional space. The projection of this polyhedron to the $\underline{X}$-space $\mathbb{R}^n$ contains $S(\underline{g})$ but will very often just be the whole of $\mathbb{R}^n$ and thus be of no help.

This idea becomes however less naive if we add a bunch of redundant inequalities before the linearization. For example, we could add certain inequalities of the form $p^2(x) \geq 0$ or $(p^2 g_i)(x) \geq 0$ with $p \in \mathbb{R}[\underline{X}]$. If we choose finitely many such inequalities in a clever way and then linearize as above, we will get a polyhedron in a higher-dimensional space whose projection to $\underline{X}$-space $\mathbb{R}^n$ might enclose $S(\underline{g})$ more tightly. Unless $S(\underline{g})$ happens to be a polyhedron, this projection can however still not equal $S(\underline{g})$ since projections of polyhedra are again polyhedra (see [Scr, Subsection 12.2] for a textbook reference).

The idea of Lasserre was therefore to add the whole (infinite) family of all redundant inequalities of the form $p^2(x) \geq 0$ or $(p^2 g_i)(x) \geq 0$ with $p \in \mathbb{R}[\underline{X}]$ before the linearization [L1, L2]. To get something that is useful in practice (for example, one would like to avoid using infinitely many of the new variables Y_α), he restricted the degree of the polynomials of the added redundant inequalities.

Therefore fix a degree bound $d \in \mathbb{N}_0$ and set $g_0 := 1 \in \mathbb{R} \subseteq \mathbb{R}[\underline{X}]$. For each $i \in \{0, \dots, m\}$ with $g_i \neq 0$, fix a (column) vector v_i whose entries are the different monomials of degree at most

$$r_i := \frac{d - \deg g_i}{2} \tag{1}$$

and set $\ell_i := \dim \mathbb{R}[\underline{X}]_{r_i}$. Note that in the case $g_i \notin \mathbb{R}[\underline{X}]_d$, r_i is negative, and consequently $\ell_i = 0$ and $v_i = () \in \mathbb{R}[\underline{X}]^0 = \{0\}$ is the empty vector. This case is usually avoided in practice and in the literature by assuming d large enough but we think it is more convenient to admit it. In the pathological case $g_i = 0$, we set $r_i := -\infty$, $\ell_i := 0$ and let v_i again be the empty vector. Then

$$\mathbb{R}[\underline{X}]_{r_i} = \{a^T v_i \mid a \in \mathbb{R}^{\ell_i}\}$$

and

$$\{p^2 g_i \mid p \in \mathbb{R}[\underline{X}]_{r_i}\} = \{(a^T v_i)^2 g_i \mid a \in \mathbb{R}^{\ell_i}\} = \{a^T (g_i v_i v_i^T) a \mid a \in \mathbb{R}^{\ell_i}\}.$$

The key observation is that instead of linearizing each $p^2 g_i$ with $p \in \mathbb{R}[\underline{X}]_{r_i}$ individually, we can just linearize the symmetric matrix polynomial $g_i v_i v_i^T \in \mathbb{R}[\underline{X}]^{\ell_i \times \ell_i}$. In this way, we get for each $i \in \{0, \dots, m\}$ a linear symmetric matrix polynomial

$M_i \in \mathbb{R}[\underline{X}, (Y_\alpha)_{2\leq|\alpha|\leq d}]_1^{\ell_i\times\ell_i}$. Instead of an *infinite* family of linear inequalities, we thus get *finitely many* linear *matrix* inequalities [BEFB] (whose size depends on d) saying that

$$M_0(x,y) \succeq 0, \ldots, M_m(x,y) \succeq 0 \qquad (x \in \mathbb{R}^n, y \in \mathbb{R}^I)$$

where $I := \{\alpha \in \mathbb{N}_0^n \mid 2 \leq |\alpha| \leq d\}$ and "$\succeq 0$" means positive semidefiniteness. By defining $M \in \mathbb{R}[\underline{X}, (Y_\alpha)_{2\leq|\alpha|\leq d}]_1^{\ell\times\ell}$ with $\ell := \ell_0 + \cdots + \ell_m$ as the block diagonal matrix with blocks $M_0, \ldots, M_n$, we could even combine this into a *single* linear matrix inequality

$$M(x,y) \succeq 0 \qquad (x \in \mathbb{R}^n, y \in \mathbb{R}^I).$$

Its solution set is a spectrahedron [Vin] (in particular a semialgebraic closed convex subset of $\mathbb{R}^n$) that projects down to the convex set

$$(*) \qquad S_d(\underline{g}) := \{x \in \mathbb{R}^n \mid \exists y \in \mathbb{R}^I : M(x,y) \succeq 0\}.$$

The description $(*)$ of $S_d(\underline{g})$ is called the *degree d Lasserre relaxation* of $\underline{g}$ (or of the system of polynomial inequalities given by $\underline{g}$). By abuse of language, we call sometimes $S_d(\underline{g})$ itself the degree d Lasserre relaxation of $\underline{g}$. By construction, it is clear that each $S_d(\underline{g})$ is convex and

$$S(\underline{g}) \subseteq \ldots \subseteq S_{d+2}(\underline{g}) \subseteq S_{d+1}(\underline{g}) \subseteq S_d(\underline{g}).$$

If $S(\underline{g})$ happens to be convex, there is a certain hope that $S_k(\underline{g})$ equals $S(\underline{g})$ for all k large enough. In this case, we say that $\underline{g}$ (or the system of polynomial inequalities given by $\underline{g}$) *has an exact Lasserre relaxation*.

In this article, we provide a new sufficient criterium for $\underline{g}$ to have an exact Lasserre relaxation. To the best of our knowledge this is the strongest result currently available for convex $S(\underline{g})$.

If $S(\underline{g})$ is not convex, one can still ask whether $S_k(\underline{g})$ equals eventually *the convex hull* of $S(\underline{g})$. This seems to require very different techniques and will be studied in our forthcoming work [KS], see also Example 4.10 below.

Here we will also not address the important question asking from what k on $S(\underline{g})$ equals $S_k(\underline{g})$ in case $\underline{g}$ has an exact Lasserre relaxation. In principle, a corresponding complexity analysis of our proof would probably be possible but would, at least for general $\underline{g}$, be extremely tedious, and in the end yield a bound that is only of theoretical interest.

The Lasserre relaxation $(*)$ is a special case of the more general *semidefinite representation* of a subset $S \subseteq \mathbb{R}^n$

$$(**) \qquad S = \{x \in \mathbb{R}^n \mid \exists y \in \mathbb{R}^h : H(x,y) \succeq 0\}$$

where $H \in \mathbb{R}[\underline{X}, Y_1, \ldots, Y_h]_1^{\ell\times\ell}$ is a symmetric linear matrix polynomial for some $h, \ell \in \mathbb{N}_0$. Sets S having such a representation $(**)$ are called *semidefinitely representable*. Other commonly used terms are *projections of spectrahedra*, *spectrahedral shadows*, *spectrahedrops*, *lifted LMI sets* and *SDP-representable sets*. If the number h of additional variables is not too large, one can optimize efficiently linear functions on such sets by the use of semidefinite programming, an important generalization of linear programming [NN]. Semidefinitely representable sets are obviously convex and they are semialgebraic by Tarski's real quantifier elimination [PD, Thm. 2.1.6]. The class of semidefinitely representable sets is closed under many operations like for example taking the interior [Net]. It was asked by Nemirovski in his plenary address at the 2006 International Congress of Mathematicians in Madrid whether each convex semialgebraic set is semidefinitely representable [Nem, Subsection 4.3.1]. Helton and Nie conjectured the answer to be positive [HN2, Section 6]. In two seminal works, Scheiderer proved this conjecture for $n = 2$ [S1, Theorem 6.8] and very recently disproved it for each $n \geq 14$ [S2, Remark 4.21].

In [NPS, Theorem 3.5], it has been shown that $\underline{g}$ cannot have an exact Lasserre relaxation if $S(\underline{g}) \subseteq \mathbb{R}^n$ is convex, has nonempty interior and has at least one nonexposed face. Other obstructions to exactness have been given by Gouveia and Netzer [GN], see Theorem 4.9 below.

On the positive side, the breakthrough was the seminal work of Helton and Nie [HN2] from 2009 preceded by their earlier work [HN1], which curiously appeared later. We will the summarize the strategy behind their approach, which builds on ideas of Lasserre [L2], and indicate where this paper introduces advantageous modifications:

Let $\underline{g} := (g_1, \dots, g_m) \in \mathbb{R}[\underline{X}]^m$ and suppose $S(\underline{g})$ is convex and has nonempty interior. We will introduce in Definition 2.10 below the *d-truncated quadratic module* $M_d(\underline{g})$ associated to $\underline{g}$. It consist of *the sums* of polynomials $p^2 g_i$ with $\deg(p^2 g_i) \leq d$ (or equivalently $\deg(p) \leq r_i$, see Equation (1) above). As explained above, these were the polynomials that we add before the linearization when we build the degree d Lasserre relaxation. The following fact is good to know although we will need from it only the trivial "if" part in order to prove our Main Theorem 4.8: We have $S(\underline{g}) = S_d(\underline{g})$ if and only if all $f \in \mathbb{R}[\underline{X}]_1$ (i.e., all linear polynomials) that are nonnegative on $S(\underline{g})$ lie in $M_d(\underline{g})$, see Proposition 2.13 below.

Denoting by $M(\underline{g}) = \bigcup_{d \in \mathbb{N}} M_d(\underline{g})$ the quadratic module generated by $\underline{g}$ introduced in Definition 2.10 below, one deduces from this (due to the compactness of S) a trivial necessary condition for $\underline{g}$ having an exact Lasserre relaxation: For each $f \in \mathbb{R}[\underline{X}]_1$, there is an $N \in \mathbb{N}$ such that $f + N \in M(\underline{g})$. If $\underline{g}$ satisfies this condition, one says that $M(\underline{g})$ is Archimedean, see Proposition 2.7(d) below. This condition is unfortunately stronger than compactness of $S(\underline{g})$. In practice, this is however not too important, since a small change of the description $\underline{g}$ of $S(\underline{g})$ always makes $M(\underline{g})$ Archimedean if $S(\underline{g})$ is compact, see Remark 2.9 below.

Therefore suppose for the rest of the introduction that $M(\underline{g})$ is Archimedean.

We saw that it suffices to look at those $f \in \mathbb{R}[\underline{X}]_1$ nonnegative on $S(\underline{g})$ whose real zero set is a supporting hyperplane of the convex set $S(\underline{g})$. By Putinar's Positivstellensatz from 1993 (see [Put, Lemma 4.1], [PD, Thm. 5.3.8], [Mar, Cor. 5.6.1], [Lau]), we know that each $f \in \mathbb{R}[\underline{X}]$ *positive* on $S(\underline{g})$ lies in $M(\underline{g})$. However, this is not really what we need here. The advantage we have is that we need to consider only $f \in \mathbb{R}[\underline{X}]_1$, i.e., only linear polynomials. The problem we have to fight is however that we have f only *nonnegative* on $S(\underline{g})$ and, most importantly, we need a uniform degree bound d for which all such f are in one and the same $M_d(\underline{g})$. Such degree bounds are known for polynomials positive on $S(\underline{g})$ but depend on a measure of how close f comes to have a zero on $S(\underline{g})$ [NS', Theorem 6].

Lasserre [L2] made a first key observation to deal with this problem: He considered without loss of generality only such $f \in \mathbb{R}[\underline{X}]_1$ nonnegative on $S(\underline{g})$ that vanish in at least one point $u \in S(\underline{g})$ (and whose real zero set therefore defines a supporting hyperplane at the point u of the convex set $S(\underline{g})$ unless $f = 0$). Under a very restrictive condition, namely that the Hessians of the defining polynomials g_i have a certain matrix sums-of-squares (*sos* for short) representation (and in particular, are globally concave, which is still very restrictive), he showed that he can produce from this finitely many matrix sos representations by the use of Karush–Kuhn–Tucker (KKT) multipliers (the Lagrange multiplier technique for inequalities instead of equations [FH, Section 2.2]).

In the aforementioned articles [HN1, HN2], Helton and Nie pushed the idea of Lasserre much further and made it fruitful in many situations. There are several important ideas in their work. For those Hessians of the g_i for which the matrix sos certificate that Lasserre assumed (and which is trivial for those g_i that happen to be *linear*) does not exist, they show that in many situations, one can with a lot of new ideas still pursue the basic strategy of Lasserre. These ideas include:

- One might exchange in a very subtle way the g_i at certain places by suitable h_i having stronger concavity properties.
- Instead of looking for matrix sos representations of the Hessians themselves, they look for matrix representations of certain matrix polynomials arising from double integrals of the Hessians and depending on a parameter u that runs over part of the boundary of $S(\underline{g})$. The matrix polynomial belonging to this parameter u serves to produce the bounded degree polynomial sos certificates for those linear polynomials f defining a supporting hyperplane containing the point u.
- Instead of *assuming* the sos certificates as Lasserre did, Helton and Nie had the idea to *prove* the existence using a matrix version of Putinar's Positivstellensatz that was already available [SH, Thm. 2]. Because of the dependence of the tangent point u of the supporting hyperplane, they had to prove a version of Putinar's theorem for matrix polynomials with degree bounds similar to the one existing already for polynomials that was mentioned above (see [HN1, Thm. 29] and Theorem 2.11 below).

We modify the approach of Helton and Nie at several places, but the most important change is a new analysis of the properties of the modified polynomials h_i which are at the same time chosen slightly more carefully (see Lemma 4.5 below). This new analysis shows that the double integral mentioned above (actually already a related single integral) is negative definite even if the term under the integral is not negative semidefinite on the whole domain of integration, see Lemma 4.6 below. Helton and Nie seem to be compelled to work with negative semidefinite terms under the integral whereas the new method enables us to be more liberal about this issue.

In this way, we will be able to show our Main Theorem 4.8: If each g_i satisfies a certain second order strict quasiconcavity condition (see Definition 3.1 below) where it vanishes on $S(\underline{g})$ (which is very natural because of the convexity of S, see Proposition 3.4(b) below) or its Hessian has a matrix sos certificate for negative-semidefiniteness on S (see Definition 2.10 below), then $\underline{g}$ has an exact Lasserre relaxation.

Helton and Nie showed under the same conditions only that $S(\underline{g})$ is semidefinitely representable [HN2, Thm. 3.3]. They obtained the semidefinite representation by glueing together Lasserre relaxations of many small pieces obtained in a non-constructive way [HN2, Prop. 4.3] (see also [NS]). With a very tedious proof (using smoothening techniques similar to those from [Gho]) they show in addition under very technical assumptions not easy to state [HN2, Section 5] that there exists $s \in \mathbb{N}_0$ and $\underline{h} \in \mathbb{R}[\underline{X}]^s$ such that $S(\underline{g}) = S(\underline{h})$ and $\underline{h}$ has an exact Lasserre relaxation [HN2, Theorem 5.1]. In his diploma thesis, Sinn thoroughly analyzed and improved this proof and showed under the same technical assumptions that one can take $\underline{h} := (g_1, \dots, g_m, g_1g_2, g_1g_3, \dots, g_{m-1}g_m)$ [Sin, Theorem 3.3.2].

2. Reminder on sums of squares

In this section, we collect all the tools from the interplay between positive polynomials and sums of squares that we need from the area of real algebraic geometry.

Definition 2.1. We call $p \in \mathbb{R}[\underline{X}]$ a *sums-of-squares (sos)* polynomial if there exist $\ell \in \mathbb{N}_0$ and polynomials $p_1, \dots, p_\ell \in \mathbb{R}[\underline{X}]$ such that

$$p = p_1^2 + \dots + p_\ell^2.$$

We say that a polynomial $p \in \mathbb{R}[\underline{X}]$ is *nonnegative* (or *positive*) on a set $S \subseteq \mathbb{R}^n$ if $p(x) \geq 0$ (or $p(x) > 0$) for all $x \in S$. In this case, we write "$p \geq 0$ on S" (or "$p > 0$ on S").

It is obvious that each sos polynomial is nonnegative on $\mathbb{R}$. In Lemma 4.5 below, we will need the well-known fact that each polynomial in *one* variable nonnegative on $\mathbb{R}$ is sos.

Proposition 2.2. Let $f \in \mathbb{R}[T]$ with $f \geq 0$ on $\mathbb{R}$. Then f is sos.

Proof. Using the fundamental theorem of algebra, one shows easily that there are $p, q \in \mathbb{R}[T]$ such that $f = (p - \mathrm{i}q)(p + \mathrm{i}q) = p^2 + q^2$ where $\mathrm{i} := \sqrt{-1} \in \mathbb{C}$ is the imaginary unit. □

A matrix $A \in \mathbb{R}^{k \times k}$ is called *positive semidefinite (psd)* (or *positive definite (pd)*) if it is symmetric and $x^T Ax \geq 0$ (or $x^T Ax > 0$) for all $x \in \mathbb{R}^k \setminus \{0\}$. Equivalently, A is symmetric and the eigenvalues of A (which are all real) are all nonnegative (or positive). In this case, we write $A \succeq 0$ (or $A \succ 0$). By $A \succeq B$, $A \succ B$, $A \preceq 0$ etc., we mean $A - B \succeq 0$, $A - B \succ 0$, $-A \succeq 0$ and so on.

The appropriate generalization of Definition 2.1 to matrix polynomials is the following.

Definition 2.3. We call $P \in \mathbb{R}[\underline{X}]^{k\times k}$ a *sums-of-squares (sos)* matrix polynomial if there exist $\ell \in \mathbb{N}_0$ and $P_1, \dots, P_m \in \mathbb{R}[\underline{X}]^{k\times k}$ such that

$$P = P_1^T P_1 + \ldots + P_\ell^T P_\ell.$$

The following is an easy exercise that is good to know when dealing with sos matrix polynomials.

Proposition 2.4. For $P \in \mathbb{R}[\underline{X}]^{k\times k}$, the following are equivalent:

(a) P is an sos matrix.
(b) There is an $\ell \in \mathbb{N}_0$ and a matrix polynomial $Q \in \mathbb{R}[\underline{X}]^{\ell\times k}$ such that $P = Q^T Q$.
(c) There are $\ell \in \mathbb{N}_0$ and $v_1, \dots, v_\ell \in \mathbb{R}[\underline{X}]^k$ such that $P = v_1 v_1^T + \ldots + v_\ell v_\ell^T$.

We say that a matrix polynomial $P \in \mathbb{R}[\underline{X}]^{k\times k}$ is *psd* (or *pd*) on a set $S \subseteq \mathbb{R}^n$ if $P(x) \succeq 0$ (or $P(x) \succ 0$) for all $x \in S$. In this case, we write "$P \succeq 0$ on S" (or "$P \succ 0$ on S").

Definition 2.5. A subset M of $\mathbb{R}[\underline{X}]$ is called a *quadratic module* of $\mathbb{R}[\underline{X}]$ if

- $1 \in M$,
- $p + q \in M$ for all $p, q \in M$ and
- $p^2 q \in M$ for all $p \in \mathbb{R}[\underline{X}]$ and $q \in M$.

For a tuple $\underline{g} := (g_1, \dots, g_m) \in \mathbb{R}[\underline{X}]^m$, the smallest quadratic module containing $g_1, \dots, g_m$ is obviously

$$M(\underline{g}) := \left\{ \sum_{i=0}^{m} s_i g_i \mid s_0, \dots, s_m \in \mathbb{R}[\underline{X}] \text{ are sos} \right\}$$

where we set $g_0 := 1$. We call it the quadratic module *generated* by $\underline{g}$.

Definition 2.6. A quadratic module M of $\mathbb{R}[\underline{X}]$ is called *Archimedean* if for all $p \in M$ there is some $N \in \mathbb{N}$ such that $N + p \in M$.

The following is well-known (see for example [PD, Lemma 5.1.13] and [Mar, Cor. 5.2.4]) but for convenience of the reader we include a compact easy proof.

Proposition 2.7. Let M be a quadratic module of $\mathbb{R}[\underline{X}]$. Then the following are equivalent:

(a) M is Archimedean.
(b) There is some $N \in \mathbb{N}$ such that $N - (X_1^2 + \ldots + X_n^2) \in M$.
(c) There are $m \in \mathbb{N}$ and $\underline{g} \in (\mathbb{R}[\underline{X}]_1 \cap M)^m$ such that the polyhedron $S(\underline{g})$ is non-empty and compact.
(d) For each $f \in \mathbb{R}[\underline{X}]_1$, there is some $N \in \mathbb{N}$ such that $N + f \in M$.

Proof. Consider the vector subspace

$$B := \{p \in \mathbb{R}[\underline{X}] \mid \exists N \in \mathbb{N} : N \pm p \in M\} \supseteq \mathbb{R}$$

of $\mathbb{R}[\underline{X}]$. If $p \in \mathbb{R}[\underline{X}]$ with $p^2 \in B$, then we can choose $N \in \mathbb{N}$ such that $(N-1) - p^2 \in M$ and thus

$$N \pm p = (N-1) - p^2 + \left(\frac{1}{2} \pm p\right)^2 + \frac{3}{4} \in M$$

and thus $p \in B$. Conversely, if $p \in B$, then one can choose $N \in \mathbb{N}$ such that $2N-1 \pm p \in M$ and thus

$$N^2(2N-1) - p^2 = \frac{1}{2}\left((N-p)^2(2N-1+p) + (N+p)^2(2N-1-p)\right) \in M,$$

showing that $p^2 \in B$ since anyway $N^2(2N-1) + p^2 \in M$. Thus, we have

$$(*) \qquad p^2 \in B \iff p \in B$$

for all $p \in \mathbb{R}[\underline{X}]$. This implies that B is a subring of $\mathbb{R}[\underline{X}]$. Indeed, for $p, q \in \mathbb{R}[\underline{X}]$ with $p, q \in B$ we have

$$pq = \frac{1}{2}(\overbrace{(\underbrace{p+q}_{\in B})^2}^{\in B} - \underbrace{p^2}_{\in B} - \underbrace{q^2}_{\in B}) \in B.$$

This shows that $\mathbb{R}[\underline{X}]_1 \subseteq B \iff \mathbb{R}[\underline{X}] = B$, which is the equivalence (d) $\iff$ (a). Condition (b) is easily seen to be equivalent to $X_1^2, \dots, X_n^2 \in B$, which in turn is by $(*)$ equivalent to $X_1, \dots, X_n \in B$. Again by using that B is a subring of $\mathbb{R}[\underline{X}]$, this shows the equivalence (a) $\iff$ (b). It remains to show (c) $\iff$ (d). If (d) holds, then one trivially finds $\underline{g}$ like in (c), e.g., with $S(\underline{g})$ being a hypercube. Conversely, suppose that we have $\underline{g}$ like in (c) and let $f \in \mathbb{R}[\underline{X}]_1$. Then there is $N \in \mathbb{N}$ such that $N + f \geq 0$ on the polytope $S(\underline{g})$. By the affine form of Farkas' lemma [Scr, Cor. 7.1h, p. 93], we have that $N + f$ is a nonnegative linear combination of the $1, g_1, \dots, g_m$ and thus lies in M. □

We mention the following important theorem although we will need it only for Example 4.10 below.

Theorem 2.8 (Schmüdgen)**.** Let M be a quadratic module of $\mathbb{R}[\underline{X}]$. The following are equivalent:

(a) There are $m \in \mathbb{N}$ and $\underline{g} = (g_1, \dots, g_m) \in \mathbb{R}[\underline{X}]^m$ such that $S(\underline{g})$ is compact and $\prod_{i \in I} g_i \in M$ for all $I \subseteq \{1, \dots, m\}$.
(b) There is some $g \in M$ with compact $S(g)$.
(c) M is Archimedean.

Proof. (a) $\implies$ (c) is the deep part of Schmüdgen's Positivstellensatz [Scm, Cor. 3], namely his characterization of Archimedean *preorders* (see [PD, Thm. 5.1.17] and [Mar, Thm. 6.1.1]). The implications (c) $\implies$ (b) $\implies$ (a) are trivial. □

Remark 2.9. For $n \geq 2$, there are examples of $\underline{g} = (g_1, \dots, g_m) \in \mathbb{R}[\underline{X}]^m$ with compact (even empty) $S(\underline{g})$ such that $M(\underline{g})$ is not Archimedean (see [Mar, Ex. 7.3.1] or [PD, Ex. 6.3.1]). However if $S(\underline{g})$ is compact, then Proposition 2.7 and Theorem 2.8 provide several ways of changing the description $\underline{g}$ of $S(\underline{g})$ such that $M(\underline{g})$ becomes Archimedean. For example, if one knows a big ball containing $S(\underline{g})$, it suffices to add its defining quadratic polynomial to $\underline{g}$ by Proposition 2.7(b). That is why for many practical purposes, the Archimedean property of $M(\underline{g})$ is not much stronger than the compactness of $S(\underline{g})$.

We use the symbols ∇ and Hess to denote the gradient and the Hessian of a real-valued function of n variables, respectively. For a *polynomial* $g \in \mathbb{R}[\underline{X}]$, we understand its gradient ∇g as a column vector from $\mathbb{R}[\underline{X}]^n$, i.e., as a vector of polynomials. Similarly, its Hessian $\operatorname{Hess} g$ is a symmetric matrix polynomial of size n, i.e., a symmetric matrix from $\mathbb{R}[\underline{X}]^{n\times n}$.

Definition 2.10. Let $\underline{g} := (g_1, \dots, g_m) \in \mathbb{R}[\underline{X}]^m$ and set again $g_0 := 1$. For $i \in \{0, \dots, m\}$, set $r_i := \frac{d-\deg g_i}{2}$ if $g_i \neq 0$ and $r_i := -\infty$ if $g_i = 0$. Then we define the *d-truncated quadratic module* $M_d(\underline{g})$ associated to $\underline{g}$ by

$$M_d(\underline{g}) := \left\{ \sum_{i=0}^{m} \sum_j p_{ij}^2 g_i \mid p_{ij} \in \mathbb{R}[\underline{X}]_{r_i} \right\} \subseteq M(\underline{g}) \cap \mathbb{R}[\underline{X}]_d.$$

More generally, we define the *d-truncated $k \times k$ matricial quadratic module* associated to $\underline{g}$ by

$$M_d^{k\times k}(\underline{g}) := \left\{ \sum_{i=0}^{m} \sum_j P_{ij}^T P_{ij} g_i \mid P_{ij} \in \mathbb{R}[\underline{X}]_{r_i}^{k\times k} \right\} \subseteq \mathbb{R}[\underline{X}]_d^{k\times k}.$$

We say that $f \in \mathbb{R}[\underline{X}]$ is *$\underline{g}$-sos-concave* if

$$-\operatorname{Hess} f \in M^{n\times n}(\underline{g}) := \bigcup_{d\in\mathbb{N}_0} M_d^{n\times n}(\underline{g}).$$

If $m = 0$, this means that the negated Hessian of f is an sos matrix polynomial and we say that f is *sos-concave*.

Any $f \in \mathbb{R}[\underline{X}]_1$ is sos-concave since $\operatorname{Hess} f = 0$. The Hessian of a $\underline{g}$-sos-concave polynomial is negative semidefinite on $S(\underline{g})$.

The following is Putinar's Positivstellensatz [Put, Lemma 4.1] for matrix polynomials with degree bounds. It has been first proven by Helton and Nie [HN1, Thm. 29] following the technical approach of Nie and the second author [NS′] for the case of polynomials. This technical approach yields explicit degree bounds. The first author found a short topological proof for the mere existence of such bounds [Kri, Thm. 3.2] that is based on knowing already the result without the degree bounds that stems from [SH, Thm. 2].

Theorem 2.11 (Helton and Nie)**.** Fix $C, d, k, m, n \in \mathbb{N}$ and fix any norm on the vector space $\mathbb{R}[\underline{X}]_d^{k\times k}$. Let $\underline{g} := (g_1, \dots, g_m) \in \mathbb{R}[\underline{X}]^m$ such that $M(\underline{g})$ is Archimedean. Then there exists $d \in \mathbb{N}_0$ such that every symmetric $H \in \mathbb{R}[\underline{X}]_d^{k\times k}$ satisfying $\|H\| \le C$ and $H \succeq \frac{1}{C}$ on $S(\underline{g})$ satisfies $H \in M_d^{k\times k}(\underline{g})$.

The following is a slight generalization of [HN1, Lemma 7] that will be needed in the proof of Theorem 4.7.

Lemma 2.12. Let $d \in \mathbb{N}_0$, $\underline{g} := (g_1, \dots, g_m) \in \mathbb{R}[\underline{X}]^m$ and $u \in \mathbb{R}^n$. If $P \in M_d^{k\times k}(\underline{g})$, then the matrix polynomial $H \in \mathbb{R}[\underline{X}]^{k\times k}$ defined by

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

for $x \in \mathbb{R}^n$ lies again in $M_d^{k\times k}(\underline{g})$.

Proof. The proof [HN1, Lemma 7] can be easily adapted. Another more conceptual proof is the following: $M_d^{k\times k}(\underline{g})$ is a convex cone in a finite-dimensional vector space. Then

$$H = \int_0^1 \int_0^t P(u + s(\underline{X} - u))\, ds\, dt$$

is an existing Bochner integral of a vector valued function with values in this convex cone and thus lies again in this convex cone [RW] (regardless of whether the cone is closed or not). □

The "if" direction of the following proposition is trivial since a closed convex set in a finite-dimensional vector space is the intersection over all half spaces containing it. We will use it to prove our Main Theorem 4.8. The "only if" direction will be needed only in Example 4.10 below.

Proposition 2.13 (Netzer, Plaumann and Schweighofer)**.** Suppose $d \in \mathbb{N}_0$, $\underline{g} := (g_1, \dots, g_m) \in \mathbb{R}[\underline{X}]_d^m$, $S(\underline{g})$ is compact and convex and has nonempty interior. Then $S_d(\underline{g}) = S(\underline{g})$ if and only if every $f \in \mathbb{R}[\underline{X}]_1$ with $f \geq 0$ on $S(\underline{g})$ lies in $M_d(\underline{g})$.

Proof. This is a special case of [NPS, Proposition 3.1]. □

3. Reminder on strict quasiconcavity

We denote the *real zero set* of g by

$$Z(g) := \{x \in \mathbb{R}^n \mid g(x) = 0\}.$$

We adopt the following notion from [HN1, p. 25], which is a local second order quasiconcavity condition.

Definition 3.1. Let $g \in \mathbb{R}[\underline{X}]$. We say that g is *strictly quasiconcave* at $x \in \mathbb{R}^n$ if for all $v \in \mathbb{R}^n \setminus \{0\}$ with $(\nabla g(x))^T v = 0$, we have that $v^T(\operatorname{Hess} g(x))v < 0$. We say that g is *strictly quasiconcave* on $A \subseteq \mathbb{R}^n$ if g is *strictly quasiconcave* at each point of A.

Remark 3.2. Let $g \in \mathbb{R}[\underline{X}]$ and $x \in \mathbb{R}^n$ such that $\nabla g(x) = 0$.

(a) g is strictly quasiconcave at x if and only if $\operatorname{Hess} g(x) \prec 0$.
(b) If g is strictly quasiconcave at x and $g(x) = 0$, then there is a neighborhood U of x such that $U \cap S(g) = \{x\}$.

If $g \in \mathbb{R}[\underline{X}]$ satisfies $g(x) = 0$ and $\nabla g(x) \neq 0$, then $Z(g)$ is locally around x a smooth hypersurface. Differential geometers will recognize that strict quasiconcavity of g at x then means that the second fundamental form of this hypersurface at x is positive definite when one chooses the "outward normal" (pointing away from $S(g)$). Thus this means that $S(g)$ is locally convex in a strong second order sense. For a detailed discussion we refer to [HN1, HN2] and the references therein. As Helton and Nie in [HN1, Subsection 3.1], we want however to help those readers who are not familiar with the basics of differential geometry by discussing strict quasiconcavity in an elementary manner. The reason why we include this is that Helton and Nie presuppose already that the reader is familiar with the geometric notion of tangent hyperplanes and knows that the gradient is a normal vector for it [HN2, p. 786]. Conversely we fit this into their arguments, see Part (a) of the following lemma and Proposition 3.4(b) below.

Formally, we will use the following lemma and the next proposition only in Example 4.10 below and even there it can be avoided by some calculations. Some readers might therefore decide to skip them.

Lemma 3.3. Let $n \in \mathbb{N}$, $g \in \mathbb{R}[\underline{X}]$ and $x \in \mathbb{R}^n$ such that $g(x) = 0$ and $\nabla g(x) \neq 0$. Suppose $v_1, \dots, v_n$ form a basis of $\mathbb{R}^n$, U is an open neighborhood of 0 in $\mathbb{R}^{n-1}$, $\varphi \colon U \to \mathbb{R}$ is smooth and satisfies $\varphi(0) = 0$ as well as

$$(*) \qquad g(x + \xi_1 v_1 + \ldots + \xi_{n-1} v_{n-1} + \varphi(\xi) v_n) = 0$$

for all $\xi = (\xi_1, \dots, \xi_{n-1}) \in U$. Then the following hold:
(a) $(\nabla g(x))^T v_1 = \ldots = (\nabla g(x))^T v_{n-1} = 0 \iff \nabla\varphi(0) = 0$
(b) If $\nabla\varphi(0) = 0$ and $(\nabla g(x))^T v_n > 0$, then

$$g \text{ is strictly quasiconcave at } x \iff \operatorname{Hess}\varphi(0) \succ 0.$$

Proof. Taking the derivative of $(*)$ with respect to ξ_i, we get

$$(**) \qquad (\nabla g(x + \xi_1 v_1 + \ldots + \xi_{n-1} v_{n-1} + \varphi(\xi) v_n))^T \left(v_i + \frac{\partial\varphi(\xi)}{\partial\xi_i} v_n \right) = 0$$

for all $i \in \{1, \dots, n-1\}$. Setting here ξ to 0, we get

$$(\nabla g(x))^T \left(v_i + \frac{\partial\varphi(\xi)}{\partial\xi_i}\bigg|_{\xi=0} v_n \right) = 0$$

for each $i \in \{1, \dots, n-1\}$. From this, (a) follows easily (for "$\implies$" use that $(\nabla g(x))^T v_n \neq 0$ since $v_1, \dots, v_n$ is a basis). Taking the derivative of $(**)$ with respect to ξ_j, we get

$$\left(v_j + \frac{\partial\varphi(\xi)}{\partial\xi_j} v_n \right)^T (\operatorname{Hess} g(x + \xi_1 v_1 + \ldots + \xi_{n-1} v_{n-1} + \varphi(\xi) v_n)) \left(v_i + \frac{\partial\varphi(\xi)}{\partial\xi_i} v_n \right)$$
$$+ (\nabla g(x + \xi_1 v_1 + \ldots + \xi_{n-1} v_{n-1} + \varphi(\xi) v_n))^T \left(\frac{\partial^2\varphi(\xi)}{\partial\xi_i\partial\xi_j} v_n \right) = 0$$

for all $i, j \in \{1, \dots, n-1\}$. To prove (b), suppose now that $\nabla\varphi(0) = 0$ and $(\nabla g(x))^T v_n > 0$. Then the preceding equation implies

$$\operatorname{Hess}\varphi(0) = -\frac{1}{(\nabla g(x))^T v_n} (v_i^T (\operatorname{Hess} g(x)) v_j)_{i,j \in \{1,\dots,n-1\}}.$$

Since $v_1, \dots, v_{n-1}$ now form a basis of the orthogonal complement of $\nabla g(x)$ by (a), the matrix $(v_i^T (\operatorname{Hess} g(x)) v_j)_{i,j \in \{1,\dots,n-1\}}$ is negative definite if and only if g is strictly quasiconcave at x (see Definition 3.1). □

The following proposition is important for understanding the notion of quasiconcavity. It is trivial that quasiconcavity of a polynomial g at x depends only on the function $V \to \mathbb{R},\ x \mapsto g(x)$ where V is an arbitrarily small neighborhood of x. But if $g(x) = 0$ and $\nabla g(x) \neq 0$, then it actually depends only on the function

$$V \to \{-1, 0, 1\},\ x \mapsto \operatorname{sgn}(g(x))$$

as the equivalence of Conditions (a) and (b) of the following proposition show.

Proposition 3.4. Let $n \in \mathbb{N}$, $g \in \mathbb{R}[\underline{X}]$ and $x \in \mathbb{R}^n$ such that

$$g(x) = 0 \text{ and } \nabla g(x) \neq 0.$$

Suppose that V is a neighborhood of x. Then the following are equivalent:

(a) g is strictly quasiconcave at x.

(b) There is a basis $v_1, \ldots, v_n$ of $\mathbb{R}^n$, an open neighborhood U of 0 in $\mathbb{R}^{n-1}$ and a smooth function $\varphi\colon U \to \mathbb{R}$ such that $\varphi(0) = 0$, $\nabla\varphi(0) = 0$, $\operatorname{Hess}\varphi(0) \succ 0$,

$$(*) \qquad x + \xi_1 v_1 + \ldots + \xi_{n-1} v_{n-1} + \varphi(\xi) v_n \in Z(g) \cap V$$

for all $\xi \in U$ and

$$(**) \qquad x + \lambda v_n \in S(g) \cap V$$

for all small enough $\lambda \in \mathbb{R}_{>0}$.

(c) Condition (b) holds with "basis" replaced by "orthogonal basis".

For any basis $v_1, \ldots, v_n$ of $\mathbb{R}^n$ like in (b), one has

$$(***) \qquad (\nabla g(x))^T v_1 = \ldots = (\nabla g(x))^T v_{n-1} = 0 \qquad \text{and} \qquad (\nabla g(x))^T v_n > 0.$$

Proof. Using Lemma 3.3(a), it is easy to show that any $v_1, \ldots, v_n$ like in (b) satisfy $(***)$ using that $(\nabla g(x))^T v_n = 0$ would contradict the hypothesis $\nabla g(x) \neq 0$ since $v_1, \ldots, v_n$ is a basis. Now Part (b) of the same lemma shows that (b) implies (a). Since it is trivial that (c) implies (b), it only remains to show that (a) implies (c).

To this end, let (a) be satisfied. In order to show (c), choose an orthogonal basis $v_1, \ldots, v_n$ of $\mathbb{R}^n$ satisfying $(***)$. The implicit function theorem yields an open neighborhood U of the origin in $\mathbb{R}^{n-1}$ such that for each $\xi = (\xi_1, \ldots, \xi_{n-1}) \in U$ there is a unique $\varphi(\xi) \in \mathbb{R}$ satisfying $(*)$, in particular $\varphi(0) = 0$. Moreover, one can choose U such that the resulting function $\varphi\colon U \to \mathbb{R}$ is smooth. From $(\nabla g(x))^T v_n > 0$, we get $(**)$. From Part (a) of Lemma 3.3, we get $\nabla\varphi(0) = 0$. From Part (b) of the same lemma and from (a), we obtain $\operatorname{Hess}\varphi(0) \succ 0$. $\square$

Another more algebraic way of understanding strict quasiconcavity is given by the following easy exercise [HN1, Lemma 11(a)].

Lemma 3.5. Let $S \subseteq \mathbb{R}^n$ be a compact set and consider a polynomial $g \in \mathbb{R}[\underline{X}]$ that is strictly quasiconcave on S. Then one can find $\lambda > 0$ such that

$$\lambda \nabla g (\nabla g)^T - \operatorname{Hess} g$$

is positive definite on S.

We will need the following lemma only in the case where f is linear. In that case, one can use for its proof a slightly weaker version of the Karush-Kuhn-Tucker theorem [Pla, Theorem 5.1].

Lemma 3.6. Suppose $\underline{g} := (g_1, \ldots, g_m) \in \mathbb{R}[\underline{X}]^m$, $S(\underline{g})$ is convex and has nonempty interior. Suppose $u \in S(\underline{g})$ and let $I := \{i \in \{1, \ldots, m\} \mid g_i(u) = 0\}$. Suppose $f \in \mathbb{R}[\underline{X}]$ and U is a neighborhood of u such that u is a minimizer of f on $S(\underline{g}) \cap U$ and $\operatorname{Hess} g_i \preceq 0$ on $S(\underline{g}) \cap U$ for all $i \in I$. Then there exist a family $(\lambda_i)_{i \in I}$ of nonnegative Lagrange multipliers $\lambda_i \in \mathbb{R}_{\geq 0}$ such that $\nabla f(u) = \sum_{i \in I} \lambda_i \nabla g_i(u)$.

Proof. By the Karush-Kuhn-Tucker theorem [FH, Theorem 2.2.5], it suffices to show that the g_i ($i \in I$) satisfy the Mangasarian-Fromowitz constraint qualification, i.e., there is some $v \in \mathbb{R}^n$ such that $(\nabla g_i(u))^T v > 0$ for all $i \in I$ [FH, Chapter 2.2.5]. By discarding those g_i that are the zero polynomial, we may assume $g_i \neq 0$ for all $i \in I$. Since $S(\underline{g})$ has nonempty interior, there is then some $x \in S(\underline{g})$ such that $g_i(x) > 0$ for all $i \in I$. Set $v := x - u$ and consider for fixed $i \in I$ the function $h\colon \mathbb{R} \to \mathbb{R}$, $t \mapsto g_i(u + tv)$. We have $0 = h(0)$ and $h(1) = g_i(x) > 0$. Therefore there is $t \in [0,1]$ such that $h'(t) > 0$. Because of $h''(t) = v^T(\operatorname{Hess} g_i(u + tv))v \leq 0$ for all $t \in [0,1]$, this implies $(\nabla g_i(u))^T v = h'(0) > 0$ as desired. □

4. The main result

In this section, we will prove our main result about the exactness of the Lasserre relaxation. The first step is to get an alternate description of the compact basic closed semialgebraic set $S(\underline{g})$ with nonempty interior. Both descriptions, the original one $\underline{g}$ and the alternate one will be used in the proof of Theorem 4.7. The new description will arise by replacing polynomials g_i that are strictly quasiconcave on $S(\underline{g}) \cap Z(g_i)$ by polynomials of the form $h_i := g_i h(g_i)$ with a univariate polynomial $h \in \mathbb{R}[T]$ such that $h \geq 1$ on $\mathbb{R}$. It will be of outmost importance that $h_i \in M(\underline{g})$ which follows from the fact that $h - 1$ and therefore h is an sos-polynomial by Lemma 2.2 above. Roughly speaking, the basic idea is that $h_i(x)$ will be, up to positive factor, approximately $1 - e^{-cg_i(x)}$ for a big constant c when x lies in $S(\underline{g})$ or x lies sufficiently close to $S(\underline{g})$. The effect of this is that h_i will be a polynomial (unfortunately of large degree) that is very close to being a positive constant on the "safe part" of $S(\underline{g})$ consisting of the points in $S(\underline{g})$ that are in "safe distance" to the boundary of $S(\underline{g})$. On the "safe part" of $S(\underline{g})$ one can hope (and it will turn out from our actual choice of h) that the Hessian of the h_i does not vary too quickly. This will be crucial in the proof of Lemma 4.6 (the interval J_3 appearing there corresponds to this "safe part").

In the proof of Lemma 4.5 below, the auxiliary polynomial h will be chosen as $h := f_{c,d} \in \mathbb{Q}[T]$ for a big real constant c and a large nonnegative *even* integer d where $f_{c,d}$ is defined in Notation 4.1 below. In [HN1, Lemma 13], Helton and Nie use exactly the same polynomial $f_{c,d}$ except that they do not care about the parity of the degree d. Lemma 4.4 below is an important observation that was probably not known to Helton and Nie. If Helton and Nie had exploited this, they could have sharpened some of their results in [HN1]. However, they would not have come close to our main result Theorem 4.8 which ultimately relies on our new refined and subtle analysis in the proofs of Lemma 4.6 and Theorem 4.7 that focuses on *integrals of* the Hessian of the h_i instead of the Hessians themselves.

Notation 4.1. For $c > 0$ and $d \in \mathbb{N}_0$, we denote by

$$e_{c,d} := \sum_{k=0}^{d} \frac{c^k}{k!} T^k \in \mathbb{Q}[T]$$

the d-th Taylor polynomial of the function

$$\mathbb{R} \to \mathbb{R},\ t \mapsto \exp(ct)$$

at the origin and we set

$$f_{c,d} := \frac{1 - e_{c,d+1}(-T)}{cT} = \sum_{k=0}^{d} \frac{c^k}{(k+1)!}(-T)^k \in \mathbb{Q}[T].$$

For any $p \in \mathbb{R}[T]$, we denote by p' its (formal) derivative (with respect to T) and by $p'' = (p')'$ its second derivative.

Proposition 4.2. For $c > 0$, we have

(a) $$e'_{c,d} = ce_{c,d-1} \qquad \text{for } d \in \mathbb{N},$$

(b) $$f'_{c,d} = \frac{e_{c,d}(-T) - f_{c,d}}{T} \qquad \text{for } d \in \mathbb{N}_0 \text{ and}$$

(c) $$f''_{c,d} = \frac{-e'_{c,d}(-T) - 2f'_{c,d}}{T} \qquad \text{for } d \in \mathbb{N}.$$

Proof. Use the chain rule, the product rule and the quotient rule for derivation. □

The following lemma has been given an easy short proof by Speyer [Spe], which we reproduce here for convenience of the reader.

Lemma 4.3 (Speyer)**.** For $c \in \mathbb{R}_{>0}$ and $d \in \mathbb{N}_0$, we have:
(a) If d is even, then $e_{c,d}(t) > 0$ for all $t \in \mathbb{R}$.
(b) If d is odd, then $e_{c,d}$ is strictly increasing on $\mathbb{R}$.

Proof. We fix $c \in \mathbb{R}_{>0}$ and proceed by induction on d. The case $d = 0$ is trivial since $e_{c,0} = 1 > 0$. Suppose the lemma is already proven for $d - 1$ instead of d where $d \in \mathbb{N}$ is fixed. First consider the case where d is even. Then by induction hypothesis the odd degree polynomial $e_{c,d-1}$ must have exactly one real root t_0. By Lemma 4.2(a) the even degree polynomial $e_{c,d}$ takes therefore its (unique) minimum in t_0. To prove the statement, it suffices to observe that

$$e_{c,d}(t_0) = \frac{(ct_0)^d}{d!} + e_{c,d-1}(t_0) = \frac{(ct_0)^d}{d!} + 0 > 0.$$

In the case where d is odd, the statement follows immediately from the induction hypothesis and Lemma 4.2(a). □

Lemma 4.4. Let $c \in \mathbb{R}_{>0}$ and suppose $d \in \mathbb{N}_0$ is even. Then $f_{c,d}(t) > 0$ for all $t \in \mathbb{R}$.

Proof. The leading coefficient of $f_{c,d}$ is $\frac{c^d(-1)^d}{(d+1)!} > 0$. Therefore it suffices to show that $f_{c,d}$ has no real roots. One easily checks that $f_{c,d}$ has no root at the origin. Assume we have a root $t \in \mathbb{R}$ different from the origin. Then $e_{c,d+1}(-t) = 1$. Observing that $e_{c,d+1}(0) = 1$, it follows from Lemma 4.3(b) that $t = 0$, a contradiction. □

The following lemma is an improved version of [HN1, Lemma 13]. Most importantly, we manage to get that $h - 1$ (defined in this lemma) is an sos polynomial (and in particular h is positive on $\mathbb{R}$) instead of just positivity of h on the interval $[0, R]$. This will come out of Lemmata 4.4 and 2.2 together with the approach we take in the proof that uses simply Taylor approximations of the exponential function instead of the nonconstructive approximation theory used in [HN1]. The second crucial improvement is the new property (c). A surprising improvement coming out of Lemma 4.3 is that we get in Condition (a) positivity on $\mathbb{R}$ instead of just the positivity on $[0, R]$ that Helton and Nie get. At the moment however, we do not have any application for this. Finally, an insignificant improvement again not used by us is the validity of Condition (b) on the interval $[-R, R]$ instead of the interval $[0, R]$ used by Helton and Nie.

Lemma 4.5. Let $H, \delta, \varepsilon, R \in \mathbb{R}$ such that $H > 0$ and $0 < \delta < \varepsilon < R$. Then there exists a univariate polynomial $h \in \mathbb{R}[T]$ such that

$$h - 1 \text{ is an sos polynomial}$$

satisfying the following conditions:

$$\text{(a)} \qquad h(t) + th'(t) > 0 \text{ for all } t \in \mathbb{R}$$

$$\text{(b)} \qquad 2h'(t) + th''(t) < -H(h(t) + th'(t)) \text{ for all } t \in [-R, R]$$

$$\text{(c)} \qquad H \max\left\{h(t) + th'(t) \mid t \in [\varepsilon, R]\right\} < \min\left\{h(t) + th'(t) \mid t \in [-R, \delta]\right\}$$

Proof. By a scaling argument, we can relax the condition that $h - 1$ is sos to the condition that $h - \gamma$ is sos for some $\gamma \in \mathbb{R}_{>0}$. By the Lemmata 4.4 and 2.2, it suffices to find $c \in \mathbb{R}_{>0}$ and $d \in \mathbb{N}_0$ even such that (a)–(c) are satisfied for $h := f_{c,d} \in \mathbb{Q}[T]$. Noting that

$$f_{c,d} + Tf'_{c,d} = e_{c,d}(-T) \qquad \text{and} \qquad 2f'_{c,d} + Tf''_{c,d} = -e'_{c,d}(-T) = -ce_{c,d-1}(-T)$$

by Proposition 4.2, this means that we are trying to find $c \in \mathbb{R}_{>0}$ and $d \in \mathbb{N}_0$ even with

$$\text{(a')} \qquad e_{c,d}(-t) > 0 \text{ for all } t \in \mathbb{R}$$

$$\text{(b')} \qquad -ce_{c,d-1}(-t) < -He_{c,d}(-t) \text{ for all } t \in [-R, R]$$

$$\text{(c')} \qquad H \max\left\{e_{c,d}(-t) \mid t \in [\varepsilon, R]\right\} < \min\left\{e_{c,d}(-t) \mid t \in [-R, \delta]\right\}.$$

Condition (a') is always satisfied by Lemma 4.3(a) if d is even. Since the functions induced by the polynomials $e_{c,d}$ on the interval $[-R, R]$ converge uniformly to the function $[-R, R] \to \mathbb{R},\ t \mapsto \exp(ct)$ as $d \in \mathbb{N}$ tends to infinity, it suffices to find $c > 0$ satisfying

$$\text{(b'')} \qquad -c\exp(-ct) < -H\exp(-ct) \text{ for all } t \in [-R, R]$$

$$\text{(c'')} \qquad H \max\{\exp(-ct) \mid t \in [\varepsilon, R]\} < \min\{\exp(-ct) \mid t \in [-R, \delta]\}.$$

These conditions can be rewritten as

$$\text{(b'')} \qquad -c < -H$$

$$\text{(c'')} \qquad H\exp(-c\varepsilon) < \exp(-c\delta).$$

Thus it suffices to choose $c > \max\left\{H, \frac{\log H}{\varepsilon - \delta}\right\}$ and $d \in \mathbb{N}_0$ even and sufficiently large. $\square$

The previous result is now used to prove the following key lemma. This key lemma is our "luxury version" of [HN1, Proposition 10] in the work of Helton and Nie. It will be used in this article only with $C := S(\underline{g})$ (when $S(\underline{g})$ is compact) but for potential future applications we formulate it in greater generality. It has several advantages over [HN1, Proposition 10]. The most important one is that we only require the g_i to be strictly quasiconcave on a set that will be very slim in general whereas Helton and Nie assume them to be strictly quasiconcave on the whole of $S(\underline{g})$. Another important advantage is that the new polynomials h_i lie in $M(\underline{g})$. The only price that we have to pay is that not the Hessian itself but only an integrated version of it satisfies the negative definiteness condition. This will however be enough for the proof of Theorem 4.7 and the Main Theorem 4.8.

Lemma 4.6. Let $\underline{g} := (g_1, \dots, g_m) \in \mathbb{R}[\underline{X}]^m$ and let C be a compact subset of $S(\underline{g})$ such that g_i is strictly *quasi*-concave on $C \cap Z(g_i)$ for each $i \in \{1, \dots, m\}$. Then there exists a polynomial $h \in \mathbb{R}[T]$ with $h - 1$ an sos polynomial such that $h_i := g_i h(g_i)$ satisfies

$$\int_0^1 (\operatorname{Hess} h_i)(u + s(x-u))\, ds \prec 0$$

for all $i \in \{1, \dots, m\}$, $u \in Z(g_i)$ and $x \in \mathbb{R}^n$ with $\{u + s(x-u) \mid 0 \le s \le 1\} \subseteq C$.

Proof. By Lemma 3.5 and the compactness of $C \cap Z(g_i)$, we find $\lambda > 0$ such that

$$F_i := \lambda(\nabla g_i)(\nabla g_i)^T - \operatorname{Hess} g_i$$

satisfies

$$F_i(x) \succ 0$$

for all $i \in \{1, \dots, m\}$ and all $x \in C \cap Z(g_i)$. The polynomial h will come out of Lemma 4.5 applied to certain values of R, H, ε and δ, which we will now adjust. First of all, we choose $R > 0$ such that

$$g_i(x) \le R$$

for all $i \in \{1, \dots, m\}$ and $x \in C$. To get ε, we observe that the compact set C is contained in the union of the chain consisting of the open sets

$$\bigcap_{i=1}^{m} \left(\{x \in \mathbb{R}^n \mid g_i(x) > \varepsilon\} \cup \{x \in \mathbb{R}^n \mid F_i(x) \succ 0\}\right) \qquad (0 < \varepsilon < R)$$

and therefore is contained in those of these sets that belong to a sufficiently small ε, i.e., there is ε with $0 < \varepsilon < R$ such that

$$\forall x \in C : \forall i \in \{1, \dots, m\} : (g_i(x) \le \varepsilon \implies F_i(x) \succ 0)\,.$$

By compactness, there exists $\xi > 0$ such that

$$\forall x \in C : \forall i \in \{1, \dots, m\} : (g_i(x) \le \varepsilon \implies F_i(x) \succ \xi\, \mathrm{I}_n)\,. \tag{2}$$

We choose δ with $0 < \delta < \varepsilon$ arbitrary and $d > 0$ such that

$$\|x - y\| \le d$$

for all $x, y \in C$. The compact subset $C \times C$ of $\mathbb{R}^{2n}$ is contained in the union of the chain consisting of the open sets

$$\begin{aligned}&\{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n \mid \|x-y\| > \sigma\} \cup \\ &\quad \bigcap_{i=1}^{m} \left(\{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n \mid g_i(x) \ne 0\} \cup \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n \mid g_i(y) < \delta\}\right) \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (0 < \sigma \le d),\end{aligned}$$

and therefore is contained in those of these sets that belong to a sufficiently small σ, i.e., there is σ with $0 < \sigma \le d$ such that

$$\forall x, y \in C : (\|x - y\| \le \sigma \implies \forall i \in \{1, \dots, m\} : (g_i(x) = 0 \implies g_i(y) < \delta)). \tag{3}$$

Because C is compact, we can choose $\tau > 0$ such that

$$\|F_i(x)\| \leq \tau$$

for all $x \in C$ and $i \in \{1, \dots, m\}$. Finally, set

$$H := \max\left\{\frac{d\tau}{\sigma\xi}, \lambda\right\}.$$

Choose $h \in \mathbb{R}[T]$ such that $h - 1$ is an sos polynomial in $\mathbb{R}[T]$ according to Lemma 4.5 and the chosen values of H, R, ε and δ. Fix $i \in \{1, \dots, m\}$ and set $h_i := g_i h(g_i)$. Using the product and chain rule, we calculate

$$\nabla h_i = g_i h'(g_i) \nabla g_i + h(g_i) \nabla g_i = (h(g_i) + g_i h'(g_i)) \nabla g_i$$

and therefore

$$\operatorname{Hess} h_i = (h(g_i) + g_i h'(g_i)) \operatorname{Hess} g_i + \nabla g_i \nabla (h(g_i) + g_i h'(g_i))^T.$$

Using

$$\nabla (h(g_i) + g_i h'(g_i)) = (2h'(g_i) + g_i h''(g_i)) \nabla g_i,$$

it follows that

$$\operatorname{Hess} h_i = (h(g_i) + g_i h'(g_i)) \operatorname{Hess} g_i + (2h'(g_i) + g_i h''(g_i))(\nabla g_i)(\nabla g_i)^T.$$

One now recognizes that conditions (a) and (b) from Lemma 4.5 guarantee that

$$\begin{aligned} \operatorname{Hess} h_i(x) &\preceq \Big((h(g_i) + g_i h'(g_i)) \Big(\operatorname{Hess} g_i - H(\nabla g_i)(\nabla g_i)^T\Big)\Big)(x) \\ &\preceq \left(-(h(g_i) + g_i h'(g_i)) F_i\right)(x) \end{aligned}$$

for all $x \in C$ since $H \geq \lambda$. Now let $u \in Z(g_i)$ and $x \in \mathbb{R}^n$ with

$$\{u + s(x - u) \mid 0 \leq s \leq 1\} \subseteq C.$$

It suffices to show

$$\int_0^1 \left((h(g_i) + g_i h'(g_i)) F_i\right)(u + s(x - u))\, ds \succ 0.$$

To this end, we split up the unit interval $[0, 1]$ into three disjoint parts

$$\begin{aligned} J_1 &:= \{s \in [0,1] \mid g_i(u + s(x-u)) < \delta\}, \\ J_2 &:= \{s \in [0,1] \mid \delta \leq g_i(u + s(x-u)) \leq \varepsilon\} \text{ and} \\ J_3 &:= \{s \in [0,1] \mid g_i(u + s(x-u)) > \varepsilon\}. \end{aligned}$$

In particular, each J_k is a union of intervals such that $[0, 1] = J_1 \,\dot\cup\, J_2 \,\dot\cup\, J_3$. We now analyze the integral in question on each of these parts separately: The integral over J_1 will contribute a guaranteed amount of positive definiteness, the integral over J_2 an unknown amount of positive semidefiniteness and the integral over J_3 will be very small in norm so that it cannot destroy the positive definiteness accumulated over J_1. For further use, we set

$$M := \max\{h(s) + h'(s)s \mid s \in [\varepsilon, R]\}.$$

Analysis on J_1. The subinterval $[0, \frac{\sigma}{d}]$ of $[0, 1]$ (note that $\frac{\sigma}{d} \leq 1$) is contained in J_1 since $\|u - (u + s(x - u))\| = s\|x - u\| \leq \frac{\sigma}{d} d = \sigma$ for $s \in [0, \frac{\sigma}{d}]$ and therefore

$$g_i(u + s(x - u)) < \delta$$

for all $s \in [0, \frac{\sigma}{d}]$ by the choice of σ (see Property (3) above). By choice of ξ, we have that

$$F_i(u + s(x - u)) \succ \xi \, \mathrm{I}_n$$

for all $s \in J_1$ (in fact also for $s \in J_2$). By Parts (a) and (c) of Lemma 4.5, we have $(h(g_i) + g_i h'(g_i))(u + s(x - u)) > HM$ for all $s \in J_1$. Hence we get with Property (2) above that

$$\int_{J_1} \left((h(g_i) + g_i h'(g_i))F_i\right)(u + s(x - u))\, ds \succ \frac{\sigma}{d} HM\xi \, \mathrm{I}_n \succeq \frac{\sigma}{d}\frac{d\tau}{\sigma\xi} M\xi \, \mathrm{I}_n = \tau M \, \mathrm{I}_n\,.$$

Analysis on J_2. We have of course

$$F_i(u + s(x - u)) \succeq 0$$

for all $s \in J_2$ (in fact also for $s \in J_1$) and, by Part (a) of Lemma 4.5,

$$(h(g_i) + g_i h'(g_i))(u + s(x - u)) \geq 0$$

for all $s \in [0, 1]$. Hence

$$\int_{J_2} \left((h(g_i) + g_i h'(g_i))F_i\right)(u + s(x - u))\, ds \succeq 0.$$

Analysis on J_3. We have of course $F_i(u + s(x - u)) \succeq -\|F_i(u + s(x - u))\| \, \mathrm{I}_n \succeq -\tau \, \mathrm{I}_n$ for all $s \in [0, 1]$ and therefore

$$\int_{J_3} \left((h(g_i) + g_i h'(g_i))F_i\right)(u + s(x - u))\, ds \succeq -M\tau \, \mathrm{I}_n$$

Total analysis. Finally, we get

$$\begin{aligned}
&\int_0^1 \left((h(g_i) + g_i h'(g_i))F_i\right)(u + s(x - u))\, ds \\
&\qquad \succeq \int_{J_1} \left((h(g_i) + g_i h'(g_i))F_i\right)(u + s(x - u))\, ds \\
&\qquad + \int_{J_3} \left((h(g_i) + g_i h'(g_i))F_i\right)(u + s(x - u))\, ds \\
&\qquad\qquad \succ \tau M \, \mathrm{I}_n - M\tau \, \mathrm{I}_n = 0
\end{aligned}$$

□

Theorem 4.7. Let $\underline{g} := (g_1, \ldots, g_m) \in \mathbb{R}[\underline{X}]^m$ such that $S(\underline{g})$ is convex with nonempty interior and $M(\underline{g})$ is Archimedean. Suppose that each g_i is strictly quasiconcave on $S(\underline{g}) \cap Z(g_i)$ or $\underline{g}$-sos-concave. Then there is $d \in \mathbb{N}_0$ such that for all $f \in \mathbb{R}[\underline{X}]_1$ with $f \geq 0$ on $S(\underline{g})$ we have $f \in M_d(\underline{g})$.

Proof. Choose I and J such that $\{1, \ldots, m\} = I \,\dot\cup\, J$, g_i is strictly quasiconcave on $S(\underline{g}) \cap Z(g_i)$ for $i \in I$ and g_j is $\underline{g}$-sos-concave for $j \in J$. Applying Lemma 4.6 with $(g_i)_{i \in I}$ instead of $\underline{g}$ and the compact subset $C := S(\underline{g})$ of $S((g_i)_{i \in I})$, we get for each $i \in I$ a polynomial

$$h_i \in M(\underline{g})$$

satisfying $S(h_i) = S(g_i)$, $Z(h_i) = Z(g_i)$ and

$$\int_0^1 (\operatorname{Hess} h_i)(u + s(x - u))\, ds \prec 0$$

for all $u \in S(\underline{g}) \cap Z(g_i)$ and $x \in S(\underline{g})$. Setting here $x = u$, we obtain in particular

$$\text{Hess}\, h_i \prec 0 \text{ on } S(\underline{g}) \cap Z(g_i) \tag{4}$$

for each $i \in I$. Set

$$h_j := g_j$$

for all $j \in J$. Then

$$S(\underline{g}) = S(\underline{h}).$$

Choose $d_1 \in \mathbb{N}_0$ such that

$$h_i \in M_{d_1}(\underline{g})$$

for all $i \in I \cup J$. Define for all $i \in I \cup J$ and $u \in \mathbb{R}^n$ a symmetric matrix polynomial $H_{i,u} \in \mathbb{R}[\underline{X}]^{n\times n}$ by

$$H_{i,u}(x) = -\int_0^1 \int_0^t (\text{Hess}\, h_i)(u + s(x-u))\, ds\, dt$$

for all $x \in \mathbb{R}^n$. Applying compactness of $S(\underline{g}) \cap Z(g_i)$, $S(\underline{g})$ and the unit sphere in $\mathbb{R}^n$ together with continuity, we find $\delta > 0$ such that

$$-\int_0^1 (\text{Hess}\, h_i)(u + s(x-u))\, ds \succeq 2\delta\, \mathrm{I}_n$$

for all $i \in I, u \in S(\underline{g}) \cap Z(g_i)$ and $x \in S(\underline{g})$. For each $t \in [0,1]$, we apply this to $u + t(x-u) \in S(\underline{g})$ instead of x to get

$$-\int_0^t (\text{Hess}\, h_i)(u + s(x-u))\, ds = -t\int_0^1 (\text{Hess}\, h_i)(u + st(x-u))\, ds \succeq 2t\delta\, \mathrm{I}_n$$

for all $i \in I, u \in S(\underline{g}) \cap Z(g_i)$ and $x \in S(\underline{g})$. Thus

$$H_{i,u}(x) \succeq \int_0^1 2t\delta\, \mathrm{I}_n\ dt = \delta\, \mathrm{I}_n$$

for all $i \in I$, $u \in S(\underline{g}) \cap Z(g_i)$ and $x \in S(\underline{g})$. Again using the compactness of $S(\underline{g}) \cap Z(g_i)$ and continuity, we find some $E > 0$ such that

$$\|H_{i,u}\| \leq E$$

for all $i \in I$ and $u \in S(\underline{g}) \cap Z(g_i)$. Theorem 2.11 yields $d_2 \in \mathbb{N}$ such that

$$H_{i,u} \in M_{d_2}^{n\times n}(\underline{g})$$

for all $i \in I$ and $u \in S(\underline{g}) \cap Z(g_i)$. Lemma 2.12 yields $d_3 \in \mathbb{N}$ such that

$$H_{j,u} \in M_{d_3}^{n\times n}(\underline{g})$$

for all $j \in J$ and $u \in \mathbb{R}^n$. For later use, set

$$d_4 := \max\{d_2, d_3\} + 2 \quad \text{and} \quad d := \max\{d_1, d_4\}.$$

Now let $f \in \mathbb{R}[\underline{X}]_1$ with $f \geq 0$ on $S(\underline{g})$. Since $S(\underline{g})$ is nonempty and compact, we can define c as the minimum of f on $S(\underline{g})$. Exchanging f by $f - c$, we can suppose without loss of generality that $c = 0$. Then there is some $u \in S(\underline{g})$ with

$$f(u) = 0.$$

Consider

$$K := \{i \in I \cup J \mid g_i(u) = 0\} = \{i \in I \cup J \mid h_i(u) = 0\}.$$

Because of $\operatorname{Hess} h_i(u) \prec 0$ (see Property (4)) and continuity, we get a neighborhood U of u such that

$$\operatorname{Hess} h_i \prec 0 \text{ on } U$$

for all $i \in I \cap K$. Since each $h_j = g_j$ with $j \in J$ is $\underline{g}$-sos-concave, we have on the other hand

$$\operatorname{Hess} h_j \preceq 0 \text{ on } S(\underline{g})$$

for all $j \in J$. Combining both, we have in particular that

$$\operatorname{Hess} h_k \preceq 0 \text{ on } S(\underline{g}) \cap U$$

for all $k \in K$. Applying Lemma 3.6, we get a family $(\lambda_k)_{k\in K}$ of nonnegative Lagrange multipliers such that $\nabla f = \sum_{k\in K} \lambda_k \nabla h_k(u)$ (recall that f is linear) and thus

$$\left(f - \sum_{k\in K} \lambda_k h_k\right)(u) = 0 \text{ and } \nabla\left(f - \sum_{k\in K} \lambda_k h_k\right)(u) = 0.$$

Fix now $x \in \mathbb{R}^n$. For the map

$$h\colon \mathbb{R} \to \mathbb{R},\ s \mapsto \left(f - \sum_{k\in K} \lambda_k h_k\right)(u + s(x-u)),$$

we have $h(0) = 0$, $h'(0) = 0$ and

$$h''(s) = -\sum_{k\in K} \lambda_k (x-u)^T ((\operatorname{Hess} h_k)(u + s(x-u)))(x-u)$$

for $s \in \mathbb{R}$. Hence

$$\left(f - \sum_{k\in K} \lambda_k h_k\right)(x) = h(1) \overset{h(0)=0}{=} \int_0^1 h'(t)\, dt \overset{h'(0)=0}{=} \int_0^1 \int_0^t h''(s)\, ds\, dt$$
$$= \sum_{k\in K} \lambda_k (x-u)^T H_{k,u}(x-u).$$

Since $x \in \mathbb{R}^n$ was arbitrary, we thus have

$$f - \sum_{k\in K} \lambda_k h_k = \sum_{k\in K} \lambda_k (\underline{X} - u)^T H_{k,u}(\underline{X} - u) \in M_{d_4}(\underline{g})$$

and thus $f \in M_d(\underline{g})$. □

Note that it is essential in the previous theorem to require f to be linear. It is even not enough to require f to be globally convex of small bounded degree [KL].

Main Theorem 4.8. Let $\underline{g} := (g_1, \dots, g_m) \in \mathbb{R}[\underline{X}]^m$ such that $S(\underline{g})$ is convex with nonempty interior and $M(\underline{g})$ is Archimedean. Suppose that each g_i is strictly quasiconcave on $S(\underline{g}) \cap Z(g_i)$ or $\underline{g}$-sos-concave. Then $\underline{g}$ has an exact Lasserre relaxation.

Proof. Directly from 4.7 by the trivial direction of Proposition 2.13. □

In the situation of this theorem, now drop the convexity assumption and consequently ask whether the *convex hull* of $S(\underline{g})$ (instead of $S(\underline{g})$ itself) equals $S_d(\underline{g})$ for large d. Helton and Nie proved that in this situation the convex hull of $S(\underline{g})$ is semidefinitely representable [HN2, Theorem 4.4]. The question arises if it even equals $S_d(\underline{g})$ for large d. This will be proven in our forthcoming paper [KS] if *all* g_i are strictly quasiconcave on $S(\underline{g}) \cap Z(g_i)$. However, Example 4.10 below shows that in this case, one cannot allow that some of the g_i are linear (or even sos-concave) instead. To prove this, we need the following important criterion from [GN, Proposition 4.1].

Theorem 4.9 (Gouveia and Netzer)**.** Suppose $\underline{g} := (g_1, \dots, g_m) \in \mathbb{R}[\underline{X}]^m$, $L \subseteq \mathbb{R}^n$ is a straight line in $\mathbb{R}^n$, $S(\underline{g}) \cap L$ has nonempty interior in L and $u \in S(\underline{g})$ is an element of the boundary of $\overline{\operatorname{conv}(S(\underline{g}))} \cap L$ in L. Suppose that for each i with $g_i(u) = 0$, $\nabla g_i(u)$ is orthogonal to L. Then $S_d(\underline{g})$ *strictly* contains the closure of the convex hull of $S(\underline{g})$ for all d.

Example 4.10. Let $n := 2$, write X, Y for X_1, X_2 and consider $\underline{g} := (g_1, g_2)$ with

$$g_1 := -(1 - X^2 - Y^2)(4 - (X-4)^2 - Y^2) \text{ and } g_2 := 1 - Y.$$

We see that $S(g_1)$ is the disjoint union of two closed disks of different radii. The affine half plane $S(g_2)$ cuts out a piece from the bigger disk and its boundary line $L := \binom{0}{1} + \mathbb{R}\binom{1}{0}$ is tangent to the smaller disk. Since $S(g_1)$ is compact, $M(\underline{g})$ is Archimedean by Theorem 2.8(b). By Proposition 3.4(b), g_1 is strictly quasiconcave on $S(\underline{g}) \cap Z(g_1)$. The line L is tangent to the smaller disk in the point $\binom{0}{1}$ and passes through the interior of the larger disk. By the criterion 4.9 of Gouveia and Netzer applied with $u := \binom{0}{1}$, $S_d(\underline{g})$ strictly contains the convex hull of $S(\underline{g})$ for all d. By inspection of the proof of Gouveia and Netzer, we see more precisely that each $S_d(\underline{g})$ contains a left neighbourhood of u inside L.

Acknowledgments

The authors would like to thank all three anonymous referees for their thorough reading that helped to improve the presentation of the material.

References

[BEFB] S. Boyd, L. El Ghaoui, E. Feron, V. Balakrishnan: Linear matrix inequalities in system and control theory, SIAM Studies in Applied Mathematics 15, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994 3

[FH] W. Forst, D. Hoffmann: Optimization – theory and practice, Springer Undergraduate Texts in Mathematics and Technology, Springer, New York, 2010 5, 13

[Gho] M. Ghomi: Optimal smoothing for convex polytopes, Bull. London Math. Soc. 36 (2004), no. 4, 483–492 6

[GN] J. Gouveia, T. Netzer: Positive polynomials and projections of spectrahedra, SIAM J. Optim. 21 (2011), no. 3, 960–976 4, 21

[HN1] J.W. Helton, J. Nie: Semidefinite representation of convex sets, Math. Program. 122 (2010), no. 1, Ser. A, 21–64 4, 5, 9, 10, 12, 13, 14, 15, 21

[HN2] J.W. Helton, J. Nie: Sufficient and necessary conditions for semidefinite representability of convex hulls and sets, SIAM J. Optim. 20 (2009), no. 2, 759–791 [this article is a continuation of [HN1] although it appeared earlier] 4, 5, 6, 10, 20

[KL] E. de Klerk, M. Laurent: On the Lasserre hierarchy of semidefinite programming relaxations of convex polynomial optimization problems, SIAM J. Optim. 21 (2011), no. 3, 824–832 20

[Kri] T. Kriel: A new proof for the existence of degree bounds for Putinar's Positivstellensatz, Ordered algebraic structures and related topics 203–209, Contemp. Math., 697, Amer. Math. Soc., Providence, RI, 2017 9

[KS] T-L. Kriel, M. Schweighofer: On the exactness of Lasserre relaxations and pure states over real closed fields, preprint, October 20, 2017 [https://arxiv.org/abs/1710.07521] 3, 20

[L1] J.B. Lasserre: Global optimization with polynomials and the problem of moments, SIAM J. Optim. 11 (2000/01), no. 3, 796–817 2

[L2] J.B. Lasserre: Convex sets with semidefinite representation, Math. Program. 120 (2009), no. 2, Ser. A, 457–477 2, 4, 5

[Lau] M. Laurent; Sums of squares, moment matrices and optimization over polynomials, Emerging applications of algebraic geometry, 157–270, IMA Vol. Math. Appl., 149, Springer, New York, 2009, [updated version available at http://homepages.cwi.nl/~monique/files/moment-ima-update-new.pdf] 4

[Mar] M. Marshall: Positive polynomials and sums of squares, Mathematical Surveys and Monographs 146, American Mathematical Society, Providence, RI, 2008 4, 7, 8

[Nem] A. Nemirovski: Advances in convex optimization: conic programming, International Congress of Mathematicians. Vol. I, 413–444, Eur. Math. Soc., Zürich, 2007 4

[Net] T. Netzer: On semidefinite representations of non-closed sets, Linear Algebra Appl. 432 (2010), no. 12, 3072–3078 4

[NN] Y. Nesterov, 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 3

[NPS] T. Netzer, D. Plaumann, M. Schweighofer: Exposed faces of semidefinitely representable sets, SIAM J. Optim. 20 (2010), no. 4, 1944–1955 4, 10

[NS] T. Netzer, R. Sinn: A note on the convex hull of finitely many projections of spectrahedra, preprint, August 24, 2009 [https://arxiv.org/abs/0908.3386] 6

[NS'] J. Nie, M. Schweighofer: On the complexity of Putinar's Positivstellensatz, J. Complexity 23 (2007), no. 1, 135–150 5, 9

[PD] A. Prestel, C.N. Delzell: Positive polynomials, From Hilbert's 17th problem to real algebra, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2001 2, 4, 7, 8

[Pla] D. Plaumann: Geometry of linear matrix inequalities, lecture notes, Universität Konstanz, 2013 [http://www.math.uni-konstanz.de/~plaumann/LMI13/lmi.pdf] 12

[Put] M. Putinar: Positive polynomials on compact semi-algebraic sets, Indiana Univ. Math. J. 42 (1993), no. 3, 969–984 4, 9

[RW] H. Rubin, O. Wesler: A note on convexity in Euclidean n-space, Proc. Amer. Math. Soc. 9, 1958, 522–523 10

[S1] C. Scheiderer: Semidefinite representation for convex hulls of real algebraic curves, preprint, September 18, 2017 [https://arxiv.org/abs/1208.3865] 4

[S2] C. Scheiderer: Spectrahedral shadows, preprint, December 4, 2017 [https://arxiv.org/abs/1612.07048] 4

[Scr] A. Schrijver: Theory of linear and integer programming, Wiley-Interscience Series in Discrete Mathematics, Chichester, 1986 2, 8

[Scm] K. Schmüdgen: The K-moment problem for compact semi-algebraic sets, Math. Ann. 289 (1991), no. 2, 203–206 8

[SH] C.W. Scherer, C.W.J. Hol: Matrix sum-of-squares relaxations for robust semi-definite programs, Math. Program. 107 (2006), no. 1–2, Ser. B, 189–211 5, 9

[Sin] R. Sinn: Spectrahedra and a relaxation of convex, semialgebraic sets, Diplomarbeit – slightly improved version, Universität Konstanz, September 2010 [http://www.math.uni-konstanz.de/~sinn/thesis.pdf] 6

[Spe] D. Speyer: entry in the interactive mathematics website MathOverflow, November 6, 2009 [http://mathoverflow.net/a/4410] 14

[Vin] C. Vinzant: What is ... a spectrahedron?, Notices Amer. Math. Soc. 61 (2014), no. 5, 492–494 3

FACHBEREICH MATHEMATIK UND STATISTIK, UNIVERSITÄT KONSTANZ, 78457 KONSTANZ, GERMANY

E-mail address: tom-lukas.kriel@uni-konstanz.de

FACHBEREICH MATHEMATIK UND STATISTIK, UNIVERSITÄT KONSTANZ, 78457 KONSTANZ, GERMANY

E-mail address: markus.schweighofer@uni-konstanz.de