HAL open science

Exact Moment Representation in Polynomial Optimization

Lorenzo Baldi, Bernard Mourrain

▶ **To cite this version:**

Lorenzo Baldi, Bernard Mourrain. Exact Moment Representation in Polynomial Optimization. 2022. hal-03082531v4

HAL Id: hal-03082531

https://hal.science/hal-03082531v4

Preprint submitted on 27 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Exact Moment Representation in Polynomial Optimization

Lorenzo Baldi, Bernard Mourrain

Inria, Université Côte d'Azur, Sophia Antipolis, France, April 27, 2022.

In memory of Carlo Casolo

Abstract

We investigate the problem of representing moment sequences by measures in the context of Polynomial Optimization Problems. This consists in finding the infimum of a real polynomial on a real semialgebraic set defined by polynomial inequalities. We analyze the exactness of Moment Matrix (MoM) relaxations, dual to the Sum of Squares (SoS) relaxations, which are hierarchies of convex cones introduced by Lasserre to approximate measures and positive polynomials. We investigate in particular flat truncation properties, which allow testing effectively when MoM exactness holds.

We consider the quadratic module Q generated by the inequalities. We show that the dual of the MoM relaxation coincides with the SoS relaxation extended with the real radical of the support of Q, and focus on the zero-dimensional case, generalizing results for equations defining a finite real variety. We deduce sufficient and necessary conditions for flat truncation, under the finite convergence assumption: flat truncation happens if and only if the support of the quadratic module associated with the minimizers is of dimension zero. We also bound the order of the relaxation at which flat truncation holds.

As corollaries, we conclude that flat truncation holds:

- when regularity conditions, known as Boundary Hessian Conditions, hold: this result implies that flat truncation and MoM exactness holds generically;
- when the support of the quadratic module Q is zero-dimensional;
- in singular cases, flat truncation holds for the MoM relaxation extended with the polar constraints when the real variety of polar points is finite.

Effective numerical computations illustrate these flat truncation properties.

1 Introduction

Let $f, g_1, \dots, g_s \in \mathbb{R}[X_1, \dots, X_n]$ be polynomials in the indeterminates $X_1, \dots, X_n$ with real coefficients. The goal of Polynomial Optimization is to find:

$$f^* := \inf \left\{ f(x) \in \mathbb{R} \mid x \in \mathbb{R}^n,\ g_i(x) \geq 0 \ \text{ for } i = 1, \dots, s \right\}, \tag{1}$$

that is the infimum f^* of the *objective function* f on the *basic semialgebraic set* $S := \{x \in \mathbb{R}^n \mid g_i(x) \geq 0 \ \text{ for } i = 1, \dots, s\}$. It is a general problem, which appears in many contexts (e.g. real solution of polynomial equations, ...) and with many applications. To cite a few of them: in graph theory [LV21], network optimization design [MH15], control [HK14], ... See [Las10] for a more comprehensive list.

To solve this NP-hard problem, Lasserre [Las01] proposed to use two hierarchies of finite dimensional convex cones depending on an order $d \in \mathbb{N}$ and he proved, for Archimedean quadratic modules, the convergence when $d \to \infty$ of the optima associated to these hierarchies to the minimum f^* of f on S. The first hierarchy replaces non-negative polynomials by Sums of Squares (SoS) and non-negative polynomials on S by polynomials of degree $\leq d$ in the truncated quadratic module $\mathcal{Q}_{2d}(\mathbf{g})$ generated by the tuple of polynomials $\mathbf{g} = g_1, \dots, g_s$.

The second and dual hierarchy replaces positive measures by linear functionals $\in \mathcal{L}_{2d}(\mathbf{g})$ which are non-negative on the polynomials of the truncated quadratic module $\mathcal{Q}_{2d}(\mathbf{g})$. We will describe more precisely these constructions in section 2.1.

This approach has many interesting properties (see e.g. [Las15], [Lau09], [Mar08]). It was proposed with the aim to recover the infimum f^* and, if this infimum is reached, the set of minimizers $\{\xi \in S \mid f(\xi) = f^*\}$. The extraction of minimizers is strongly connected to the so called *flat truncation* property, that will be the focus of the paper.

To tackle these challenges, one can first address the finite convergence problem, that is when the value f^* can be obtained at a given order of the relaxation(s). The second problem is the exactness of the relaxations, which is the main topic of this paper. The Sum of Squares (SoS) exactness is when the non-negative polynomial $f - f^*$ belongs to the truncated quadratic module $\mathcal{Q}_{2d}(\mathbf{g})$ for some $d \in \mathbb{N}$. The Moment Matrix (MoM) exactness is when, for some $d \in \mathbb{N}$, any optimal linear functional $\sigma^* \in \mathcal{L}_{2d}(\mathbf{g})$ for f is coming from a positive measure supported on S. We are going to investigate in detail this MoM exactness property.

Several works have been developed over the last decades to address SoS representation problems. [Par02] showed that if the complex variety $\mathcal{V}_{\mathbb{C}}(I)$ defined by an ideal I generated by real polynomials is finite and I is radical, then $f - f^*$ has a representation as a sum of squares modulo I. [Lau07] showed the finite convergence property if the complex variety $\mathcal{V}_{\mathbb{C}}(I)$ is finite, and a moment sequence representation property, if moreover the ideal I is radical. [Nie13c] showed that if the semialgebraic set S is finite, then the finite convergence property holds for a preordering defining S. [Sch05a] proved that $f - f^*$ is in the quadratic module $\mathcal{Q}$ defining S modulo $(f - f^*)^2$ if and only if $f - f^* \in \mathcal{Q}$ and then the SoS relaxation is exact. [Mar06], [Mar09] proved that, under regularity conditions at the minimizers, known as Boundary Hessian Conditions (BHC), $f - f^*$ is in the quadratic module and the SoS exactness property holds. [NDS06], [DNP07] showed that, by adding gradient constraints when $S = \mathbb{R}^n$ or KKT constraints when S is a general basic semialgebraic set, the SoS exactness property holds when the corresponding Jacobian ideal is radical. [Nie13a] showed that, by adding the Jacobian constraints, the finite convergence property holds under some regularity assumption. [Nie13b] showed that finite convergence and the flat truncation property are equivalent under generic assumptions, if the SoS relaxation is exact and strong duality holds. In [Nie14], it is shown that BHC imply finite convergence and that BHC are generic. [KS19] showed the SoS exactness property if the quadratic module defining S is Archimedian and some strict concavity properties of f at the finite minimizers are satisfied.

Though many works focused on the SoS relaxation and the representation of positive polynomials with Sums of Squares, the MoM relaxation has been much less studied. We mention [LLR08] and [Las+13], which prove that if S is finite, the value f^*, the minimizers and the vanishing ideal of S can be recovered from moment matrices associated with the truncated preordering defining S; and [Nie13b], which shows that finite convergence and the flat truncation property are equivalent under generic assumptions, if the SoS relaxation is exact and strong duality holds.

From a methodological and practical point of view, flat truncation tests on moment matrices [CF98], [LM09] are essentially the only known way to decide finite convergence, i.e. whether the minimum f^* is reached at some order of the relaxation. But flat truncation also implies MoM exactness. Moreover, it allows extracting the finite minimizers from moment matrices [HL05], [Mou18], whereas SoS exactness does not yield the minimizers. Therefore a natural question, of theoretical and practical importance, is:

When flat truncation holds in a Polynomial Optimization Problem?

It is known that truncated minimizing (pseudo-)moment sequences are not always coming from measures, as shown for instance by Examples 2.7, 2.8, 2.9, 2.10, 2.11. Thus flat truncation does not hold in general. But surprisingly, for regular Polynomial Optimization Problems, this question remained open for more than a decade.

Contributions. Our first main contribution to this problem is given in Theorem 4.4. We prove that, under finite convergence assumption, flat truncation is equivalent to having the support of the

quadratic module associated with the minimizers of dimension zero. Moreover, we give conditions on the order of the relaxation to have flat truncation.

As a corollary we deduce Theorem 4.7, where we show that flat truncation and MoM exactness hold for regular problems which satisfy the Boundary Hessian Conditions. It also provides new bounds on the order where flat truncation holds. This generalizes the results on finite convergence and SoS exactness proved in [Mar06], [Mar09], and [Nie14]. It also shows that flat truncation and MoM exactness holds generically (Corollary 4.8).

Another consequence of Theorem 4.4, shown in Theorem 4.10, is that when the set S is finite, flat truncation holds if the quotient by the support of the quadratic module $\mathcal{Q}$ is of dimension zero. This generalizes results of [LLR08] and [Las+13] on semi-definite moment representations on finite sets. Moreover, we provide new bounds on the order where flat truncation holds.

When the problem is not regular, we introduce polar equations, that can help achieving flat truncation and MoM exactness. In Theorem 4.14, we prove that if the real variety of polar points is finite then the relaxation extended with Jacobian constraints is MoM exact and the flat truncation property is satisfied. This generalizes the results of finite convergence and SoS exactness of the KKT and Jacobian relaxations under regularity conditions, proved in [NDS06], [DNP07], [Nie13a].

To prove these results, we investigate in detail the properties of truncated moment relaxations and their duals. In Theorem 3.4, we provide a new description of the dual of the MoM relaxation in terms of the initial SoS relaxation and the real radical of the support of the associated quadratic module Q. A key ingredient to analyze flat truncation is Theorem 3.21. When the quotient by the support of Q is of dimension zero, we prove that the (truncated) linear functionals in $\mathcal{L}_d(\mathbf{g})$ coincide with the measures supported on S, that is the convex hull of the evaluations at the points of S. Moreover, the ideal generated by the annihilator of a generic element in $\mathcal{L}_d(\mathbf{g})$ is the vanishing ideal of S.

The paper is structured as follows. In the next section of the introduction, we define the algebraic objects that we will use and recall their main properties. In Section 2, we describe in detail the notions of finite convergence and exactness for the Sum of Squares (SoS) and Moment Matrix (MoM) relaxations. We give several examples showing how these notions are related.

In Section 3, we investigate the properties of truncated moment sequences (Section 3.1), of their annihilators (Section 3.2) and we analyze when flat truncation holds and relate it with the regularity of S (Section 3.3).

In Section 4, we apply these results to Polynomial Optimization Problems (POPs). In Section 4.1, we prove necessary and sufficients conditions for flat truncation and analyze at which degree flat truncation holds and yields the minimizers. We prove that exactness and flat truncation hold for POPs satisfying the Boundary Hessian Conditions (Section 4.2), for finite semialgebraic sets (Section 4.3), and finally for POPs with polar constraints (Section 4.4).

For the numerical computations performed on the examples, which illustrate these developments, we use the Julia package `MomentTools.jl`[1] with the SDP solvers `Mosek` and `SDPA`, based on interior point methods.

1.1 Notation

We provide here the basic definitions we need hereafter and refer to [Mar08] and [Mou18] for more details.

If A is a subset of a $\mathbb{R}$-vector space V, we denote by $\mathrm{cone}(A)$ the convex cone generated by A, by $\mathrm{conv}(A)$ its convex hull and by $\langle A\rangle$ its linear span.

Let $\mathbb{R}[\mathbf{X}] := \mathbb{R}[X_1,\dots,X_n]$ be the $\mathbb{R}$-algebra of polynomials in n indeterminates $X_1,\dots,X_n$. We denote $(h_1,\dots,h_r) \subset \mathbb{R}[\mathbf{X}]$ the *ideal* generated by $h_1,\dots,h_r \in \mathbb{R}[\mathbf{X}]$.

We say that the tuple of polynomials $\mathbf{h} = h_1,\dots,h_s$ is a *graded basis* of an ideal I if for all $p \in I$, there exists $q_i \in \mathbb{R}[\mathbf{X}]$ with $\deg(q_i) \le \deg(p) - \deg(h_i)$ such that $p = \sum_{i=1}^s h_i\, q_i$. Equivalently, we have for all $t \in \mathbb{N}$, $\langle \mathbf{h}\rangle_t := \{p = \sum_{i=1}^s h_i\, q_i \mid \deg(q_i) \le t - \deg(h_i)\} = I_t := \{f \in I \mid \deg f \le t\}$, i.e. for all t the

[1] https://gitlab.inria.fr/AlgebraicGeometricModeling/MomentTools.jl

truncated ideal in degree t generated by $\mathbf{h}$ is equal to all the polynomials of degree $\leq t$ in the ideal. If $\mathbf{h}$ is not a graded basis, the inclusion $\langle \mathbf{h} \rangle_t \subset I_t$ is strict. A graded basis of an ideal $I = (\mathbf{h})$ can be computed as a Grobner basis using a monomial ordering $\prec$, which refines the degree ordering (see e.g. [CLO15]). It can also be computed as a border basis for a monomial basis of least degree of $\mathbb{R}[\mathbf{X}]/I$ (see e.g. [MT05]).

$\dim \frac{\mathbb{R}[\mathbf{X}]}{I}$ will denote the Krull dimension of $\frac{\mathbb{R}[\mathbf{X}]}{I}$.

If $A \subset \mathbb{R}[\mathbf{X}]$ and $t \in \tilde{\mathbb{N}} := \mathbb{N} \cup \{\infty\}$, $A_t := \{f \in A \mid \deg f \leq t\}$. In particular $\mathbb{R}[\mathbf{X}]_t$ is the vector space of polynomials of degree $\leq t$. When $t = \infty$, we will omit the index: $A_\infty = A$. By convention, ∞ is an idempotent element for all the operations $+, *$ in the ordered ring $\tilde{\mathbb{N}}$.

Given a finite tuple of polynomials $\mathbf{g} = g_1, \dots, g_r$, we define $\Pi\mathbf{g} := \prod_{j \in J} g_j \colon J \subset \{1, \dots, r\} = g_1, \dots, g_r, g_1 g_2, \dots, g_1 \cdots g_r$, the tuple of all the products of the g_i's, and $\pm\mathbf{g} := g_1, -g_1, \dots, g_r, -g_r$.

If $A \subset \mathbb{R}[\mathbf{X}]$ we define $\mathcal{S}(A) := \left\{x \in \mathbb{R}^n \mid f(x) \geq 0 \ \forall f \in A\right\}$. In particular we denote $\mathcal{S}(\mathbf{g}) = \left\{x \in \mathbb{R}^n \mid g(x) \geq 0 \ \forall g \in \mathbf{g}\right\}$ (the *basic semialgebraic set defined by* $\mathbf{g}$).

We denote by $\mathrm{Pos}(S) = \{f \in \mathbb{R}[\mathbf{X}] : \forall x \in S, f(x) \geq 0\}$ the cone of positive polynomials on S.

Quadratic modules. Let $\Sigma^2 = \Sigma^2[\mathbf{X}] := \left\{f \in \mathbb{R}[\mathbf{X}] \mid \exists r \in \mathbb{N},\ g_i \in \mathbb{R}[\mathbf{X}] \colon f = g_1^2 + \cdots + g_r^2\right\}$ be the convex cone of *Sum of Squares polynomials* (SoS).

$Q \subset \mathbb{R}[\mathbf{X}]$ is called *quadratic module* if $1 \in Q$, $\Sigma^2 \cdot Q \subset Q$ and $Q + Q \subset Q$. If in addition $Q \cdot Q \subset Q$, Q is a *preordering*. For $Q \subset \mathbb{R}[\mathbf{X}]$, we define $\operatorname{supp} Q := Q \cap -Q$. If Q is a quadratic module then $\operatorname{supp} Q$ is an ideal.

We say that a quadratic module Q is *finitely generated* (f.g.) if $\exists g_1 \dots g_l \in \mathbb{R}[\mathbf{X}] \colon Q = \mathcal{Q}(g_1, \dots, g_l) := \Sigma^2 + \Sigma^2 \cdot g_1 + \cdots + \Sigma^2 \cdot g_l$ (it is the smallest quadratic module containing $g_1, \dots, g_l$).

We say that a quadratic module Q is *Archimedean* if $\exists\ 0 \leq r \in \mathbb{R} \colon r - \|\mathbf{X}\|^2 \in Q$. Notice that if Q is Archimedean then $\mathcal{S}(Q)$ is compact.

For $d \in \tilde{\mathbb{N}}$ and $\mathbf{g} \subset \mathbb{R}[\mathbf{X}]$, let

$$\mathcal{Q}_d(\mathbf{g}) := \left\{s_0 + \sum_{j=1}^{r} s_j g_j \in \mathbb{R}[\mathbf{X}]_d \mid r \in \mathbb{N},\ g_j \in \mathbf{g},\ s_0 \in \Sigma_d^2,\ s_j \in \Sigma_{d - \deg g_j}^2\right\}$$

be the *truncated quadratic module* generated by $\mathbf{g}$. Notice that $\mathcal{Q}_d(\mathbf{g}) \subset \mathcal{Q}(\mathbf{g})_d = \mathcal{Q}(\mathbf{g}) \cap \mathbb{R}[\mathbf{X}]_d$, but the inclusion is strict in general. We denote $\mathcal{O}_d(\mathbf{g}) := \mathcal{Q}_d(\Pi\mathbf{g})$ the *truncated preordering* generated by $\mathbf{g}$.

Linear functionals. For a $\mathbb{R}$-vector space V, $v \in V$, $\sigma \in V^*$, we denote $\langle \sigma | v \rangle = \sigma(v)$ the application of σ to $v \in V$. For $A \subset V$, we define $A^\perp := \left\{\sigma \in V^* \mid \langle \sigma | a \rangle = 0 \ \forall a \in A\right\}$ and $A^\vee := \left\{\sigma \in V^* \mid \langle \sigma | a \rangle \geq 0 \ \forall a \in A\right\}$.

Recall that $(\mathbb{R}[\mathbf{X}])^* \cong \mathbb{R}[[\mathbf{Y}]] := \mathbb{R}[[Y_1, \dots, Y_n]]$, with the isomorphism given by: $(\mathbb{R}[\mathbf{X}])^* \ni \sigma \mapsto \sum_{\alpha \in \mathbb{N}^n} \langle \sigma | \mathbf{X}^\alpha \rangle \frac{\mathbf{Y}^\alpha}{\alpha!} \in \mathbb{R}[[\mathbf{Y}]]$, where $\{\frac{\mathbf{Y}^\alpha}{\alpha!}\}$ is dual to $\{\mathbf{X}^\alpha\}$, i.e. $\left\langle \mathbf{Y}^\alpha \middle| \mathbf{X}^\beta \right\rangle = \alpha! \, \delta_{\alpha,\beta}$. See [Mou18] for more details. We can identify $\sigma \in (\mathbb{R}[\mathbf{X}])^*$ with its sequence of coefficients $(\sigma_\alpha)_\alpha$ (called *moments*, in analogy to the case of a measure), where $\sigma_\alpha := \langle \sigma | \mathbf{X}^\alpha \rangle$.

If $t \leq s \in \tilde{\mathbb{N}}$ and $\sigma \in (\mathbb{R}[\mathbf{X}]_s)^*$, then $\sigma^{[t]} \in (\mathbb{R}[\mathbf{X}]_t)^*$ denotes its restriction to $\mathbb{R}[\mathbf{X}]_t$. Similarly if $B \subset (\mathbb{R}[\mathbf{X}]_s)^*$ then $B^{[t]} := \{\sigma^{[t]} \in (\mathbb{R}[\mathbf{X}]_t)^* \mid \sigma \in B\}$.

We give to $\mathbb{R}[\mathbf{X}]$ and $(\mathbb{R}[\mathbf{X}])^*$ the locally convex topology defined as follows. If $V = \mathbb{R}[\mathbf{X}]$ or $V = (\mathbb{R}[\mathbf{X}])^*$ and $W \subset V$ is a finitely dimensional vector subspace, W is equipped with the Euclidean topology. We define $U \subset V$ open if and only if $U \cap W$ is open in W for every finitely dimensional vector subspace W.

Measures. We will consider Borel measures with support included in $S \subset \mathbb{R}^n$, denoted as $\mathcal{M}(S)$, as linear fuctionals, i.e. $\mathcal{M}(S) \subset (\mathbb{R}[\mathbf{X}])^*$. In this case the sequence $(\mu_\alpha)_\alpha$ associated with a measure μ is the sequence of *moments*: $\mu_\alpha = \int X^\alpha \, \mathrm{d}\mu$. Moreover $\mathcal{M}^{(1)}(S)$ will denote the Borel *probability* measures supported on S. We recall a version of Haviland's theorem [Mar08, th. 3.1.2]: if $\sigma \in (\mathbb{R}[\mathbf{X}])^*$, then $\sigma \in \mathcal{M}(S)$ if and only if $\forall f \in \mathrm{Pos}(S)$, $\langle \sigma | f \rangle \geq 0$. In particular we are interested in *evaluations* or *Dirac measures*: if $\xi \in \mathbb{R}^n$ then $\mathbf{e}_\xi(f) = \left\langle \mathbf{e}_\xi \middle| f \right\rangle = \int f \, \mathrm{d}\mathbf{e}_\xi = f(\xi)$ for all $f \in \mathbb{R}[\mathbf{X}]$.

Moment matrices. Let m_g denote the multiplication operator by the polynomial g. For $t \geq r \in \tilde{\mathbb{N}}$, $\sigma \in (\mathbb{R}[\mathbf{X}]_r)^*$ and $g \in \mathbb{R}[\mathbf{X}]_t$, we define the *convolution of g and σ* as $g \star \sigma := \sigma \circ m_g \in (\mathbb{R}[\mathbf{X}]_{r-t})^*$ (i.e.

$\langle g \star \sigma | f \rangle = \langle \sigma | g f \rangle \ \forall f$) and the *Hankel operator* $H_\sigma^t \colon \mathbb{R}[\mathbf{X}]_t \to (\mathbb{R}[\mathbf{X}]_{r-t})^*$, $g \mapsto g \star \sigma$. If $\sigma = (\sigma_\alpha)_\alpha$ and $g = \sum_\alpha g_\alpha X^\alpha$ then $g \star \sigma = (\sum_\beta g_\beta \sigma_{\alpha+\beta})_\alpha$. Notice that $g \star \sigma = 0 \iff H_{g\star\sigma} = 0$.

We denote by $\mathrm{Ann}_t(\sigma)$ the *annihilator* of σ w.r.t. $\star$ in degree $\leq t$, that is $\mathrm{Ann}_t(\sigma) = \ker H_\sigma^t = \{p \in \mathbb{R}[\mathbf{X}]_t \mid p \star \sigma = 0\}$. The *moment matrix* of σ in degree t is the matrix $H_\sigma^t = (\sigma_{\alpha+\beta})_{|\alpha|\leq t, |\beta|\leq t}$ of the Hankel operator H_σ^t with respect to the basis $\{\mathbf{X}^\beta\}$ and $\{\frac{\mathbf{Y}^\alpha}{\alpha!}\}$. Therefore the kernel of the moment matrix is the annihilator of σ.

Notice that, if $s \leq t$, we can identify the matrix of H_σ^s with the submatrix of H_σ^t indexed by monomials of degree $\leq s$.

The *localizing matrix* of $g \in \mathbb{R}[\mathbf{X}]_t$ is the matrix $H_{g\star\sigma}^t = ((g \star \sigma)_{\alpha+\beta})_{\alpha,\beta} = (\sum_\gamma g_\gamma \sigma_{\alpha+\beta+\gamma})_{\alpha,\beta}$ of the Hankel operator $H_{g\star\sigma}^t$.

Positive linear functionals. We say that $\sigma \in (\mathbb{R}[\mathbf{X}]_{2t})^*$ is *positive semidefinite* (psd) $\iff H_\sigma^t$ is psd, i.e. $\left\langle H_\sigma^t(f) \middle| f \right\rangle = \left\langle \sigma \middle| f^2 \right\rangle \geq 0 \ \forall f \in \mathbb{R}[\mathbf{X}]_t$ (see [Sch17] or [Mar08] for basic properties of psd matrices). Recall that if $\sigma \in (\mathbb{R}[\mathbf{X}]_{2t})^*$ is psd and $\left\langle \sigma \middle| f^2 \right\rangle = 0$ then $\forall p \in \mathbb{R}[\mathbf{X}]_t, \langle \sigma | f p \rangle = 0$, see [Las+13, lem. 3.12].

For $d \in \tilde{\mathbb{N}}$ and $\mathbf{g} \subset \mathbb{R}[\mathbf{X}]_d$, we define:

$$\mathcal{L}_d(\mathbf{g}) := \mathcal{Q}_d(\mathbf{g})^\vee = \{\sigma \in (\mathbb{R}[\mathbf{X}]_d)^* \mid \forall q \in \mathcal{Q}_d(\mathbf{g}) \ \langle \sigma | q \rangle \geq 0\}$$

the cone of *positive linear functionals* on $\mathcal{Q}_d(\mathbf{g})$. Equivalently $\sigma \in \mathcal{L}_d(\mathbf{g})$ if and only if $\langle \sigma | s \rangle \geq 0 \ \forall s \in \Sigma_d^2$ and $\langle \sigma | s f \rangle \geq 0 \ \forall f \in \mathbf{g}, \forall s \in \Sigma^2$ such that $\deg f s \leq d$. Another equivalent way to describe $\mathcal{L}_d(\mathbf{g})$ is using positive semidefinite matrices (*Linear Matrix Inequalities*). Indeed $\sigma \in \mathcal{L}_d(\mathbf{g})$ if and only if the matrices $H_\sigma^{\lfloor \frac{d}{2} \rfloor}, H_{g_1\star\sigma}^{\lfloor \frac{d-\deg g_1}{2} \rfloor}, \ldots, H_{g_r\star\sigma}^{\lfloor \frac{d-\deg g_r}{2} \rfloor}$ are positive semidefinite.

Notice that $\mathcal{L}_d(\mathbf{g})$ is the *dual convex cone* to $\mathcal{Q}_d(\mathbf{g})$, see [Mar08, sec. 3.6]. By conic duality, $\overline{\mathcal{Q}_d(\mathbf{g})} = \mathcal{L}_d(\mathbf{g})^\vee$.

An important construction is the restriction of positive linear functionals. For $d \geq k$:

$$\mathcal{L}_d(\mathbf{g})^{[k]} = \{\sigma^{[k]} \mid \sigma \in \mathcal{L}_d(\mathbf{g})\}.$$

As $\mathcal{L}_d(\mathbf{g})$ is defined by Linear Matrix Inequalities, it is a so called *spectrahedron*. Geometrically, $\mathcal{L}_d(\mathbf{g})^{[k]}$ is the projection of $\mathcal{L}_d(\mathbf{g})$ onto the moments of degree $\leq k$, i.e. it is a projected spectrahedron. Then for every d, $\mathcal{L}_d(\mathbf{g})^{[k]}$ is an outer, spectrahedral approximation of the cone $\mathcal{M}(S)^{[k]}$.

The case $k = 1$ is of particular geometric interest. Indeed, identifying moments of degree one and points of $\mathbb{R}^n$, one can see the section of $\mathcal{L}_d(\mathbf{g})$ given by $\sigma_0 = \langle \sigma | 1 \rangle = 1$ as a spectrahedral outer approximation of the convex hull of S, see for instance [Las15, ex. 4.4].

Real algebra. We list here for reader's convenience some results of real algebra that will be frequently used through the paper, and refer to [Mar08] and [BCR98] for more details.

Definition 1.1 (Real Radical)**.** Let $I \subset \mathbb{R}[\mathbf{X}]$ be an ideal. The *real radical* of I, denoted $\sqrt[\mathbb{R}]{I}$ is the ideal:

$$\sqrt[\mathbb{R}]{I} := \{f \in \mathbb{R}[\mathbf{X}] \mid \exists m \in \mathbb{N},\ s \in \Sigma^2 \text{ with } f^{2m} + s \in I\} = (I + \Sigma^2) \cap -(I + \Sigma^2).$$

We say that the ideal I is *real* or *real radical* if $I = \sqrt[\mathbb{R}]{I}$.

Notice that $\sqrt[\mathbb{R}]{I}$ is a radical ideal. We are in particular interested in the case $I = \operatorname{supp} Q = Q \cap -Q$ for an arbitrary quadratic module $Q \subset \mathbb{R}[\mathbf{X}]$.

Lemma 1.2. *Let $Q \subset \mathbb{R}[\mathbf{X}]$ be a quadratic module and $I = \operatorname{supp} Q$. Then:*

(i) I is an ideal;

(ii) $\sqrt{I} = \sqrt[\mathbb{R}]{I}$, i.e. the radical of I is equal to the real radical of I;

(iii) $\dim \frac{\mathbb{R}[\mathbf{X}]}{I} = \dim \frac{\mathbb{R}[\mathbf{X}]}{\sqrt[\mathbb{R}]{I}}$, where $\dim$ denotes the Krull dimension.

Proof. We briefly prove these known results for the sake of completeness and refer to [Mar08] for more details.

For the first point, closure by addition is trivial. Closure by multiplication follows observing that for all $f \in \mathbb{R}[\mathbf{X}]$ we have $f = (\frac{f+1}{2})^2 - (\frac{f-1}{2})^2 \in \Sigma^2 - \Sigma^2$.

For the second point, since $\sqrt[\mathbb{R}]{I}$ is a radical ideal we have $\sqrt{I} \subset \sqrt[\mathbb{R}]{I}$. Recall that $\sqrt{I}$ is the intersection of all the minimal prime ideals $\mathfrak{p}$ lying over I. From [Mar08, prop. 2.1.7] we deduce that for such a minimal prime $\mathfrak{p}$, $(M+\mathfrak{p}) \cap -(M+\mathfrak{p}) = \mathfrak{p}$ and thus $(\Sigma^2+\mathfrak{p}) \cap -(\Sigma^2+\mathfrak{p}) = \mathfrak{p}$, i.e. $\mathfrak{p}$ is real radical, see Definition 1.1. As the intersection of real radical ideals is real radical, we see that $\sqrt{I}$ is a real radical ideal and thus $\sqrt{I} = \sqrt[\mathbb{R}]{I}$.

The last point follows from the second point and the property of the Krull dimension: $\dim \frac{\mathbb{R}[\mathbf{X}]}{I} = \dim \frac{\mathbb{R}[\mathbf{X}]}{\sqrt{I}}$. □

We will then use $\sqrt[\mathbb{R}]{\operatorname{supp} Q}$ to denote both the radical and the real radical of $\operatorname{supp} Q$.

The real radical ideals can be used to describe the polynomials vanishing on a semialgebraic or algebraic set.

Theorem 1.3 (Real Nullstellensatz, [Mar08, note 2.2.2 (vi)]). *Let $S = \mathcal{S}(\mathbf{g})$ be a basic closed semi-algebraic set. Then for $f \in \mathbb{R}[\mathbf{X}]$, $f = 0$ on S if and only if $f \in \sqrt[\mathbb{R}]{\operatorname{supp} \mathcal{O}(\mathbf{g})}$. In other words, $\mathcal{I}(S) = \sqrt[\mathbb{R}]{\operatorname{supp} \mathcal{O}(\mathbf{g})}$.*

In particular, for an ideal $I \subset \mathbb{R}[\mathbf{X}]$ we have $\mathcal{I}(\mathcal{V}_{\mathbb{R}}(I)) = \sqrt[\mathbb{R}]{I}$.

The preordering $\mathcal{O}(\mathbf{g})$ can be replaced with the quadratic module $\mathcal{Q}(\mathbf{g})$ when the Krull dimension of the quotient $\frac{\mathbb{R}[\mathbf{X}]}{\operatorname{supp} \mathcal{Q}(\mathbf{g})}$ is ≤ 1, as shown in the book of Marshall.

Theorem 1.4 ([Mar08, cor. 7.4.2 (3)]). *If $\dim \frac{\mathbb{R}[\mathbf{X}]}{\operatorname{supp} \mathcal{Q}(\mathbf{g})} \leq 1$, then $\mathcal{I}(S) = \sqrt[\mathbb{R}]{\operatorname{supp} \mathcal{Q}(\mathbf{g})}$.*

We will often use Theorem 1.4 in the case $\dim \frac{\mathbb{R}[\mathbf{X}]}{\operatorname{supp} \mathcal{Q}(\mathbf{g})} = 0$.

2 Finite Convergence and Exactness

We describe now Lasserre SoS and MoM relaxations [Las01], and we define the *exactness* property. Hereafter we assume that the minimum f^* of the objective function f is always attained on S, that is: $S^{\min} := \{x \in S \mid f(x) = f^*\} \neq \emptyset$.

2.1 Polynomial optimization relaxations

Lasserre's SoS relaxations. We define the *SoS relaxation of order d* of problem (1) as $\mathcal{Q}_{2d}(\mathbf{g})$ and the supremum:

$$f^*_{\mathrm{SoS},d} := \sup\Big\{\lambda \in \mathbb{R} \mid f - \lambda \in \mathcal{Q}_{2d}(\mathbf{g})\Big\}. \tag{2}$$

When necessary we will replace $\mathbf{g}$ by $\Pi\mathbf{g}$ (that is $\mathcal{Q}(\mathbf{g})$ by $\mathcal{O}(\mathbf{g})$).

Lasserre's MoM ralaxations. To define the dual approximation of the polynomial optimization problem, we consider an affine hyperplane section of the cone $\mathcal{L}_{2d}(\mathbf{g})$:

$$\mathcal{L}^{(1)}_{2d}(\mathbf{g}) := \Big\{\sigma \in \mathcal{L}_{2d}(\mathbf{g}) \mid \langle \sigma | 1 \rangle = 1\Big\}.$$

This will be the set of feasible (pseudo-)moment sequences of the MoM relaxation of order d. Notice that $\mathcal{L}_{2d}(\mathbf{g})$ is the cone over $\mathcal{L}^{(1)}_{2d}(\mathbf{g})$, since for $\sigma \in \mathcal{L}_{2d}(\mathbf{g})$ we have $\langle \sigma | 1 \rangle = 0 \Rightarrow \sigma = 0$ (see [Las+13, lem. 3.12]), and $\sigma \in \mathcal{L}_{2d}(\mathbf{g}) \neq 0$ implies $\frac{1}{\langle \sigma | 1 \rangle}\sigma \in \mathcal{L}^{(1)}_{2d}(\mathbf{g})$.

The convex sets $\mathcal{L}^{(1)}_{2d}(\mathbf{g})$ are spectrahedra: they are defined by the Linear Matrix Inequalities $H_\sigma^{\lfloor \frac{d}{2} \rfloor} \succcurlyeq 0$, $H^{N_1}_{g_1 \star \sigma} \succcurlyeq 0, \ldots, H^{N_r}_{g_r \star \sigma} \succcurlyeq 0$, where $N_i = d - \lceil \frac{\deg g_i}{2} \rceil$. The convex set $\mathcal{L}^{(1)}_d(\mathbf{g})$ is also called the state space of $(\mathbb{R}[\mathbf{X}]_d, \mathcal{Q}_d(\mathbf{g}), 1)$ in [KS19]. The pure states are the extreme points of this convex set.

With this notation we define the *MoM relaxation of order d* of problem (1) as $\mathcal{L}_{2d}(\mathbf{g})$ and the infimum:

$$f^*_{\mathrm{MoM},d} := \inf\Big\{ \langle \sigma | f \rangle \in \mathbb{R} \mid \sigma \in \mathcal{L}^{(1)}_{2d}(\mathbf{g}) \Big\}. \tag{3}$$

We easily verify that $f^*_{\mathrm{SoS},d} \leq f^*_{\mathrm{MoM},d} \leq f^*$. When necessary we will replace $\mathbf{g}$ by $\Pi\mathbf{g}$ (that is $\mathcal{Q}(\mathbf{g})$ by $\mathcal{O}(\mathbf{g})$).

When $S^{\min} := \{\xi \in S \mid f(\xi) = f^*\} \neq \emptyset$, the infimum $f^*_{\mathrm{MoM},d}$ is reached since $\mathcal{L}^{(1)}_{2d}(\mathbf{g})$ is closed and bounded. We are interested, in particular, in the linear functionals that realize this minimum.

Definition 2.1. Consider the problem of minimizing $f \in \mathbb{R}[\mathbf{X}]$ on $\mathcal{S}(\mathbf{g})$. We define the set of *functional minimizers* at relaxation order d as the σ minimizing (3), i.e.:

$$\mathcal{L}^{\min}_{2d}(\mathbf{g}) := \Big\{ \sigma \in \mathcal{L}^{(1)}_{2d}(\mathbf{g}) \mid \langle \sigma | f \rangle = f^*_{\mathrm{MoM},d} \Big\}.$$

We now introduce two convergence properties that will be central in the article.

Definition 2.2 (Finite Convergence)**.** We say that the SoS relaxation $(\mathcal{Q}_{2d}(\mathbf{g}))_{d\in\mathbb{N}}$ (resp. the MoM relaxation $\mathcal{L}_{2d}(\mathbf{g}))_{d\in\mathbb{N}}$ has the *Finite Convergence* property for f if $\exists k \in \mathbb{N}$ such that for every $d \geq k$, $f^*_{\mathrm{SoS},d} = f^*$ (resp. $f^*_{\mathrm{MoM},d} = f^*$).

Notice that if the SoS relaxation has finite convergence then the MoM relaxation has finite convergence too, since $f^*_{\mathrm{SoS},d} \leq f^*_{\mathrm{MoM},d} \leq f^*$. Moreover, if $f^*_{\mathrm{MoM},d} = f^*$ then $\mathcal{L}^{\min}_{2d}(\mathbf{g}) = \{\sigma \in \mathcal{L}^{(1)}_{2d}(\mathbf{g}) \mid \langle \sigma | f \rangle = f^*\}$.

Definition 2.3 (SoS Exactness)**.** We say that the SoS relaxation $(\mathcal{Q}_{2d}(\mathbf{g}))_{d\in\mathbb{N}}$ is *exact* for f if it has the finite convergence property and for all d big enough, we have $f - f^* \in \mathcal{Q}_{2d}(\mathbf{g})$ (in other words sup = max in the definition of $f^*_{\mathrm{SoS},d}$).

For the moment relaxation we can ask the property that every truncated functional minimizer is coming from a measure:

Definition 2.4 (MoM Exactness)**.** We say that the MoM relaxation $(\mathcal{L}_{2d}(\mathbf{g}))_{d\in\mathbb{N}}$ is *exact* for f on the basic closed semialgebraic set S if:

- it has the finite convergence property, and
- for every $k \in \mathbb{N}$, there exists $d = d(k) \in \mathbb{N}$ such that every truncated functional minimizer is coming from a probability measure supported on S, i.e. $\mathcal{L}^{\min}_{2d}(\mathbf{g})^{[k]} \subset \mathcal{M}^{(1)}(S)^{[k]}$.

If not specified, S will be the semialgebraic set $S = \mathcal{S}(\mathbf{g})$ defined by $\mathbf{g}$.

MoM exactness may be considered as a particular instance of the so called *Moment Problem* (i.e. asking if $\sigma \in \mathbb{R}[\mathbf{X}]^*$ is coming from a measure) or of the *Strong Moment Problem* (i.e. asking that the measure has a specified support). More precisely, MoM exactness can be considered as a *Truncated Strong Moment Property* (since we are considering functionals restricted to polynomials up to a certain degree).

Notice that in the definition we require the property $\mathcal{L}^{\min}_{2d(k)}(\mathbf{g})^{[k]} \subset \mathcal{M}^{(1)}(S)^{[k]}$ to hold for every k, and in general the fact that the property is verified for particular k does not imply that it is verified for every k.

We show now an example where we investigate the properties of finite convergence and exactness.

Example 2.5. Consider the problem of minimizing $f = X^2$ on the semialgebraic set $S = \mathcal{S}(\mathbf{g}) = \mathcal{S}(1 - X^2 - Y^2, X + Y - 1) \subset \mathbb{R}^2$ defined by $g_1 = 1 - X^2 - Y^2$ and $g_2 = X + Y - 1$. Clearly, the minimum is $f^* = 0$ and the only minimizer is $(0, 1)$. Notice that $f - f^* = X^2 \in \mathcal{Q}_2(1 - X^2 - Y^2, X + Y - 1)$ and therefore $f^*_{\mathrm{SoS},1} = f^*_{\mathrm{MoM},1} = f^* = 0$, we have finite convergence and the SoS relaxation is exact.

We now investigate MoM exactness. If a truncated moment sequence σ is coming from a probability measure $\mu \in \mathcal{M}^{(1)}(S)$ such that $\int f \, \mathrm{d}\mu = f^*$, then the support of μ should be contained

in the set of minimizers $S^{\min} = \{(0,1)\}$ of f. Thus $\mu = \mathbf{e}_{(0,1)}$ is the evaluation at $(0,1)$ (or in other words, the Dirac measure concentrated at $(0,1)$). Its moments are easily computed: $\mu_{00} = 1$, $\mu_{10} = 0$, $\mu_{01} = 1$, $\mu_{20} = 0, \ldots$.

Analyzing the constraints on the degree one and two moments of an optimal moment sequence $\sigma \in \mathcal{L}_2^{\min}(\mathbf{g})$, where

$$\begin{aligned}\mathcal{L}_2^{\min}(\mathbf{g}) &= \{\sigma \in \mathbb{R}[\mathbf{X}]_2^* \mid H_\sigma^1 \succcurlyeq 0,\ H_{g_1\star\sigma}^0 \succcurlyeq 0,\ H_{g_2\star\sigma}^0 \succcurlyeq 0, \langle\sigma|1\rangle = 1, \langle\sigma|f\rangle = f^* = 0\} \\ &= \{\sigma \in \mathbb{R}[\mathbf{X}]_2^* \mid \begin{pmatrix}\sigma_{00} & \sigma_{10} & \sigma_{01} \\ \sigma_{10} & \sigma_{20} & \sigma_{11} \\ \sigma_{01} & \sigma_{11} & \sigma_{02}\end{pmatrix} \succcurlyeq 0,\ \sigma_{00} - \sigma_{20} - \sigma_{02} \geq 0,\ \sigma_{10} + \sigma_{01} - \sigma_{00} \geq 0, \sigma_{00} = 1, \sigma_{20} = 0\},\end{aligned}$$

we deduce that $\sigma_{00} = 1$, $\sigma_{10} = 0$, $\sigma_{01} = 1$, $\sigma_{20} = 0$, $\sigma_{11} = 0$ and $\sigma_{02} = 1$: this shows that the only element of $\mathcal{L}_2^{\min}(\mathbf{g})$ is $\sigma = \mathbf{e}_{(0,1)}^{[2]}$. In particular notice that $\left\langle\sigma\middle|X^2\right\rangle = \left\langle\sigma\middle|(Y-1)^2\right\rangle = 0$.

For any order $d \geq 1$ and any element $\sigma \in \mathcal{L}_{2d}^{\min}(\mathbf{g})$, its truncation $\sigma^{[2]}$ is in $\mathcal{L}_2^{\min}(\mathbf{g})$ since $\left\langle\sigma^{[2]}\middle|X^2\right\rangle = \left\langle\sigma\middle|X^2\right\rangle = 0$. This implies that $\left\langle\sigma\middle|X^2\right\rangle = \left\langle\sigma\middle|(Y-1)^2\right\rangle = 0$ and that $\forall p \in \mathbb{R}[\mathbf{X}]_d, \langle\sigma|Xp\rangle = \langle\sigma|(Y-1)p\rangle = 0$. We deduce from Proposition 3.16 that the moments of $\sigma^{[d]} = \mathbf{e}_{(0,1)}^{[d]}$ are coming from the Dirac measure $\mathbf{e}_{(0,1)}$. Therefore the relaxation is MoM exact.

Another equivalent way to certify MoM exactness is to check flat truncation (see Definition 3.18). For $\sigma \in \mathcal{L}_{2d}^{\min}(\mathbf{g})$ with $d \geq 2$, we have computed the moments of degree ≤ 2. Since the moment matrices in degree ≤ 2:

$$H_\sigma^0 = \begin{pmatrix}1\end{pmatrix}, \quad H_\sigma^1 = \begin{pmatrix}1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1\end{pmatrix}$$

have the same rank, the flat extension property is satisfied. This certifies that $\sigma^{[2]} = \mathbf{e}_{(0,1)}^{[2]}$ is coming from a measure supported at the minimizer of f on S and the MoM relaxation is exact, see Theorem 4.1.

In practice, to check the finite convergence, one tests the flat extension or the flat truncation property of moment matrices (see [CF98], [LM09], [Nie13b]). But flat truncation certifies MoM exactness, and not only finite convergence. We will investigate flat truncation for POPs in Section 4.

Notice that in the previous example the rank condition is satisfied by the full sequence of moments of $\sigma \in \mathcal{L}_2^{\min}(\mathbf{g})$ (flat extension). In general this is not true, as the high degree moments may be increasing the rank of the moment matrix, see for in instance [Nie13b, ex. 1.1]. Therefore it is necessary to discard the high degree moments, i.e. to consider $\mathcal{L}_{2d}^{\min}(\mathbf{g})^{[t]}$, for some $t \leq 2d$, instead of simply $\mathcal{L}_{2d}^{\min}(\mathbf{g})$. This implies that we look for rank conditions on the moment matrix of the truncated moment sequence (i.e. flat truncation).

We recall results of strong duality, i.e. cases when we know that $f_{\mathrm{SoS},d}^* = f_{\mathrm{MoM},d}^*$, that we will be using. See also Proposition 3.9.

Theorem 2.6 (Strong duality). *Let $Q = \mathcal{Q}(\mathbf{g})$ be a quadratic module and f the objective function. Then:*

(i) if $\operatorname{supp} Q = 0$ then $\forall d$: $f_{\mathrm{SoS},d}^*$ is attained (i.e. $f - f_{\mathrm{SoS},d}^* \in \mathcal{Q}_d(\mathbf{g})$) and $f_{\mathrm{SoS},d}^* = f_{\mathrm{MoM},d}^*$ [Mar08, prop. 10.5.1];

(ii) if $r^2 - \|\mathbf{X}\|^2 \in \mathbf{g}$ then $f_{\mathrm{SoS},d}^* = f_{\mathrm{MoM},d}^*$ for all d [JH16].

Remark. [JH16] applies when the ball constraint $r^2 - \|\mathbf{X}\|^2$ appears explicitely in the description of S. But if we consider a problem with MoM finite convergence and such that $\mathcal{Q}(\mathbf{g})$ is Archimedean, then we can use [JH16] to prove that we have also SoS finite convergence. Indeed, if $\mathcal{Q}(\mathbf{g})$ is Archimedean there exists r, d such that $r^2 - \|\mathbf{X}\|^2 \in \mathcal{Q}_{2t}(\mathbf{g})$. This means that $\mathcal{Q}_{2d}(\mathbf{g}, r^2 - \|\mathbf{X}\|^2) \subset \mathcal{Q}_{2d+2t}(\mathbf{g})$. If we define:

- $f_{\mathrm{SoS},d}^* = \sup\left\{\lambda \in \mathbb{R} \mid f - \lambda \in \mathcal{Q}_{2d}(\mathbf{g})\right\}$

- $f^{*'}_{\mathrm{SoS},d} = \sup\left\{\lambda \in \mathbb{R} \mid f - \lambda \in \mathcal{Q}_{2d}(\mathbf{g}, r^2 - \|X\|^2)\right\}$

and $f^*_{\mathrm{MoM},d}$, $f^{*'}_{\mathrm{MoM},d}$ the corresponding MoM relaxations, then:

$$f^*_{\mathrm{MoM},d} \leq f^{*'}_{\mathrm{MoM},d} = f^{*'}_{\mathrm{SoS},d} \leq f^*_{\mathrm{SoS},d+t} \leq f^*.$$

Then finite convergence of the MoM relaxation implies finite convergence of the SoS one.

We recall that we are assuming $S^{\min} \neq \emptyset$ (in particular f^* is finite: otherwise it may happen that $f^*_{\mathrm{SoS},d} = -\infty$). Notice that if strong duality holds, then SoS finite convergence is equivalent to MoM finite convergence.

2.2 Examples and counterexamples

In this section, we give examples showing how the notions of finite convergence and exactness of the SoS and MoM relaxations are (and are not) related.

No finite convergence. The first example shows that SoS and MoM relaxations for polynomial optimization on algebraic curves do not have necessarily the finite convergence property.

Example 2.7 ([Sch00])**.** Let $\mathcal{C} \subset \mathbb{R}^n$ be a smooth connected curve of genus ≥ 1, with only real points at infinity. Let $\mathbf{h} = \{h_1, \ldots, h_s\} \subset \mathbb{R}[\mathbf{X}]$ be a graded basis of $I = \mathcal{I}(\mathcal{C}) = (\mathbf{h})$. Then there exists $f \in \mathbb{R}[\mathbf{X}]$ such that the SoS relaxation $\mathcal{Q}_{2d}(\pm\mathbf{h})$ and the MoM relaxation $\mathcal{L}_{2d}(\pm\mathbf{h})$ have no finite convergence and are not exact.

Indeed by [Sch00, Theorem 3.2], there exists $f \in \mathbb{R}[\mathbf{X}]$ such that $f \geq 0$ on $\mathcal{C} = \mathcal{S}(\pm\mathbf{h})$, which is not a sum of squares in $\mathbb{R}[\mathcal{C}] = \mathbb{R}[\mathbf{X}]/I$. Consequently, $f \notin \Sigma^2[\mathbf{X}] + I = \mathcal{Q}(\pm\mathbf{h})$. As $f \geq 0$ on $\mathcal{C}$, its infimum f^* is non-negative and we also have $f - f^* \notin \mathcal{Q}(\pm\mathbf{h})$.

Using Proposition 3.9 we deduce that $\mathcal{Q}_{2d}(\pm\mathbf{h})$ is closed, that there is no duality gap and that the supremum $f^*_{\mathrm{SoS},d}$ is reached. Thus if the SoS relaxation has finite convergence then $f - f^* \in \mathcal{Q}_{2d}(\pm\mathbf{h})$ for some $d \in \mathbb{N}$. This is a contradiction, showing that the SoS and the MoM relaxations have no finite convergence and cannot be SoS exact for f.

In dimension 2, there are also cases where the SoS and MoM relaxations cannot have finite convergence or be exact.

Example 2.8 ([Mar08])**.** Let $g_1 = X_1^3 - X_2^2, g_2 = 1 - X_1$. Then $S = \mathcal{S}(\mathbf{g})$ is a compact semialgebraic set of dimension 2 and $\mathcal{O}(\mathbf{g})$ is Archimedean. We have $f = X_1 \geq 0$ on S but $X_1 \notin \mathcal{O}(\mathbf{g})$ (see [Mar08, Example 9.4.6(3)]). The infimum of f on S is $f^* = 0$. Assume that we have MoM finite convergence. By Theorem 2.6 and remark below, $\mathcal{Q}_{2d}(\Pi\mathbf{g})$ is closed, the supremum $f^*_{\mathrm{SoS},d}$ is reached and strong duality holds: $f^*_{\mathrm{SoS},d} = f^*_{\mathrm{MoM},d} = f^* = 0$ for $d \in \mathbb{N}$ big enough. Then $f - f^* = f \in \mathcal{O}(\mathbf{g})$: but this is a contradiction. Therefore, the relaxations $\mathcal{Q}_{2d}(\Pi\mathbf{g})$ and $\mathcal{L}_{2d}(\Pi\mathbf{g})$ cannot have finite convergence and thus cannot be exact for $f = X_1$.

The next example shows that non-finite convergence and non-exactnesss is always possible in dimension ≥ 3.

Example 2.9. Let $n \geq 3$. Let Q be an Archimedean quadratic module generated by $g_1, \ldots, g_s \in \mathbb{R}[\mathbf{X}]$ such that $\mathcal{S}(\mathbf{g}) \subset \mathbb{R}^n$ is of dimension $m \geq 3$ and let $\mathbf{h}$ be a graded basis of $\sqrt[\mathbb{R}]{\operatorname{supp} Q}$ (in particular $\mathbf{h} = 0$ if $\operatorname{supp} Q = 0$ or if $m = n$, i.e. $\mathcal{S}(Q)$ is of maximal dimension), then there exists $f \in \mathbb{R}[\mathbf{X}]$ such that the SoS relaxation $(\mathcal{Q}_{2d}(\mathbf{g}, \pm\mathbf{h}))_{d\in\mathbb{N}}$ and MoM relaxation $(\mathcal{L}_{2d}(\mathbf{g}, \pm\mathbf{h}))_{d\in\mathbb{N}}$ do not have the finite convergence property (and thus are not exact).

Indeed by Proposition 3.9 $f^*_{\mathrm{SoS},d} = f^*_{\mathrm{MoM},d}$ for d big enough and the supremum $f^*_{\mathrm{SoS},d}$ is reached. By [Sch00, Prop. 6.1] for $m \geq 3$, $\operatorname{Pos}(\mathcal{S}(\mathbf{g})) = \operatorname{Pos}(\mathcal{S}(Q + (\mathbf{h}))) \supsetneq Q + (\mathbf{h})$. So let $f \in \operatorname{Pos}(\mathcal{S}(Q)) \setminus Q + (\mathbf{h})$ and let f^* be its minimum on $\mathcal{S}(Q)$. Suppose that $f - f^* \in Q + (\mathbf{h})$, then $f \in Q + (\mathbf{h}) + f^* = Q + (\mathbf{h})$, a contradiction. Then the SoS and the MoM relaxations do not have the finite convergence property (and they are not exact).

SoS exactness, no MoM exactness.

Example 2.10. We want to find the global minimum of $f = X_1^2 \in \mathbb{R}[X_1,\dots,X_n] = \mathbb{R}[\mathbf{X}]$ for $n \geq 3$. Let $d \geq 2$, $\mathbf{X}' = (X_2,\dots,X_n)$ and $\overline{\sigma} \in \mathcal{L}_{2d}(\Sigma^2[\mathbf{X}'])$ such that $\overline{\sigma} \notin \mathcal{M}(\mathbb{R}^{n-1})^{[d]}$. Such a linear functional exists because when $n > 2$ there are non-negative polynomials in $\mathbb{R}[\mathbf{X}']$ which are not sum of squares, such as the Motzkin polynomial (see [Rez96]). As $\mathcal{Q}_{2d}(\Sigma^2[\mathbf{X}'])$ is closed, such a polynomial can be separated from $\mathcal{Q}_{2d}(\Sigma^2[\mathbf{X}'])$ by a linear functional $\overline{\sigma} \in \mathcal{L}_{2d}(\Sigma^2[\mathbf{X}'])$, which cannot be the truncation of a measure (i.e. $\Sigma^2[\mathbf{X}']$ does not have the truncated moment property). Define $\sigma : h \mapsto \langle \sigma | h \rangle = \langle \overline{\sigma} | h(0, X_2,\dots,X_n) \rangle$. We have $\sigma \in \mathcal{L}_{2d}(\Sigma^2[\mathbf{X}]) = \Sigma^2[\mathbf{X}]^{\vee}_{2d}$ since $\overline{\sigma} \in \mathcal{L}_{2d}(\Sigma^2[\mathbf{X}']) = \Sigma^2[\mathbf{X}']^{\vee}_{2d}$. Obviously $\langle \sigma | f \rangle = 0 = f^*$ (the minimum of X_1^2), $f - f^* = X_1^2 \in \Sigma^2$ and the SoS relaxation is exact. Since σ is coming from a measure if and only if $\overline{\sigma}$ is coming from a measure, the MoM relaxation cannot be exact.

The previous example generalizes easily to quadratic modules Q with $\operatorname{supp}(Q) \neq \{0\}$, which do not have the (truncated) moment property, i.e. there exists $\sigma \in \mathcal{L}_{2d}(Q)$ such that $\sigma \notin \mathcal{M}(\mathcal{S}(Q))^{[2d]}$. Taking $f = h^2$ with $h \in \operatorname{supp}(Q)$, $h \neq 0$, we have $\langle \sigma | f \rangle = 0 = f^*$ and the MoM relaxation cannot be exact since $\sigma \notin \mathcal{M}(\mathcal{S}(Q))^{[2d]}$, while the SoS relaxation is exact ($f - f^* = h^2 \in Q$).

Example 2.10 is an example where the number of minimizers of f on S is infinite. We show that non exactness can happen also when the minimizers are finite (and even when S is finite!).

Example 2.11 ([Sch05a, ex. 3.2], [Sch05b, rem. 3.15], Example 3.5)**.** We want to minimize the constant function $f = 1$ on the origin $S = \mathcal{S}(Q) = \{\mathbf{0}\}$, where $Q = \mathcal{Q}(1 - X^2 - Y^2, -XY, X - Y, Y - X^2) \subset \mathbb{R}[X,Y]$. In this case $\operatorname{supp} Q = \sqrt[\mathbb{R}]{\operatorname{supp} Q} = (0)$. Notice that the SoS relaxation is exact and the MoM relaxation has finite convergence, since f is a square. Now suppose that the MoM relaxation is exact, i.e. $\mathcal{L}_{2d}^{\min}(\mathbf{g})^{[2k]} = \mathcal{L}_{2d}^{(1)}(\mathbf{g})^{[2k]} \subset \mathcal{M}^{(1)}(S)^{[2k]} = \{\mathbf{e}_{\mathbf{0}}^{[2k]}\}$ for some d, k. Then for $\sigma^* \in \mathcal{L}_{2d}(\mathbf{g})$ generic we have $(\operatorname{Ann}_k(\sigma^*)) = (\operatorname{Ann}_k(\mathbf{e}_{\mathbf{0}})) = (X, Y)$. But from Theorem 3.12 we know that for d, k big enough $(\operatorname{Ann}_k(\sigma^*)) = \sqrt[\mathbb{R}]{\operatorname{supp} Q} = (0)$, a contradiction. Then the MoM relaxation is not exact. Moreover the flat truncation property is not satisfied in this case: see Theorem 4.4.

We investigate concretely this example for $d = 1$. We show in Figure 1[2] the plot of $\mathcal{L}_2(\mathbf{g})^{[1]}$, that is the moments of degree one of the the linear functionals in dual cone of $\mathcal{Q}_2(\mathbf{g})$. Notice that this is an outer approximation of $\mathbf{e}_{(0,0)} \in \mathcal{L}_2(\mathbf{g})^{[1]}$ or, identifying moments of degree one with points of $\mathbb{R}^n$, a convex outer approximation of $S = \{(0,0)\}$.

One can also verify explicitly that $\mathcal{L}_2(\mathbf{g})$ has nonempty interior, since $\sigma = \sigma(\varepsilon)$ defined by $\sigma_{10} = 2\varepsilon$, $\sigma_{01} = \varepsilon$, $\sigma_{20} = \frac{\varepsilon}{2}$, $\sigma_{11} = -\varepsilon^2$ and $\sigma_{02} = \frac{1}{2}$ lies in the interior of $\mathcal{L}_2(\mathbf{g})$ for $\varepsilon > 0$ small enough.

Notice that $\mathcal{L}_2(\mathbf{g})^{[1]} \supset \mathcal{L}_3(\mathbf{g})^{[1]} \supset \mathcal{L}_4(\mathbf{g})^{[1]} \supset \cdots \supset \{\mathbf{e}_{(0,0)}^{[1]}\}$, and we have convergence in this case since $\mathcal{Q}(\mathbf{g})$ is Archimedean. This nested outer approximations, shown in Figure 1[2], never coincide with $\{\mathbf{e}_{(0,0)}^{[1]}\}$, as we have proven before.

SoS finite convergence, MoM exactness.

Example 2.12. Let $f = (X^4Y^2 + X^2Y^4 + Z^6 - 2X^2Y^2Z^2) + X^8 + Y^8 + Z^8 \in \mathbb{R}[X,Y,Z]$. We want to optimize f over the gradient variety $\mathcal{V}_{\mathbb{R}}\left(\frac{\partial f}{\partial X}, \frac{\partial f}{\partial Y}, \frac{\partial f}{\partial Z}\right)$ which is zero dimensional (see [NDS06]). By Theorem 4.10 the flat truncation is satisfied and the MoM relaxation is exact, and by Theorem 2.6 and remark below the SoS has the finite convergence property (notice that $\mathcal{Q}(\pm\frac{\partial f}{\partial X}, \pm\frac{\partial f}{\partial Y}, \pm\frac{\partial f}{\partial Z}) = \mathcal{O}(\pm\frac{\partial f}{\partial X}, \pm\frac{\partial f}{\partial Y}, \pm\frac{\partial f}{\partial Z})$ is Archimedean since $\mathcal{V}_{\mathbb{R}}\left(\frac{\partial f}{\partial X}, \frac{\partial f}{\partial Y}, \frac{\partial f}{\partial Z}\right)$ is compact). But the SoS relaxation is not exact, as shown in [NDS06].

Example 2.13. Let $f = X_1$. We want to find its value at the origin, defined by $\|\mathbf{X}\|^2 = 0$. As proved in [Nie13c] there is finite convergence but not exactness for the SoS relaxation. On the other hand by Theorem 4.10 the flat truncation property is satisfied and the MoM relaxation is exact.

[2] the variables X, Y in the plots, done using `SDPA`, have been scaled by 100 to reduce floating points errors

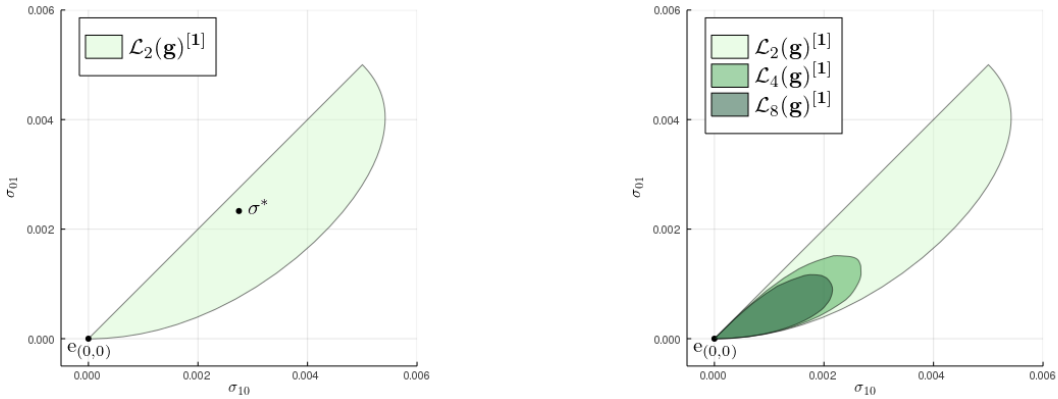

Figure 1: A generic point $\sigma^* \in \mathcal{L}_2^{(1)}(\mathbf{g})^{[1]}$ and moment outer approximations of $\mathcal{L}^{(1)}(\mathbf{g})^{[1]} = \{\mathbf{e}_{0,0}^{[1]}\}$.

Table 1: Summary of convergence results.

Expl.	SoS f. c.	SoS ex.	MoM f. c.	MoM ex.	m
2.7	NO	NO	NO	NO	1
2.8	NO	NO	NO	NO	2
2.9	NO	NO	NO	NO	≥ 3
2.10	YES	YES	YES	NO	≥ 3
2.11	YES	YES	YES	NO	0
2.12	YES	NO	YES	YES	0
2.13	YES	NO	YES	YES	0

We summarize the previous examples in Table 1 in terms of the properties of finite convergence (SoS f.c. and MoM f.c.) exactness (SoS ex. and MoM ex.) and the dimension m of the semialgebraic set S.

3 Geometry of Moment Representations

By Haviland's theorem (see [Mar08, th. 3.1.2] and [Sch17, th. 1.12]) an infinite moment sequence or a linear functional $\sigma \in (\mathbb{R}[\mathbf{X}])^*$ comes from a measure, if and only if σ is positive on positive polynomials. Since checking this is a computationally hard task, a motivation supporting Sum of Squares relaxations is to find (proper) subsets of positive polynomials that can have the same property, chosen in such a way that checking this condition is easy. Important results in this direction are theorems of Schmüdgen and Putinar.

Theorem 3.1 ([Sch91],[Put93]). *Let Q be an Archimedean finitely generated quadratic module and $S = \mathcal{S}(Q)$. Then $\mathcal{L}(Q) = \mathcal{M}(S) = \overline{\mathrm{cone}(\mathbf{e}_\xi : \xi \in S)}$ and $\overline{Q} = \mathrm{Pos}(S)$, where the closures are taken with respect tto the locally convex topology in $\mathbb{R}[\mathbf{X}]^*$ and $\mathbb{R}[\mathbf{X}]$.*

However, describing effectively complete moment sequences in $\mathcal{L}(Q)$ remains a challenging problem since $\mathcal{L}(Q)$ is an infinite dimensional cone. Therefore, it is natural to consider truncated moment problems in order to work in finite dimensional vector spaces. But the cones of truncated moment sequences do not share the properties of the truncated cones of complete moment sequences and special cares are needed.

In this section, we analyze in detail the properties of these finite dimensional *truncated* cones of moment sequences. We provide a new and explicitly description of the dual of the hierarchy of truncated moment sequences, in terms of a quadratic module (Theorem 3.4), and consequently prove properties of the cones $\mathcal{L}_d(\mathbf{g})$ (Lemma 3.6) and of their generic elements (Theorem 3.12).

Finally we apply these results to the zero dimensional case (Theorem 3.21) and we investigate the connections with the flat truncation property (Section 3.3).

3.1 Truncated moment representations

For a finitely generated quadratic module $Q = \mathcal{Q}(\mathbf{g}) \subset \mathbb{R}[\mathbf{X}]$, we have $\mathcal{L}_k(\mathbf{g}) = \mathcal{Q}_k(\mathbf{g})^\vee = \overline{\mathcal{Q}_k(\mathbf{g})}^\vee$ and $\mathcal{L}_k(\mathbf{g})^\vee = \overline{\mathcal{Q}_k(\mathbf{g})}$, where $^\vee$ denotes the dual cone and the closure is taken w.r.t. the euclidean topology on $\mathbb{R}[\mathbf{X}]_k$. Thus the following definition is natural for the study of the MoM relaxations.

Definition 3.2. Let $Q = \mathcal{Q}(\mathbf{g})$ be a finitely generated quadratic module. We define $\widetilde{Q} = \bigcup_d \overline{\mathcal{Q}_d(\mathbf{g})}$.

Notice that $\widetilde{Q}$ depends a priori on the generators $\mathbf{g}$ of Q: we will prove that $\widetilde{Q}$ is a finitely generated quadratic module and that it does not depend on the particular choice of generators. Moreover notice that $Q \subset \widetilde{Q} = \bigcup_d \overline{\mathcal{Q}_d(\mathbf{g})} \subset \overline{\bigcup_d \mathcal{Q}_k(\mathbf{g})} = \overline{Q}$, but these inclusions can be strict.

Lemma 3.3. *Let $Q = \mathcal{Q}(\mathbf{g})$ and $J = \sqrt[\mathbb{R}]{\operatorname{supp} Q}$. Then for every $d \in \mathbb{N}$ there exists $k \geq d$ such that $J_d \subset \overline{\mathcal{Q}_k(\mathbf{g})}$.*

Proof. We denote $\mathcal{Q}_d(\mathbf{g}) = Q_{[d]}$. Let m be big enough such that $\forall f \in J = \sqrt[\mathbb{R}]{\operatorname{supp} Q} = \sqrt{\operatorname{supp} Q}$ we have: $f^{2^m} \in \operatorname{supp} Q$ (if $\sqrt{J} = (h_1, \dots, h_t)$ and $h_i^{a_i} \in I$, we can take m such that $2^m \geq a_1 + \cdots + a_t$). Let $f \in J_d$ with $\deg f \leq d$. Then $f^{2^m} \in \operatorname{supp} \mathcal{Q}_{[k']} \subset Q_{[k']}$ for $k' \in \mathbb{N}$ big enough. Using the identity [Sch05b, remark 2.2]:

$$m - a = (1 - \frac{a}{2})^2 + (1 - \frac{a^2}{8})^2 + (1 - \frac{a^4}{128})^2 + \cdots + (1 - \frac{a^{2^{m-1}}}{2^{2^m - 1}})^2 - \frac{a^{2^m}}{2^{2^{m+1}-2}},$$

substituting a by $-\frac{mf}{\varepsilon}$ and multiplying by $\frac{\varepsilon}{m}$, we have that $\forall \varepsilon > 0$, $f + \varepsilon \in Q_{[k]}$ for $k = \max\{k', 2^m d\}$ (the degree of the representation of $f + \varepsilon$ does not depend on ε). This implies that $f \in \overline{\mathcal{Q}_{[k]}}$. □

We can now describe $\widetilde{Q}$.

Theorem 3.4. *Let $Q = \mathcal{Q}(\mathbf{g})$ be a finitely generated quadratic module and let $J = \sqrt[\mathbb{R}]{\operatorname{supp} Q}$. Then $\widetilde{Q} = Q + J$ and $\operatorname{supp} \widetilde{Q} = J$. In particular, $\widetilde{Q}$ is a finitely generated quadratic module and does not depend on the particular choice of generators of Q.*

Proof. We denote $\mathcal{Q}_d(\mathbf{g}) = Q_{[d]}$. By [Mar08, lemma 4.1.4] $Q_{[d]} + J_d$ is closed in $\mathbb{R}[\mathbf{X}]_d$, thus $\overline{Q_{[d]}} \subset Q_{[d]} + J_d$. Taking unions we prove that $\widetilde{Q} \subset Q + J$.

Conversely by Lemma 3.3 for $d \in \mathbb{N}$ and $k \geq d \in \mathbb{N}$ big enough, $J_d \subset \overline{Q_{[k]}}$. Then, we have $Q_{[d]} + J_d \subset Q_{[k]} + \overline{Q_{[k]}} \subset \overline{Q_{[k]}} + \overline{Q_{[k]}} \subset \overline{Q_{[k]}}$. Taking unions on both sides gives $Q + J \subset \widetilde{Q}$.

Finally $\operatorname{supp} \widetilde{Q} = \operatorname{supp}(Q + J) = J$ by [Sch05b, lemma 3.16]. □

Remark. We proved that $\widetilde{Q} = Q + \sqrt[\mathbb{R}]{\operatorname{supp} Q}$. We also have $\operatorname{supp} \widetilde{Q} = \sqrt[\mathbb{R}]{\operatorname{supp} Q}$ so that if $\operatorname{supp} Q$ is not real radical then $Q \subsetneq \widetilde{Q}$. Example 2.13 is such a case where $\operatorname{supp} Q \neq \sqrt[\mathbb{R}]{\operatorname{supp} Q}$. We notice that, by Theorem 3.4 and [Sch05b, th. 3.17], if Q is *stable*[3] then $\widetilde{Q} = \overline{Q}$. But the inclusion $\widetilde{Q} = Q + \sqrt{\operatorname{supp} Q} \subset \overline{Q}$ can be strict, as shown by the following example.

Example 3.5 ([Sch05a, ex. 3.2], [Sch05b, rem. 3.15], Example 2.11)**.** Let $Q = \mathcal{Q}(1 - X^2 - Y^2, -XY, X - Y, Y - X^2) \subset \mathbb{R}[X, Y]$. Notice that $S = \mathcal{S}(Q) = \{\mathbf{0}\}$ and that Q is Archimedean. Therefore by Theorem 3.1, $\overline{Q} = \operatorname{Pos}(\{\mathbf{0}\})$. We verify that $\operatorname{supp} Q = (0)$ and that $\mathcal{I}(S) = \operatorname{supp} \overline{Q} = (X, Y)$. Thus we have $Q + \sqrt{\operatorname{supp} Q} = \widetilde{Q} \subsetneq \overline{Q}$.

Theorem 3.4 suggests the idea that, when we consider the MoM relaxation, we are extending the quadratic module $\mathcal{Q}(\mathbf{g})$ to $\mathcal{Q}(\mathbf{g}, \pm\mathbf{h})$, where $\mathbf{h}$ are generators of $\sqrt[\mathbb{R}]{\operatorname{supp} \mathcal{Q}(\mathbf{g})}$. We specify this idea in Lemma 3.6, Proposition 3.9 and Theorem 3.12, investigating the relations between the truncated parts of $\mathcal{L}_d(\mathbf{g})$.

[3] $\mathcal{Q}(\mathbf{g})$ is stable if $\forall d \in \mathbb{N}$ there exists $k \in \mathbb{N}$ such that $\mathcal{Q}(\mathbf{g}) \cap \mathbb{R}[\mathbf{X}]_d = \mathcal{Q}_k(\mathbf{g}) \cap \mathbb{R}[\mathbf{X}]_d$.

Lemma 3.6. *Let $J = \sqrt[\mathbb{R}]{\operatorname{supp}\mathcal{Q}(\mathbf{g})}$. If $(\mathbf{h}) \subset J$, $\deg \mathbf{h} \leq t$, then $\exists d \geq t$ such that $\langle \mathbf{h} \rangle_t \subset \overline{\mathcal{Q}_d(\mathbf{g})}$. In this case:*

$$\mathcal{L}_d(\mathbf{g})^{[t]} \subset \mathcal{L}_t(\mathbf{g}, \pm\mathbf{h}) \subset \mathcal{L}_t(\mathbf{g}),$$

and in particular $\mathcal{L}_d(\mathbf{g})^{[t]} \subset \mathcal{L}_t(\pm\mathbf{h})$. Moreover, $\mathcal{L}_{d+2k}(\mathbf{g})^{[t+k]} \subset \mathcal{L}_{t+k}(\pm\mathbf{h})$ for all $k \in \mathbb{N}$.

Proof. By Lemma 3.3, $\langle \mathbf{h} \rangle_t \subset (\mathbf{h})_t \subset \overline{\mathcal{Q}_d(\mathbf{g})}$ for some $d \geq t$. Let $h \in \mathbf{h}$ and $f \in \mathbb{R}[\mathbf{X}]_{t-\deg h}$. Then $\pm f h \in \overline{\mathcal{Q}_d(\mathbf{g})}$, and for $\sigma \in \mathcal{L}_d(\mathbf{g})$, we have $\left\langle \sigma^{[t]} \middle| f h \right\rangle = \langle \sigma | f h \rangle = 0$, i.e. $\mathcal{L}_d(\mathbf{g})^{[t]} \subset \mathcal{L}_t(\mathbf{g}, \pm\mathbf{h})$. The other inclusion $\mathcal{L}_t(\mathbf{g}, \pm\mathbf{h}) \subset \mathcal{L}_t(\mathbf{g})$ follows by definition.

For the second part, notice that $\langle \mathbf{h} \rangle_{t+k} \subset \overline{\mathcal{Q}_{d+2k}(\mathbf{g})}$. Indeed, if $p \in \langle \mathbf{h} \rangle_{t+k}$ then $p = \sum_i \mathbf{X}^{\alpha(i)} p_i$, where $p_i \in \langle \mathbf{h} \rangle_t \subset \overline{\mathcal{Q}_d(\mathbf{g})}$ and $|\alpha(i)| \leq k$. Writing $\mathbf{X}^{\alpha(i)} = (\frac{\mathbf{X}^{\alpha(i)}+1}{2})^2 - (\frac{\mathbf{X}^{\alpha(i)}-1}{2})^2$, we deduce that $p = \sum_i (\frac{\mathbf{X}^{\alpha(i)}+1}{2})^2 p_i + (\frac{\mathbf{X}^{\alpha(i)}-1}{2})^2 (-p_i) \in \overline{\mathcal{Q}_{d+2k}(\mathbf{g})}$, i.e. $\langle \mathbf{h} \rangle_{t+k} \subset \overline{\mathcal{Q}_{d+2k}(\mathbf{g})}$. Then we can conclude the proof as in the first part. □

Remark. Lemma 3.6 says that the MoM relaxation $(\mathcal{L}_{2d}(\mathbf{g}))_{d \in \mathbb{N}}$ is equivalent to the MoM relaxation $(\mathcal{L}_{2d}(\mathbf{g}, \pm\mathbf{h}))_{d \in \mathbb{N}}$, where $(\mathbf{h}) = \sqrt[\mathbb{R}]{\operatorname{supp}\mathcal{Q}(\mathbf{g})}$. Lemma 3.6 is an algebraic result, in the sense that $\operatorname{supp}\mathcal{Q}(\mathbf{g})$ may be unrelated to the geometry $\mathcal{S}(\mathbf{g})$ that $\mathbf{g}$ defines. If some additional conditions hold (namely if we have only equalities, or a preordering, or a small dimension), it can however provide geometric characterizations.

Corollary 3.7. *Suppose that $\mathcal{S}(\mathbf{g}) \subset \mathcal{V}_{\mathbb{R}}(\mathbf{h})$. Then for every $t_0 \geq \deg \mathbf{h}$ there exists $t_1 \geq t_0$ such that:*

$$\mathcal{L}_{t_1}(\Pi\mathbf{g})^{[t_0]} \subset \mathcal{L}_{t_0}(\pm\mathbf{h}).$$

In particular this holds when $(\mathbf{h}) = \mathcal{I}(\mathcal{S}(\mathbf{g}))$.

Moreover, $\mathcal{L}_{t_1+2k}(\mathbf{g})^{[t_0+k]} \subset \mathcal{L}_{t_0+k}(\pm\mathbf{h})$ for all $k \in \mathbb{N}$.

Proof. $\mathcal{S}(\mathbf{g}) \subset \mathcal{V}_{\mathbb{R}}(\mathbf{h})$ if and only if $\sqrt[\mathbb{R}]{(\mathbf{h})} \subset \mathcal{I}(\mathcal{S}(\mathbf{g})) = \sqrt[\mathbb{R}]{\operatorname{supp}\mathcal{Q}(\Pi\mathbf{g})}$ by the Real Nullstellensatz, Theorem 1.3. Then we can apply Lemma 3.6. □

Corollary 3.8. *Let $Q = \mathcal{Q}(\mathbf{g})$. Suppose that $\mathcal{S}(\mathbf{g}) \subset \mathcal{V}_{\mathbb{R}}(\mathbf{h})$ and $\dim \frac{\mathbb{R}[\mathbf{X}]}{\operatorname{supp} Q} \leq 1$. Then for every $t_0 \geq \deg \mathbf{h}$ there exists $t_1 \geq t_0$ such that $(\mathbf{h})_{t_0} \subset \overline{\mathcal{Q}_{t_1}(\mathbf{g})}$. In this case:*

$$\mathcal{L}_{t_1}(\mathbf{g})^{[t_0]} \subset \mathcal{L}_{t_0}(\pm\mathbf{h}),$$

ans in particular this holds when $(\mathbf{h}) = \mathcal{I}(\mathcal{S}(\mathbf{g}))$.

Moreover, $\mathcal{L}_{t_1+2k}(\mathbf{g})^{[t_0+k]} \subset \mathcal{L}_{t_0+k}(\pm\mathbf{h})$ for all $k \in \mathbb{N}$.

Proof. We prove it as Corollary 3.7, using Theorem 1.4 instead of the Real Nullstellensatz. □

We mention now a strong duality result, that is useful to produce examples and counterexamples for exactness and finite convergence, This result, very similar to a result in [Mar03], generalizes the condition $\operatorname{supp} Q = 0$ in Theorem 2.6. We conjecture that there is no duality gap when Q is reduced (i.e. $\operatorname{supp} Q = \sqrt[\mathbb{R}]{\operatorname{supp} Q}$) without adding the generators of the radical of the support.

Proposition 3.9. *Let $Q = \mathcal{Q}(\mathbf{g})$ be a finitely generated quadratic module, and let $\mathbf{h}$ be a graded basis of $\sqrt[\mathbb{R}]{\operatorname{supp} Q}$. Then for any d we have $\mathcal{Q}_d(\mathbf{g}, \pm\mathbf{h}) = \overline{\mathcal{Q}_d(\mathbf{g}, \pm\mathbf{h})}$ is closed. Moreover, if we consider the extended relaxations $\mathcal{Q}_{2d}(\mathbf{g}, \pm\mathbf{h})$ and $\mathcal{L}_{2d}(\mathbf{g}, \pm\mathbf{h})$, then for any $f \in \mathbb{R}[\mathbf{X}]$ such that $f^* > -\infty$ we have that $f^*_{\mathrm{SoS},d}$ is attained (i.e. $f - f^*_{\mathrm{SoS},d} \in \mathcal{Q}_{2d}(\mathbf{g}, \pm\mathbf{h})$) and there is no duality gap: $f^*_{\mathrm{SoS},d} = f^*_{\mathrm{MoM},d}$.*

Proof. By [Mar08, lemma 4.1.4], $\mathcal{Q}_{2d}(\mathbf{g}, \pm\mathbf{h}) = \mathcal{Q}_{2d}(\mathbf{g}) + I_{2d}$ is closed. Therefore we have $\mathcal{L}_{2d}(\mathbf{g}, \pm\mathbf{h})^\vee = (\mathcal{Q}_{2d}(\mathbf{g}, \pm\mathbf{h}))^{\vee\vee} = \overline{\mathcal{Q}_{2d}(\mathbf{g}, \pm\mathbf{h})} = \mathcal{Q}_{2d}(\mathbf{g}, \pm\mathbf{h})$, from which we deduce that there is not duality gap, by classical convexity arguments, as follows.

If $f \in \mathbb{R}[\mathbf{X}]$ such that $f^* > -\infty$, then $\left\{\lambda \in \mathbb{R} \mid f - \lambda \in \mathcal{Q}_{2d}(\mathbf{g}, \pm\mathbf{h})\right\}$ is bounded from above. Since $\mathcal{Q}_{2d}(\mathbf{g}, \pm\mathbf{h})$ is closed $f^*_{\mathrm{SoS},d} = \sup\left\{\lambda \in \mathbb{R} \mid f - \lambda \in \mathcal{Q}_{2d}(\mathbf{g}, \pm\mathbf{h})\right\}$ is attained. If $f^*_{\mathrm{SoS},d} < f^*_{\mathrm{MoM},d}$, then $f - f^*_{\mathrm{MoM},d} \notin \mathcal{Q}_{2d}(\mathbf{g}, \pm\mathbf{h})$. Thus there exists a separating functional $\sigma \in \mathcal{L}^{(1)}_{2d}(\mathbf{g}, \pm\mathbf{h})$ such that $\left\langle \sigma \middle| f - f^*_{\mathrm{MoM},d} \right\rangle < 0$, which implies that $\langle \sigma | f \rangle < f^*_{\mathrm{MoM},d}$ in contradiction with the definition of $f^*_{\mathrm{MoM},d}$. Consequently, $f^*_{\mathrm{SoS},d} = f^*_{\mathrm{MoM},d}$. □

3.2 Annihilators of truncated moment sequences

Recall that the annihilator $\mathrm{Ann}_t(\sigma)$ is the kernel of the moment matrix of σ (or of the Hankel operator). With the characterization of $\widetilde{Q}$ we can now describe these kernels of moment matrices associated to truncated positive linear functionals.

We recall the definition of genericity in the truncated setting and equivalent characterizations.

Definition 3.10. We say that $\sigma^* \in \mathcal{L}_d(\mathbf{g})$ is *generic* if $\operatorname{rank} H^d_{\sigma^*} = \max\{\operatorname{rank} H^d_\eta \mid \eta \in \mathcal{L}_d(\mathbf{g})\}$.

This genericity can be characterized as follows, see [Las+13, prop. 4.7].

Proposition 3.11. *Let $\sigma \in \mathcal{L}_{2d}(\mathbf{g})$. The following are equivalent:*

(i) σ is generic;

(ii) $\mathrm{Ann}_d(\sigma) \subset \mathrm{Ann}_d(\eta)\ \forall \eta \in \mathcal{L}_{2d}(\mathbf{g})$;

(iii) $\forall d' \leq d$, we have: $\operatorname{rank} H^{d'}_\sigma = \max\{\operatorname{rank} H^{d'}_\eta \mid \eta \in \mathcal{L}_{2d}(\mathbf{g})\}$.

Remark. By Proposition 3.11 notice that $\forall d' \leq d$, if $\sigma^* \in \mathcal{L}_{2d}(\mathbf{g})$ is generic then $(\sigma^*)^{[2d']}$ is generic in $\mathcal{L}_{2d}(\mathbf{g})^{[2d']}$. In particular, $\mathrm{Ann}_{d'}(\sigma^*) \subset \mathrm{Ann}_{d'}(\eta)\ \forall \eta \in \mathcal{L}_{2d}(\mathbf{g})$.

The linear functionals in the relative interior of $\mathcal{L}_d(\mathbf{g})$ are generic. If we use an SDP solver based on interior point method we will (approximately) get a moment sequence in the relative interior of the face $\mathcal{L}_{2d}(\mathbf{g}) \cap \{\langle \sigma | f \rangle = f^*_{\mathrm{MoM},d}\}$, which is then generic in this face. We will use generic linear functionals to recover the minimizers when we have exactness or the flat truncation property.

We are now ready to describe the annihilator of generic elements.

Theorem 3.12. *Let $Q = \mathcal{Q}(\mathbf{g})$ and $J = \sqrt[\mathbb{R}]{\operatorname{supp} Q}$. Then for all $d, t \in \mathbb{N}$ big enough and for $\sigma^* \in \mathcal{L}_d(\mathbf{g})$ generic, we have $J = (\mathrm{Ann}_t(\sigma^*))$. Moreover if $Q = O$ is a preordering, then $(\mathrm{Ann}_t(\sigma^*)) = I(\mathcal{S}(\mathbf{g}))$.*

Proof. Let $t \in \mathbb{N}$ such that J is generated in degree $\leq t$, by the graded basis $\mathbf{h} = h_1, \ldots, h_s$. From Lemma 3.3 we deduce that there exists $d \in \mathbb{N}$ such that $J_{2t} \subset \overline{\mathcal{Q}_d(\mathbf{g})}$. Let $\sigma^* \in \mathcal{L}_d(\mathbf{g})$ generic.

We first prove that $J \subset (\mathrm{Ann}_t(\sigma^*))$. By Proposition 3.11 we have $\mathrm{Ann}_t(\sigma^*) = \bigcap_{\sigma \in \mathcal{L}_d(\mathbf{g})} \mathrm{Ann}_t(\sigma)$. Then it is enough to prove that $J_t \subset \mathrm{Ann}_t(\sigma)$ for all $\sigma \in \mathcal{L}_d(\mathbf{g})$.

By Lemma 3.6 $\mathcal{L}_d(\mathbf{g})^{[2t]} \subset \mathcal{L}_{2t}(\pm\mathbf{h}) \subset \langle \mathbf{h} \rangle_{2t}^\perp$. Then $\forall f \in J_t = \langle \mathbf{h} \rangle_t$, $\forall p \in \mathbb{R}[\mathbf{X}]_t$, $\forall \sigma \in \mathcal{L}_d(\mathbf{g})$, we have $fp \in \langle \mathbf{h} \rangle_{2t}$ and $\left\langle \sigma^{[2t]} \middle| fp \right\rangle = 0$. This shows that $H^t_\sigma(f)(p) = \left\langle (f \star \sigma)^{[t]} \middle| p \right\rangle = \langle \sigma | fp \rangle = 0$, i.e. $f \in \mathrm{Ann}_t(\sigma) = \ker H^t_\sigma$.

Conversely, we show that $(\mathrm{Ann}_t(\sigma^*)) \subset J$ for σ^* generic in $\mathcal{L}_d(\mathbf{g})$. Since $J = \operatorname{supp} \widetilde{Q} = \operatorname{supp} \bigcup_j \overline{\mathcal{Q}_j(\mathbf{g})}$ (by Theorem 3.4) it is enough to prove that $\mathrm{Ann}_t(\sigma^*) \subset \overline{\mathcal{Q}_d(\mathbf{g})} \cap -\overline{\mathcal{Q}_d(\mathbf{g})} = \operatorname{supp} \overline{\mathcal{Q}_d(\mathbf{g})} = \operatorname{supp} \mathcal{L}_d(\mathbf{g})^\vee$.

Let $f \in \mathrm{Ann}_t(\sigma^*) = \bigcap_{\sigma \in \mathcal{L}_k(\mathbf{g})} \mathrm{Ann}_t(\sigma)$ (we use again Proposition 3.11) and let $\sigma \in \mathcal{L}_d(\mathbf{g})$. Then $\langle \sigma | f \rangle = \left\langle (f \star \sigma)^{[t]} \middle| 1 \right\rangle = H^t_\sigma(f)(1) = 0$. In particular $f \in \mathcal{L}_d(\mathbf{g})^\vee$. We prove that $-f \in \mathcal{L}_d(\mathbf{g})^\vee$ in the same way. Then $f \in \mathcal{L}_d(\mathbf{g})^\vee \cap -\mathcal{L}_d(\mathbf{g})^\vee = \overline{\mathcal{Q}_d(\mathbf{g})} \cap -\overline{\mathcal{Q}_d(\mathbf{g})} = \operatorname{supp} \overline{\mathcal{Q}_d(\mathbf{g})}$, and finally we deduce from Definition 3.10 and Theorem 3.4 that $\mathrm{Ann}_t(\sigma^*) \subset \operatorname{supp} \widetilde{Q} = J$.

The second part follows from the first one and the Real Nullstelensatz, Theorem 1.3. □

Theorem 3.12 shows the possibilities and the limits of MoM relaxations. For instance we cannot expect exactness of the MoM relaxation $\mathcal{L}_{2d}(\mathbf{g})$ for any objective function f (i.e. $\mathcal{L}_{2d}(\mathbf{g})^{[k]} \subset \mathcal{M}(S)^{[k]}$) if $\sqrt[\mathbb{R}]{\operatorname{supp} Q} \neq \mathcal{I}(S)$: see Example 2.11.

In Proposition 3.13 we investigate the infinite dimensional case. We say that $\sigma^* \in \mathcal{L}(Q) = Q^\vee$ is *generic* if $\mathrm{Ann}(\sigma^*) \subset \mathrm{Ann}(\sigma)$ for all $\sigma \in \mathcal{L}(Q)$, using Proposition 3.11 to have the analogy with the finite dimensional case.

Proposition 3.13. *Let Q be a quadratic module, $S = \mathcal{S}(Q)$ and $\sigma^* \in \mathcal{L}(Q) = Q^\vee$ be generic. Then $\sqrt[\mathbb{R}]{\operatorname{supp} Q} \subset \mathrm{Ann}(\sigma^*) \subset \mathcal{I}(S)$. Moreover:*

(i) if Q is Archimedean then $\mathrm{Ann}(\sigma^*) = \mathcal{I}(S)$;

(ii) if $Q = O$ is a preordering, $\mathrm{Ann}(\sigma^*) = \mathcal{I}(S)$;

(iii) if I is an ideal of $\mathbb{R}[\mathbf{X}]$ and $\sigma^* \in \mathcal{L}(I)$ is generic, then $\mathrm{Ann}(\sigma^*) = \sqrt[\mathbb{R}]{I}$.

Proof. For $x \in S$, notice that $\mathbf{e}_x \in \mathcal{L}(Q)$. Then, since σ^* is generic:

$$\mathrm{Ann}(\sigma^*) \subset \bigcap_{x \in S} \mathrm{Ann}(\mathbf{e}_x) = \bigcap_{x \in S} \mathcal{I}(x) = \mathcal{I}(S).$$

Now observe that $\operatorname{supp} Q \subset \mathrm{Ann}(\sigma^*)$ by definition. Since $\mathrm{Ann}(\sigma^*)$ is a real radical ideal (see [Las+13, prop. 3.13]), then $\sqrt[\mathbb{R}]{\operatorname{supp} Q} \subset \mathrm{Ann}(\sigma^*)$.

If Q is Archimedean, then by Theorem 3.1 $\mathcal{L}(Q) = \mathcal{M}(S)$. In particular σ^* is a measure $\mu \in \mathcal{M}(S)$ supported on S: $\forall f \in \mathbb{R}[\mathbf{X}]$, $\langle \sigma^* | f \rangle = \int f \, \mathrm{d}\mu$. Let $h \in \mathcal{I}(S)$ and $f \in \mathbb{R}[\mathbf{X}]$. Then:

$$\langle \sigma^* | f h \rangle = \int f h \, \mathrm{d}\mu = \int 0 \, \mathrm{d}\mu = 0,$$

i.e. $h \in \mathrm{Ann}(\sigma^*)$, which proves the reverse inclusion.

If $Q = O$ is a preordering then $\sqrt[\mathbb{R}]{\operatorname{supp} O} = \mathcal{I}(S)$ by the Real Nullstellensatz, Theorem 1.3. Then $\mathrm{Ann}(\sigma^*) = \mathcal{I}(S)$.

As $\mathcal{L}(I) = \mathcal{L}(I + \Sigma^2)$, the last point follows from the previous one applied to $O = I + \Sigma^2$. □

If we compare Theorem 3.12 and Proposition 3.13, we see that the description in the infinite dimensional setting is more complicated, as we don't always have the equality $\mathrm{Ann}(\sigma^*) = \sqrt[\mathbb{R}]{\operatorname{supp} Q}$, see the case of an Archimedean quadratic module. This happens because limit properties that appear in the infinite dimensional case do not show up in the truncated setting.

3.3 Regularity, moment sequences and flat truncation

In this section, we analyze the properties of moment sequences in $\mathcal{L}_d(\mathbf{g})$ when $S = \mathcal{S}(\mathbf{g})$ is finite. We will use the results in this section to study the case of finitely many minimizers in Polynomial Optimization problems, and in particular flat truncation.

Let $\Xi = \{\xi_1, \dots, \xi_r\} \in \mathbb{C}^n$ be a finite set of (complex) points and let $\mathcal{I}(\Xi) = \{p \in \mathbb{C}[X] \mid p(\xi_i) = 0 \ \forall i \in 1, \dots, r\}$ be the complex vanishing ideal of the points Ξ. The *Castelnuovo-Mumford regularity* of an ideal I (resp. Ξ) is $\max_i(\deg S_i - i)$ where S_i is the ith module of syzygies in a minimal resolution of I (resp. $\mathcal{I}(\Xi)$). Let denote it by $\rho(I)$ (resp. $\rho(\Xi)$).

It is well known that Ξ admits a family of interpolator polynomials $(u_i) \subset \mathbb{C}[\mathbf{X}]$ such that $u_i(\xi_j) = \delta_{i,j}$. The minimal degree $\iota(\Xi)$ of a family of interpolator polynomials is called the *interpolation degree* of Ξ. Since a family of interpolator polynomials (p_i) is a basis of $\mathbb{C}[\mathbf{X}]/\mathcal{I}(\Xi)$, the ideal $\mathcal{I}(\Xi)$ is generated in degree $\leq \iota(\Xi) + 1$ and $\rho(\Xi) \leq \iota(\Xi) + 1$. A classical result [Eis05, th. 4.1] relates the interpolation degree of Ξ with its regularity, and the minimal degree of a basis of $\mathbb{C}[\mathbf{X}]/\mathcal{I}(\Xi)$. This result can be stated as follows, for real points $\Xi \subset \mathbb{R}^n$:

Proposition 3.14. *Let $\Xi = \{\xi_1, \dots, \xi_r\} \subset \mathbb{R}^r$ with regularity $\rho(\Xi)$. Then the minimal degree $\iota(\Xi)$ of a basis of $\mathbb{R}[\mathbf{X}]/\mathcal{I}(\Xi)$ is $\rho(\Xi) - 1$ and there exist interpolator polynomials $u_1, \dots, u_r \in \mathbb{R}[\mathbf{X}]_{\rho(\Xi)-1}$.*

Another property that we will use is the following:

Proposition 3.15 ([BS87])**.** *Any ideal $I \subset \mathbb{R}[\mathbf{X}]$ has a graded basis in degree less than or equal to its regularity $\rho(I)$.*

In particular, for a set of points $\Xi = \{\xi_1, \dots, \xi_r\}$, the ideal $\mathcal{I}(\Xi)$ has a graded (resp. Grobner, resp. border) basis of degree equal to the regularity $\rho(\Xi)$. The minimal degree of a monomial basis B of $\mathbb{R}[\mathbf{X}]/\mathcal{I}(\Xi)$ is $\iota(\Xi) = \rho(\Xi) - 1$. Such a finite basis B can be chosen so that it is stable by monomial division.

The next result shows that (resp. positive) moment sequences orthogonal to the vanishing ideal of the points, truncated above twice the regularity are coming from (resp. positive) measures:

Proposition 3.16. *Let $\Xi = \{\xi_1,\dots,\xi_r\} \subset \mathbb{R}^n$, $I = \mathcal{I}(\Xi)$ its real vanishing ideal and let $\rho = \rho(\Xi)$ the regularity of Ξ. For $t \geq \rho - 1$, $\sigma \in I_t^\perp$ if and only if $\sigma \in \langle \mathbf{e}_{\xi_1}^{[t]},\dots,\mathbf{e}_{\xi_r}^{[t]}\rangle$. Moreover if $t \geq \rho - 1$ and $\sigma \in \mathcal{L}_{2t}(I_{2t})$, then $\sigma \in \operatorname{cone}(\mathbf{e}_{\xi_1}^{[2t]},\dots,\mathbf{e}_{\xi_r}^{[2t]})$ and $\operatorname{rank} H_\sigma^t = r$.*

Proof. Let $u_1,\dots,u_r \in \mathbb{R}[\mathbf{X}]_t$ be interpolation polynomials of degree $\leq \rho - 1 \leq t$ (Proposition 3.14). Consider the sequence of vector space maps:

$$\begin{array}{rcl} 0 \to I_t \to \mathbb{R}[\mathbf{X}]_t & \xrightarrow{\psi} & \langle u_1,\dots,u_r\rangle \to 0 \\ p & \mapsto & \displaystyle\sum_{i=1}^r p(\xi_i)u_i, \end{array}$$

which is exact since $\ker\psi = \{p \in \mathbb{R}[\mathbf{X}]_t \mid p(\xi_i) = 0\} = I_t$. Therefore we have $\mathbb{R}[\mathbf{X}]_t = \langle u_1,\dots,u_r\rangle \oplus I_t$.

Let $\sigma \in I_t^\perp$. Then $\tilde{\sigma} = \sigma - \sum_{i=1}^r \langle \sigma | u_i\rangle \mathbf{e}_{\xi_i}^{[t]} \in I_t^\perp$ is such that $\langle \tilde{\sigma} | u_i\rangle = 0$ for $i = 1,\dots,r$. Thus, $\tilde{\sigma} \in \langle u_1,\dots,u_r\rangle^\perp \cap I_t^\perp = (\langle u_1,\dots,u_r\rangle \oplus I_t)^\perp = \mathbb{R}[\mathbf{X}]_t^\perp$, i.e. $\tilde{\sigma} = 0$ showing that $I_t^\perp \subset \langle \mathbf{e}_{\xi_1}^{[t]},\dots,\mathbf{e}_{\xi_r}^{[t]}\rangle$. The reverse inclusion is direct since I_t is the space of polynomials of degree $\leq t$ vanishing at ξ_i for $i = 1,\dots,r$.

Assume that $t \geq \rho - 1$ and $\sigma \in \mathcal{L}_{2t}(I_{2t})$. Then $\sigma \in I_{2t}^\perp$ and $\left\langle \sigma \middle| p^2 \right\rangle \geq 0$ for any $p \in \mathbb{R}[\mathbf{X}]_t$. By the previous analysis,

$$\sigma = \sum_{i=1}^r \omega_i \mathbf{e}_{\xi_i}^{[2t]}$$

As $0 \leq \left\langle \sigma \middle| u_i^2 \right\rangle = \omega_i$ for $i = 1,\dots,r$, we deduce that $\sigma \in \operatorname{cone}(\mathbf{e}_{\xi_1}^{[t]},\dots,\mathbf{e}_{\xi_r}^{[t]})$.

We verify that the image of $H_\sigma^t : p \in \mathbb{R}[\mathbf{X}]_t \mapsto \sum_{i=1}^r \omega_i p(\zeta_i)\mathbf{e}_{\xi_i}^{[t]}$ is $\langle \mathbf{e}_{\xi_1}^{[t]},\dots,\mathbf{e}_{\xi_r}^{[t]}\rangle$, computing $H_\sigma^t(u_i)$ for $i = 1,\dots,r$. Thus $\operatorname{rank} H_\sigma^t = \dim\langle \mathbf{e}_{\xi_1}^{[t]},\dots,\mathbf{e}_{\xi_r}^{[t]}\rangle = r$ since $(\mathbf{e}_{\xi_i}^{[t]})_{i=1,\dots,r}$ is the dual basis of $(u_i)_{i=1,\dots,r}$. □

We deduce another corollary, giving degree bounds for the case of a graded basis of a real radical ideal.

Proposition 3.17. *Let $I = \sqrt[\mathbb{R}]{I}$ be a real radical ideal and $\mathbf{h} = h_1,\dots,h_m$ be a graded basis of I. Then for all $d \geq \max_i(\deg h_i)$ and $\sigma^* \in \mathcal{L}_{2d}(\pm\mathbf{h})$ generic, we have $\operatorname{Ann}_d(\sigma^*) = I_d$.*

Proof. Let $d \geq \max_i(\deg h_i)$ and $\sigma^* \in \mathcal{L}_{2d}(\pm\mathbf{h})$ generic. Then for all i we have $(h_i \star \sigma^*)^{[d-\deg h_i]} = 0$. Now let $p \in I_d$. Since h is a graded basis we have $p = \sum_{i=1}^m p_i h_i$, where $\deg p_i \leq d - \deg h_i$. Notice that $(p \star \sigma^*)^{[d]} = \sum_{i=1}^m (p_i \star (h_i \star \sigma^*))^{[d]}$. As $(h_i \star \sigma^*)^{[d-\deg h_i]} = 0$ we conclude $(p \star \sigma^*)^{[d]} = 0$ and thus $p \in \operatorname{Ann}_d(\sigma^*)$. Therefore $I_d \subset \operatorname{Ann}_d(\sigma^*)$.

Conversely, let $p \in \operatorname{Ann}_d(\sigma^*)$. For all $\xi \in \mathcal{V}(I)$ we have $\mathbf{e}_\xi^{[2d]} \in \mathcal{L}_{2d}(\pm\mathbf{h})$. Since σ^* is generic, we have $p \in \operatorname{Ann}_d(\mathbf{e}_\xi)$. In particular $\left\langle \mathbf{e}_\xi \middle| p \right\rangle = p(\xi) = 0$, and therefore p vanishes on all the points of $\mathcal{V}(I)$. Since I is real radical, $p \in I \cap \mathbb{R}[\mathbf{X}]_d = I_d$ and thus $\operatorname{Ann}_d(\sigma^*) \subset I_d$, which concludes the proof. □

We describe now a property, known as *flat truncation*, which allows to test effectively if truncated moment sequences are coming from sums of evaluations.

Definition 3.18 (Flat truncation)**.** Let $d_{\mathbf{g}} = \lceil \frac{1}{2}\max_{i=1,\dots,s}\deg(g_i)\rceil$. The *flat truncation* property holds for $\sigma \in \mathcal{L}_d(\mathbf{g})$ at degree t if $t \leq \frac{d}{2} - d_{\mathbf{g}}$ and

$$\operatorname{rank} H_\sigma^t = \operatorname{rank} H_\sigma^{t+d_{\mathbf{g}}}. \tag{4}$$

This definition coincides with the definition of flat truncation used in [CF96], [Lau09] or [Nie13b]. We investigate more in detail rank conditions for the moment matrix of $\sigma \in \mathcal{L}_d(\mathbf{g})$.

Lemma 3.19. *If $\sigma \in \mathcal{L}_d(\mathbf{g})$ is such that* $\operatorname{rank} H_\sigma^t = \operatorname{rank} H_\sigma^{t+s} = r$ *with $t+1 \le t+s \le \frac{d}{2}$, then*

$$\sigma^{[t+s+\frac{d}{2}]} = \omega_1 \mathbf{e}_{\xi_1}^{[t+s+\frac{d}{2}]} + \cdots + \omega_r \mathbf{e}_{\xi_r}^{[t+s+\frac{d}{2}]}$$

for some points $\xi_i \in \mathbb{R}^n$ and weights $\omega_i > 0$, $i = 1,\dots,r$. Denoting $\Xi = \{\xi_1,\dots,\xi_r\}$, we also have $\operatorname{Ann}_{t+1}(\sigma) = \mathcal{I}(\Xi)_{t+1}$ *and* $\mathcal{V}(\operatorname{Ann}_{t+1}(\sigma)) = \Xi$ *(or, in other words,* $(\operatorname{Ann}_{t+1}(\sigma)) = \mathcal{I}(\Xi)$*).*

Moreover, if $t \le \frac{d}{2} + s - \deg(\mathbf{g})$, where $\deg(\mathbf{g}) = \max_{i=1,\dots,s} \deg(g_i)$*, the inclusion $\Xi \subset \mathcal{S}(\mathbf{g})$ holds true.*

Proof. From [Lau09, th. 5.29], there exists unique $\Xi = \{\xi_1,\dots,\xi_r\} \subset \mathbb{R}^n$ and $\omega_1,\dots,\omega_r > 0$ such that $\sigma^{[2(t+s)]} = \omega_1 \mathbf{e}_{\xi_1}^{[2(t+s)]} + \cdots + \omega_r \mathbf{e}_{\xi_r}^{[2(t+s)]}$, $(\operatorname{Ann}_{t+s}(\sigma)) = \mathcal{I}(\Xi)$ and $\mathcal{V}(\operatorname{Ann}_{t+s}(\sigma)) = \Xi$. In particular $(\operatorname{Ann}_{t+s}(\sigma))$ is a zero dimensional ideal and $\operatorname{Ann}_{t+s}(\sigma) \subset I(\Xi)_{t+s}$. Conversely, for any $h \in I(\Xi)_{t+s}$, we have

$$\left\langle \sigma \middle| h^2 \right\rangle = \left\langle \sigma^{[2(t+s)]} \middle| h^2 \right\rangle = \sum_{i=1}^r \omega_i \left\langle \mathbf{e}_{\xi_i}^{[2(t+s)]} \middle| h^2 \right\rangle = \sum_{i=1}^r \omega_i h^2(\xi_i) = 0.$$

Thus $h \in \operatorname{Ann}_{t+s}(\sigma)$ (see see [Las+13, lem. 3.12]) and $I(\Xi)_{t+s} = \operatorname{Ann}_{t+s}(\sigma)$.

As $\operatorname{rank} H_\sigma^t = \operatorname{rank} H_\sigma^{t+1} = r$, we deduce from above, that $(\operatorname{Ann}_{t+1}(\sigma)) = I(\Xi)$ is generated in degree $\le t+1$ and that $\rho(\Xi) \le t+1$. Therefore Ξ has interpolator polynomials $u_1,\dots,u_r$ of degree $\le t$.

Let us show that the description of σ on polynomials of degree $\le 2(t+s)$, can be extended to higher degree. For any $h \in \operatorname{Ann}_{t+s}(\sigma) = I(\Xi)_{s+t}$, i.e. such that $\left\langle \sigma \middle| h^2 \right\rangle = 0$, and any $p \in \mathbb{R}[\mathbf{X}]_{\frac{d}{2}}$ we have $\langle \sigma | hp \rangle = 0$. This shows that $\sigma \in (\mathcal{I}(\Xi)_{t+s+\frac{d}{2}})^\perp$. We deduce from Proposition 3.16 that $\sigma^{[t+s+\frac{d}{2}]} \in \operatorname{cone}(\mathbf{e}_{\xi_1},\dots,\mathbf{e}_{\xi_r})^{[t+s+\frac{d}{2}]}$. This implies that $\sigma^{[t+s+\frac{d}{2}]} = \omega_1 \mathbf{e}_{\xi_1}^{[t+s+\frac{d}{2}]} + \cdots + \omega_r \mathbf{e}_{\xi_r}^{[t+s+\frac{d}{2}]}$, evaluating $\langle \sigma | u_i \rangle = \left\langle \sigma^{[t+s+\frac{d}{2}]} \middle| u_i \right\rangle = \omega_i$ at the interpolator polynomials $u_1,\dots u_r$ of Ξ of degree $\le t$.

We show now that $\Xi = \{\xi_1,\dots,\xi_r\} \subset S$ if $t \le \frac{d}{2} + s - \deg(\mathbf{g})$. For $i = 1,\dots,r$ and $j = 1,\dots,m$ the polynomial $u_i^2 g_j$ has degree $\le 2t + \deg(\mathbf{g}) \le t + s + \frac{d}{2}$. Then we obtain:

$$0 \le \left\langle \sigma \middle| u_i^2 g_j \right\rangle = \left\langle \sigma^{[t+s+\frac{d}{2}]} \middle| u_i^2 g_j \right\rangle = \left\langle \omega_1 \mathbf{e}_{\xi_1}^{[t+s+\frac{d}{2}]} + \cdots + \omega_r \mathbf{e}_{\xi_r}^{[t+s+\frac{d}{2}]} \middle| u_i^2 g_j \right\rangle = g_j(\xi_i),$$

showing that $g_j(\xi_i) \ge 0$ for all i and j, i.e. $\Xi \subset \mathcal{S}(\mathbf{g})$. □

Remark. Lemma 3.19 can be used to test flat truncation in a simpler way when d is big, as we explain in the following. Assume for simplicity that $2d_{\mathbf{g}} = \deg(\mathbf{g})$. Then, if $\operatorname{rank} H_\sigma^t = \operatorname{rank} H_\sigma^{t+s}$ with $t \le \frac{d}{2} + s - \deg(\mathbf{g})$, then $2(t + d_{\mathbf{g}}) = 2t + \deg(\mathbf{g}) \le t + s + \frac{d}{2}$. Then from Lemma 3.19 we deduce that σ restricted to polynomials of degree $\le 2(t + d_{\mathbf{g}})$ is equal to a sum of evaluations at points of S with positive weights, and the flat truncation is satisfied: $\operatorname{rank} H_\sigma^t = \operatorname{rank} H_\sigma^{t+d_{\mathbf{g}}}$. In particular, when $s = 1$ and $d \ge t - 2 + 2\deg(\mathbf{g})$, $\operatorname{rank} H_\sigma^t = \operatorname{rank} H_\sigma^{t+1}$ implies $\operatorname{rank} H_\sigma^t = \operatorname{rank} H_\sigma^{t+d_{\mathbf{g}}}$.

We now show that we can use flat truncation to describe semialgebraic sets with a finite number of points.

Theorem 3.20. *If a positive linear functional $\sigma^* \in \mathcal{L}_d(\mathbf{g})$ is such that $(\sigma^*)^{[2(t+d_{\mathbf{g}})]}$ is generic in $\mathcal{L}_d(\mathbf{g})^{[2(t+d_{\mathbf{g}})]}$ (that is $\operatorname{Ann}_{t+d_{\mathbf{g}}}(\sigma^*) \subset \operatorname{Ann}_{t+d_{\mathbf{g}}}(\sigma)$ for all $\sigma \in \mathcal{L}_d(\mathbf{g})$) and σ^* satisfies the flat truncation property at degree $t \le \frac{d}{2} - d_{\mathbf{g}}$, then:*

(i) $S = \mathcal{S}(\mathbf{g}) = \{\xi_1,\dots,\xi_r\}$ *is non-empty and finite;*

(ii) $\mathcal{L}_d(\mathbf{g})^{[t+d_{\mathbf{g}}+\frac{d}{2}]} = \operatorname{cone}(\mathbf{e}_{\xi_1},\dots,\mathbf{e}_{\xi_r})^{[t+d_{\mathbf{g}}+\frac{d}{2}]}$;

(iii) $t+1 \ge \rho(\xi_1,\dots,\xi_r)$ *and* $\operatorname{Ann}_{t+1}(\sigma^*) = \mathcal{I}(\xi_1,\dots,\xi_r)_{t+1} = \mathcal{I}(S)_{t+1}$ *is the vanishing ideal of S truncated in degree $t+1$.*

(iv) $\mathcal{I}(S)_{2(t+d_{\mathbf{g}})} \subset \overline{\mathcal{Q}_d(\mathbf{g})}$ *and* $(\operatorname{Ann}_{t+1}(\sigma^*)) = \sqrt[\mathbb{R}]{\operatorname{supp} \mathcal{Q}(\mathbf{g})} = \mathcal{I}(S)$.

Proof. Let $\sigma^* \in \mathcal{L}_d(\mathbf{g})$ be such that $(\sigma^*)^{[2(t+d_\mathbf{g})]}$ is generic in $\mathcal{L}_d(\mathbf{g})^{[2(t+d_\mathbf{g})]}$, and assume that $\operatorname{rank} H^t_{\sigma^*} = \operatorname{rank} H^{t+d_\mathbf{g}}_{\sigma^*}$ with $t \leq \frac{d}{2} - d_\mathbf{g}$. By Lemma 3.19 applied with $s = d_\mathbf{g}$,

$$(\sigma^*)^{[t+d_\mathbf{g}+\frac{d}{2}]} = \omega_1 \mathbf{e}_{\xi_1}^{[t+d_\mathbf{g}+\frac{d}{2}]} + \cdots + \omega_r \mathbf{e}_{\xi_r}^{[t+d_\mathbf{g}+\frac{d}{2}]}$$

with $\omega_i > 0$, $\Xi = \{\xi_1, \ldots, \xi_r\} \subset S$, $\operatorname{Ann}_{t+1}(\sigma^*) = I(\Xi)_{t+1}$ and $(\operatorname{Ann}_{t+1}(\sigma^*)) = I(\Xi)$.

Let $\mathbf{h} = h_1, \ldots, h_m \subset \operatorname{Ann}_{t+1}(\sigma^*)$ be a graded basis of $I(\Xi)$ of degree $\leq t+1$. As $(\sigma^*)^{[2(t+d_\mathbf{g})]}$ is generic, for any $\sigma \in \mathcal{L}_d(\mathbf{g})$ we have $\operatorname{Ann}_{t+d_\mathbf{g}}(\sigma^*) \subset \operatorname{Ann}_{t+d_\mathbf{g}}(\sigma)$ and $\left\langle \sigma \middle| h_i^2 \right\rangle = 0$. Then for any $p \in \mathbb{R}[\mathbf{X}]_{d_\mathbf{g}+\frac{d}{2}}$ we have $\langle \sigma | h_i p \rangle = 0$, proving that $\sigma \in (\mathbf{h})^\perp_{t+d_\mathbf{g}+\frac{d}{2}} = (\mathcal{I}(\Xi)_{t+d_\mathbf{g}+\frac{d}{2}})^\perp$, i.e. $\mathcal{L}_d(\mathbf{g})^{[t+d_\mathbf{g}+\frac{d}{2}]} \subset (\mathcal{I}(\Xi)_{t+d_\mathbf{g}+\frac{d}{2}})^\perp$. We deduce from Proposition 3.16 that $\sigma^{[t+d_\mathbf{g}+\frac{d}{2}]} \in \operatorname{cone}(\mathbf{e}_{\xi_1}, \ldots, \mathbf{e}_{\xi_r})^{[t+d_\mathbf{g}+\frac{d}{2}]}$. This shows that $\mathcal{L}_d(\mathbf{g})^{[t+d_\mathbf{g}+\frac{d}{2}]} \subset \operatorname{cone}(\mathbf{e}_{\xi_1}, \ldots, \mathbf{e}_{\xi_r})^{[t+d_\mathbf{g}+\frac{d}{2}]}$. On the other hand the inclusion $\mathcal{L}_d(\mathbf{g})^{[t+d_\mathbf{g}+\frac{d}{2}]} \supset \operatorname{cone}(\mathbf{e}_{\xi_1}, \ldots, \mathbf{e}_{\xi_r})^{[t+d_\mathbf{g}+\frac{d}{2}]}$ holds true since $\Xi \subset S$. Therefore

$$\mathcal{L}_d(\mathbf{g})^{[t+d_\mathbf{g}+\frac{d}{2}]} = \operatorname{cone}(\mathbf{e}_{\xi_1}, \ldots, \mathbf{e}_{\xi_r})^{[t+d_\mathbf{g}+\frac{d}{2}]}.$$

Let us show that $\Xi = S$. For $\zeta \in S$ we have $\mathbf{e}_\zeta^{[t+d_\mathbf{g}+\frac{d}{2}]} \in \mathcal{L}_d(\mathbf{g})^{[t+d_\mathbf{g}+\frac{d}{2}]} \subset (\mathbf{h})^\perp_{t+d_\mathbf{g}+\frac{d}{2}}$, and thus for $i = 1, \ldots, m$, $\left\langle \mathbf{e}_\zeta \middle| h_i \right\rangle = h_i(\zeta) = 0$. This shows that ζ is a root of $\mathbf{h}$ and thus $\zeta \in \Xi$. We conclude that $\Xi = \{\xi_1, \ldots, \xi_r\} = S$.

The inclusion $\mathcal{I}(S)_{2(t+d_\mathbf{g})} \subset \overline{\mathcal{Q}_d(\mathbf{g})}$ follows from $\mathcal{L}_d(\mathbf{g})^{[t+d_\mathbf{g}+\frac{d}{2}]} \subset (\mathbf{h})^\perp_{t+d_\mathbf{g}+\frac{d}{2}}$. Indeed $2(t+d_\mathbf{g}) \leq t + d_\mathbf{g} + \frac{d}{2}$ and thus $\mathcal{L}_d(\mathbf{g})^{[2(t+d_\mathbf{g})]} \subset (\mathbf{h})^\perp_{2(t+d_\mathbf{g})}$. Now notice that $(\mathcal{L}_d(\mathbf{g})^{[2(t+d_\mathbf{g})]})^\vee \subset \overline{\mathcal{Q}_d(\mathbf{g})}$, using convex duality. Therefore dualizing $\mathcal{L}_d(\mathbf{g})^{[2(t+d_\mathbf{g})]} \subset (\mathbf{h})^\perp_{2t}$ we obtain the desired inclusion. Moreover $\mathcal{I}(S)_{2(t+d_\mathbf{g})} \subset \overline{\mathcal{Q}_d(\mathbf{g})} \cap -\overline{\mathcal{Q}_d(\mathbf{g})} \subset \operatorname{supp} \widetilde{Q} = \sqrt[\mathbb{R}]{\operatorname{supp} Q}$, by Theorem 3.4, and finally:

$$(\operatorname{Ann}_{t+1}(\sigma^*)) = \mathcal{I}(S) = (\mathcal{I}(S)_{2(t+d_\mathbf{g})}) \subset \sqrt[\mathbb{R}]{\operatorname{supp} \mathcal{Q}(\mathbf{g})} \subset \sqrt[\mathbb{R}]{\operatorname{supp} \mathcal{O}(\mathbf{g})} = \mathcal{I}(S),$$

where the last equality is the Real Nullstellenstatz, Theorem 1.3. This shows that $(\operatorname{Ann}_{t+1}(\sigma^*)) = \sqrt[\mathbb{R}]{\operatorname{supp} \mathcal{Q}(\mathbf{g})} = \mathcal{I}(S)$. □

This theorem tells us that if the flat truncation property holds at degree $t \leq \frac{d}{2} - d_\mathbf{g}$, then any element of $\mathcal{L}_d(\mathbf{g})$ truncated in degree $t + \frac{d}{2} + d_\mathbf{g}$ coincides with a positive measure supported in $S = \{\xi_1, \ldots, \xi_r\}$.

In the following theorem we show that when $\operatorname{supp}(Q)$ is a zero-dimensional ideal (and thus S is finite), the rank condition is satisfied for any moment matrix.

Theorem 3.21. *Suppose that* $\dim \frac{\mathbb{R}[\mathbf{X}]}{\operatorname{supp} \mathcal{Q}(\mathbf{g})} = 0$. *Then* $S = \mathcal{S}(\mathbf{g})$ *is finite and there exists* $d \geq 2(\rho - 1 + d_\mathbf{g})$ *such that* $\mathcal{I}(S)_{2(\rho-1+d_\mathbf{g})} \subset \operatorname{supp} \overline{\mathcal{Q}_d(\mathbf{g})}$, *where* $\rho = \rho(S)$ *is the regularity of* S, *and for any* $\sigma \in \mathcal{L}_d(\mathbf{g})$ *the flat truncation property holds at degree* $\rho - 1$.

Proof. Let $I = \operatorname{supp} \mathcal{Q}(\mathbf{g})$ and $J = \sqrt[\mathbb{R}]{\operatorname{supp} \mathcal{Q}(\mathbf{g})}$. From Lemma 1.2 we deduce $\dim \frac{\mathbb{R}[\mathbf{X}]}{J} = \dim \frac{\mathbb{R}[\mathbf{X}]}{I} = 0$ and by Theorem 1.4 we have $\mathcal{I}(\mathcal{S}(\mathbf{g})) = \sqrt[\mathbb{R}]{\operatorname{supp} \mathcal{Q}(\mathbf{g})} = J$. Then $\mathcal{V}_\mathbb{R}(J) = \mathcal{V}_\mathbb{R}(\mathcal{I}(\mathcal{S}(\mathbf{g}))) = \mathcal{S}(\mathbf{g}) = \{\xi_1, \ldots, \xi_r\}$ is finite.

We choose a graded basis $\mathbf{h}$ of J with $\deg \mathbf{h} \leq \rho = \rho(\xi_1, \ldots, \xi_r)$ (by Proposition 3.15). By Corollary 3.8, there exists $d \in \mathbb{N}$ such that $\mathcal{I}(S)_{2(\rho-1+d_\mathbf{g})} \subset \operatorname{supp} \overline{\mathcal{Q}_d(\mathbf{g})}$. From Corollary 3.8 and Proposition 3.16 we deduce that positive linear functionals in $\mathcal{L}_d(\mathbf{g})$ restricted to degree $\leq 2(\rho - 1 + d_\mathbf{g})$ are conical sums of evaluations at $\xi_1, \ldots, \xi_r$:

$$\mathcal{L}_d(\mathbf{g})^{[2(\rho-1+d_\mathbf{g})]} \subset \mathcal{L}_{2(\rho-1+d_\mathbf{g})}(\pm\mathbf{h}) = \mathcal{L}_{2(\rho-1+d_\mathbf{g})}(J_{2(\rho-1+d_\mathbf{g})}) = \operatorname{cone}(\mathbf{e}_{\xi_1}, \ldots, \mathbf{e}_{\xi_r})^{[2(\rho-1+d_\mathbf{g})]},$$

and for all $\sigma \in \mathcal{L}_d(\mathbf{g})$, we have $\operatorname{rank} H^{\rho-1}_\sigma = \operatorname{rank} H^{\rho-1+d_\mathbf{g}}_\sigma$. □

Theorem 3.21 says that if $\dim \frac{\mathbb{R}[\mathbf{X}]}{\operatorname{supp}\mathcal{Q}(\mathbf{g})} = 0$ then the minimal order for which we have flat truncation is not bigger than $d \geq 2(\rho - 1 + d_{\mathbf{g}})$ such that $\mathcal{I}(S)_{2(\rho-1+d_{\mathbf{g}})} \subset \operatorname{supp}\overline{\mathcal{Q}_d(\mathbf{g})}$. This degree is related to the minimal d for which $\mathcal{I}(S) = \sqrt[\mathbb{R}]{\operatorname{supp}\mathcal{Q}(\mathbf{g})}$ is generated by $\operatorname{supp}\overline{\mathcal{Q}_d(\mathbf{g})}$, that is, the minimal degree d such that $I(S)_{\rho-1+d_{\mathbf{g}}} \subset \operatorname{Ann}_{\frac{d}{2}}(\sigma^*)$ for a generic $\sigma^* \in \mathcal{L}_d(\mathbf{g})$. Moreover, as in the remark after Lemma 3.19, we can replace $\rho - 1 + d_{\mathbf{g}}$ with ρ if d is big enough.

Theorem 3.21 and Theorem 3.20 show that if $\dim \frac{\mathbb{R}[\mathbf{X}]}{\operatorname{supp}\mathcal{Q}(\mathbf{g})} = 0$ then S is a finite set of points and for a high enough degree d, all moment sequences in $\mathcal{L}_d(\mathbf{g})$, truncated in degree twice the regularity are coming from a weighted sum of Dirac measures at these points. In particular, it is possible to recover all the points in S from a generic truncated moment sequence, see [HL05], [ABM15] and [Mou18].

Remark. There exist examples with $\mathcal{S}(\mathbf{g})$ finite and $\dim \frac{\mathbb{R}[\mathbf{X}]}{\operatorname{supp}\mathcal{Q}(\mathbf{g})} > 1$, see Example 3.5. However the hypotheses (i) $\dim \frac{\mathbb{R}[\mathbf{X}]}{\operatorname{supp}\mathcal{Q}(\mathbf{g})} = 0$; and (ii) $\mathcal{S}(\mathbf{g})$ is finite and $\dim \frac{\mathbb{R}[\mathbf{X}]}{\operatorname{supp}\mathcal{Q}(\mathbf{g})} \leq 1$; are equivalent: (i) $\Rightarrow$ (ii) is shown in the proof of Theorem 3.21, while (ii) $\Rightarrow$ (i) follows from $\mathcal{I}(\mathcal{S}(\mathbf{g})) = \sqrt[\mathbb{R}]{\operatorname{supp}\mathcal{Q}(\mathbf{g})}$ (see Theorem 1.4).

Results related to Theorem 3.21 and Theorem 3.20 were obtained in [LLR08] and [Las+13], where they focus on the case of equations $\mathbf{h}$ defining a finite real variety. They prove that, for d big enough and for every positive linear functional $\sigma \in \mathcal{L}_{2d}(\pm\mathbf{h})$, the flat truncation property holds for H_σ^d, and that $\sigma^{[2d]}$ is a conic linear combination of evaluations at the points of $\mathcal{V}_{\mathbb{R}}(\mathbf{h})$. This can be deduced from Theorem 3.21, since in the case where $\mathcal{V}_{\mathbb{R}}(\mathbf{h}) = \{\xi_1, \ldots, \xi_r\}$ is non-empty and finite, $\dim \frac{\mathbb{R}[\mathbf{X}]}{\operatorname{supp}\mathcal{Q}(\pm\mathbf{h})} = 0$.

In [LLR08, rem. 4.9] it is also mentioned that the same can be proved for a preordering defining a finite semialgebraic set. This result can also be deduced from Theorem 3.21, since when $S = S(\mathbf{g}) = \{\xi_1, \ldots, \xi_r\}$ is non-empty and finite, we have by the Real Nullstellensatz , $\dim \frac{\mathbb{R}[\mathbf{X}]}{\operatorname{supp}\mathcal{O}(\mathbf{g})} = \dim \frac{\mathbb{R}[\mathbf{X}]}{\sqrt[\mathbb{R}]{\operatorname{supp}\mathcal{O}(\mathbf{g})}} = \dim \frac{\mathbb{R}[\mathbf{X}]}{\mathcal{I}(\mathcal{S}(\mathbf{g}))} = 0$.

But Theorem 3.21 is more general, as shown by the following example of a quadratic module, whose support is zero dimensional, but that is not a preordering.

Example 3.22 ([Mar08, ex. 7.4.5 (1)])**.** Let $Q = \mathcal{Q}(X, Y, 1-X, 1-Y, -X^4, -Y^4) \subset \mathbb{R}[X, Y]$. In this case $\operatorname{supp} Q$, which contains X^4 and Y^4, is zero dimensional and Q is not a preordering since $XY \notin Q$. Theorem 3.21 applies in this case, but the results cannot be deduced from [LLR08] or [Las+13].

As we will see, in Polynomial Optimization problems, flat truncation implies MoM exactness and thus finite convergence. Moreover, it allows extracting the minimizers from an optimal sequence.

4 Flat truncation in polynomial optimization problems

In this section, we analyze when flat truncation occurs in the Polynomial Optimization Problem, which consists of minimizing $f \in \mathbb{R}[\mathbf{X}]$ on the basic semialgebraic set $S = \mathcal{S}(\mathbf{g})$ where $\mathbf{g} = g_1, \ldots, g_s$ is a tuple of polynomials. Recall that we denote f^* the minimum of f on S. We will consider the semialgebraic set $S^{\min} = \mathcal{S}(\mathbf{g}, \pm(f - f^*)) = \mathcal{S}(\mathbf{g}) \cap \{x \in \mathbb{R}^n \mid f(x) = f^*\}$ and assume that it is nonempty.

4.1 Flat truncation degree

Hereafter, we analyze the degree at which flat truncation holds and yields the minimizers.

Theorem 4.1. *Consider the problem of minimizing f on $\mathcal{S}(\mathbf{g})$. If the flat truncation property holds for a generic $\sigma^* \in \mathcal{L}_{2d}^{\min}(\mathbf{g})$ at a degree t such that $\deg(f) - d_{\mathbf{g}} - d \leq t \leq d - d_{\mathbf{g}}$, then:*

(i) $f^* = f^*_{\mathrm{MoM},d}$ (i.e. we have MoM finite convergence);

(ii) the set of minimizers $S^{\min} = \{\xi_1, \dots, \xi_r\}$ is non-empty and finite;

(iii) $\ker H_{\sigma^*}^{t+1} = \mathrm{Ann}_{t+1}(\sigma^*) = \mathcal{I}(S^{\min})_{t+1}$ *(i.e. the kernel of the truncated moment matrix equals the truncated ideal of the minimizers) and* $\mathcal{V}(\mathrm{Ann}_{t+1}(\sigma^*)) = S^{\min}$*;*

(iv) $\mathcal{L}_d^{\min}(\mathbf{g})^{[t+d_{\mathbf{g}}+d]} = \mathrm{cone}(\mathbf{e}_{\xi_1}, \dots, \mathbf{e}_{\xi_r})^{[t+d_{\mathbf{g}}+d]}$ *(i.e. all the minimizing truncated feasible moment sequences are conic sums of evaluations at the minimizers);*

(v) the MoM relaxation is exact.

Proof. Let $\sigma^* \in \mathcal{L}_{2d}^{\min}(\mathbf{g})$ be generic such that $\mathrm{rank}\, H_{\sigma^*}^t = \mathrm{rank}\, H_{\sigma^*}^{t+d_{\mathbf{g}}}$ with $\deg(f) \le t+d_{\mathbf{g}}+d$ and $t+d_{\mathbf{g}} \le d$. Then by Lemma 3.19, $(\sigma^*)^{[t+d_{\mathbf{g}}+d]} = \sum_{i=1}^r \omega_i \mathbf{e}_{\xi_i}^{[t+d_{\mathbf{g}}+d]}$ with $\xi_i \in S = \mathcal{S}(\mathbf{g})$, $\omega_i > 0$, $\mathrm{Ann}_{t+1}(\sigma^*) = \mathcal{I}(\xi_1, \dots \xi_r)_{t+1} = \mathcal{I}(\Xi)_{t+1}$ and $\mathcal{V}(\mathrm{Ann}_{t+1}(\sigma^*)) = \Xi$. Notice that $f(\xi_i) \ge f^*$ since $\xi_i \in S$.

We show now that $S^{\min} = \Xi$. As $\langle \sigma^* | 1 \rangle = 1$ we have $\sum_{i=1}^r \omega_i = 1$. Moreover $f_{\mathrm{MoM},d}^* = \langle \sigma^* | f \rangle \le f^*$ and since $\deg(f) \le t + d_{\mathbf{g}} + d$ we obtain:

$$f^* \ge \langle \sigma^* | f \rangle = \left\langle (\sigma^*)^{[t+d_{\mathbf{g}}+d]} \middle| f \right\rangle = \sum_{i=1}^r \omega_i \left\langle \mathbf{e}_{\xi_i}^{[t+d_{\mathbf{g}}+d]} \middle| f \right\rangle = \sum_{i=1}^r \omega_i f(\xi_i) \ge f^*.$$

This implies that $f(\xi_i) = f^*$ for $i = 1, \dots, r$. Therefore $f^* = f_{\mathrm{MoM},d}^*$ and $S^{\min} \supset \Xi$.

From Proposition 3.11 we have that $\sigma^* \in \mathcal{L}_{2d}^{\min}(\mathbf{g})$ generic implies that $(\sigma^*)^{[2(t+d_{\mathbf{g}})]}$ is generic in $\mathcal{L}_{2d}^{\min}(\mathbf{g})^{[2(t+d_{\mathbf{g}})]}$. Moreover $(\sigma^*)^{[2(t+d_{\mathbf{g}})]} = \sum_{i=1}^r \omega_i \mathbf{e}_{\xi_i}^{[2(t+d_{\mathbf{g}})]} \in \mathcal{L}_{2d}^{(1)}(\mathbf{g}, \pm(f - f^*))^{[2(t+d_{\mathbf{g}})]}$ since $\Xi \subset S^{\min} = \mathcal{S}(\mathbf{g}, \pm(f - f^*))$. Then, as $\mathcal{L}_{2d}^{(1)}(\mathbf{g}, \pm(f - f^*)) \subset \mathcal{L}_{2d}^{\min}(\mathbf{g})$ and $(\sigma^*)^{[2(t+d_{\mathbf{g}})]}$ is generic in $\mathcal{L}_{2d}^{\min}(\mathbf{g})^{[2(t+d_{\mathbf{g}})]}$, we have

$$\forall \sigma \in \mathcal{L}_{2d}(\mathbf{g}, \pm(f - f^*)) \quad \mathrm{Ann}_{t+d_{\mathbf{g}}}(\sigma^*) \subset \mathrm{Ann}_{t+d_{\mathbf{g}}}(\sigma),$$

i.e. $(\sigma^*)^{[2(t+d_{\mathbf{g}})]}$ is generic in $\mathcal{L}_{2d}(\mathbf{g}, \pm(f - f^*))^{[2(t+d_{\mathbf{g}})]}$. We can then conclude from Theorem 3.20 that $S^{\min} = \Xi$ and $\mathcal{L}_d^{\min}(\mathbf{g})^{[t+d_{\mathbf{g}}+d]} = \mathrm{cone}(\mathbf{e}_{\xi_1}, \dots, \mathbf{e}_{\xi_r})^{[t+d_{\mathbf{g}}+d]}$.

Finally we show MoM exactness. For every $d' \ge d$ and $\sigma \in \mathcal{L}_{2d'}^{\min}(\mathbf{g})$, we have $\sigma^{[2d]} \in \mathcal{L}_{2d}^{\min}(\mathbf{g})$ since $\langle \sigma | f \rangle = f^*$. Therefore σ has flat truncation in degree t and by Lemma 3.19, $\sigma^{[t+d_{\mathbf{g}}+d']}$ is coming from a convex sum of Dirac measures at points in S (that are the minimizers $\xi_1, \dots, \xi_r$). This shows that the moment relaxation is exact, since increasing d' we increase also the truncation degree where σ coincides with a weighted sum of evaluations at the minimizers. □

Theorem 4.1 slightly relaxes previous degree conditions. In [Lau09, th. 6.18], the degree condition is $\deg(f) \le 2t + 2d_{\mathbf{g}} \le t + d_{\mathbf{g}} + d$. It also shows that the kernel of the moment matrix of a generic truncated moment sequence, $\mathrm{Ann}_{t+1}(\sigma^*)$, is the truncated vanishing ideal of the minimizers and that the relaxation is exact. This means that any element in $\mathcal{L}_{2d}^{\min}(\mathbf{g})$ truncated in any degree t is coming from a measure, provided $d \ge t$ is big enough.

A key ingredient in this analysis is Lemma 3.19. From Lemma 3.19 and the remark after it, the results of Theorem 4.1 hold true, if we replace the condition $\mathrm{rank}\, H_{\sigma^*}^t = \mathrm{rank}\, H_{\sigma^*}^{t+d_{\mathbf{g}}}$ with $\mathrm{rank}\, H_{\sigma^*}^t = \mathrm{rank}\, H_{\sigma^*}^{t+1}$ and d big enough.

We show in Example 4.2 that the condition $\mathrm{rank}\, H_{\sigma^*}^t = \mathrm{rank}\, H_{\sigma^*}^{t+1}$ is in general not sufficient to conclude that the points extracted from the moment matrix are inside the semialgebraic set. To the best of our knowledge, this is the first example where such a pathological behaviour is explicit.

Example 4.2. We consider the problem of minimizing $f = (1+X)(X-1)^2$ on $\mathcal{S}(1-X^2, -X^3) = [-1, 0]$. Notice that the SoS relaxation is exact, since $f^* = 0$ and:

$$(1+X)(X-1)^2 = \frac{1}{2}\Big((1+X)^2 + 1 - X^2\Big)(X-1)^2 \in \mathcal{Q}_4(1-X^2, -X^3).$$

This implies that $f_{\mathrm{SoS},2}^* = f_{\mathrm{MoM},2}^* = f^*$. The only minimizer of f on S is -1, and $\mathcal{I}(-1) = (X+1)$: therefore we would expect to get flat truncation at degree zero for a generic element, and in

particular $\operatorname{rank} H^0_{\sigma^*} = \operatorname{rank} H^1_{\sigma^*} = 1$. But this is not the case if we consider the MoM relaxation of order 2. Indeed an explicit calculation shows that $\sigma = \frac{1}{2}(\mathbf{e}^{[4]}_{-1} + \mathbf{e}^{[4]}_{1}) \in \mathcal{L}^{\min}_4(\mathbf{g})$, and $\operatorname{rank} H^1_\sigma = \operatorname{rank} H^2_\sigma = 2$. Therefore a generic $\sigma^* \in \mathcal{L}^{\min}_4(\mathbf{g})$ cannot satisfy the rank condition for $t = 0$. More precisely, it is possible to show that $\mathcal{L}^{\min}_4(\mathbf{g}) = \operatorname{conv}\left(\mathbf{e}^{[4]}_{-1}, \frac{1}{2}(\mathbf{e}^{[4]}_{-1} + \mathbf{e}^{[4]}_{1})\right)$. Therefore a generic $\sigma^* \in \mathcal{L}^{\min}_4(\mathbf{g})$ will also satisfy $\operatorname{rank} H^1_\sigma = \operatorname{rank} H^2_\sigma = 2$.

We confirm numerically the computation above, using the package `MomentTools.jl` to compute f^* and a generic $\sigma^* \in \mathcal{L}^{\min}_4(\mathbf{g})$: the moments that we obtain are

$$\sigma^*_0 = 0.9999999989784975, \quad \sigma^*_1 = -0.3530324749675295 \quad \sigma^*_2 = 0.9998474115299072$$
$$\sigma^*_3 = -0.3531851571450224 \quad \sigma^*_4 = 0.9996947364721432.$$

We compute the singular values of $H^0_{\sigma^*}$, $H^1_{\sigma^*}$ and $H^2_{\sigma^*}$ to have a numerically stable indication of the ranks:

Sing. Val. of $H^0_{\sigma^*}$: 0.9999999989784975

Sing. Val. of $H^1_{\sigma^*}$: $1.352956188465637, 0.6468912220427679$

Sing. Val. of $H^2_{\sigma^*}$: $2.2063794508570065, 0.7931627759613444, 7.983780245045715 \cdot 10^{-8}$

This confirms the theoretical description and shows that the rank condition is numerically satisfied for $t = 1$. The points extracted from the matrix are $\xi_1 \approx 0.9997640487211856$ and $\xi_2 \approx -1.0000000483192044$: notice that $\xi_1 \notin S$. This happens because the condition $\operatorname{rank} H^t_{\sigma^*} = \operatorname{rank} H^{t+d_{\mathbf{g}}}_{\sigma^*}$ is not satisfied (we cannot compute $H^{t+d_{\mathbf{g}}}_{\sigma^*} = H^3_{\sigma^*}$ as $3 = t + d_{\mathbf{g}} > d = 2$).

On the other hand, if we increase the order of the relaxation and compute $\sigma^* \in \mathcal{L}^{\min}_6(\mathbf{g})$ generic, we can verify flat truncation for $t = 0$ and the only point extracted is -1. Moreover notice, from Lemma 3.19 applied with $s = 1$ and the remark below, that it is enough to check $\operatorname{rank} H^0_{\sigma^*} = \operatorname{rank} H^1_{\sigma^*}$ to verify that $\operatorname{rank} H^t_{\sigma^*} = \operatorname{rank} H^{t+d_{\mathbf{g}}}_{\sigma^*}$, since the condition $0 = t \le d + s - \deg(\mathbf{g}) = 1$ is satisfied.

We have seen that flat truncation implies MoM exactness and a finite set of minimizers. We show now that, under the assumption of MoM finite convergence, flat truncation is equivalent to a zero dimensional support for the quadratic module $Q + (f - f^*)$ defining the minimizers.

We first need a technical lemma, that will be important to investigate the relationship between $\mathcal{L}^{\min}_{2d}(\mathbf{g})$ and $\mathcal{L}^{(1)}_{2d}(\mathbf{g}, \pm(f - f^*))$. Indeed, notice that $\mathcal{L}^{(1)}_{2d}(\mathbf{g}, \pm(f - f^*)) \subset \mathcal{L}^{\min}_{2d}(\mathbf{g})$, by definition, but the converse inclusion is not true in general, since for $\sigma \in \mathcal{L}^{\min}_{2d}(\mathbf{g})$ we only have $\langle \sigma | f \rangle = f^*$, and not $f - f^* \in \operatorname{Ann}_{d - \frac{\deg(f)}{2}}(\sigma)$.

Lemma 4.3. *Let $f \in \mathcal{Q}_{2k}(\mathbf{g})$, $\sigma \in \mathcal{L}_{2d}(\mathbf{g})$ and $t \in \mathbb{N}$ with $0 \le t \le d - k$. Then $\langle \sigma | f \rangle = 0$ implies for all $q \in \mathbb{R}[X]_t$, $\langle \sigma | qf \rangle = 0$. In other words, $f \in \operatorname{Ann}_t(\sigma)$.*

Proof. We set $g_0 = 1$ for notation convenience. Let $f = \sum_i s_i g_i = \sum_{i,j} p^2_{i,j} g_i \in \mathcal{Q}_{2k}(\mathbf{g})$, that is $\deg p^2_{i,j} g_i \le 2k$. We want to prove that for all $q \in \mathbb{R}[\mathbf{X}]$ such that $\deg(q) \le t$ we have $\langle \sigma | qf \rangle = 0$. In particular, it is enough to prove that:

$$\left\langle \sigma \middle| q p^2_{i,j} g_i \right\rangle = 0 \text{ for all } i, j \text{ and } q \in \mathbb{R}[\mathbf{X}]. \tag{5}$$

Now, notice that $\langle \sigma | f \rangle = 0$ implies $\left\langle \sigma \middle| p^2_{i,j} g_i \right\rangle = 0$ for all i, j, and consider for all $T \in \mathbb{R}$ and $h \in \mathbb{R}[\mathbf{X}]_{t + \deg p_{i,j}}$:

$$0 \le \left\langle \sigma \middle| (p_{i,j} - Th)^2 g_i \right\rangle = T^2 \left\langle \sigma \middle| h^2 g_i \right\rangle + 2T \left\langle \sigma \middle| h p_{i,j} g_i \right\rangle$$

(we can apply σ to $(p_{i,j} - Th)^2 g_i$ since $\deg(p_{i,j} - Th)^2 g_i) \le 2t + 2k \le 2d$). The polynomial $T \mapsto T^2 \left\langle \sigma \middle| h^2 g_i \right\rangle + 2T \left\langle \sigma \middle| h p_{i,j} g_i \right\rangle$ has therefore a double root at $T = 0$, and this implies $\left\langle \sigma \middle| h p_{i,j} g_i \right\rangle = 0$ for all $h \in \mathbb{R}[\mathbf{X}]_{t + \deg p_{i,j}}$. If we subsitute $h = q p_{i,j}$, we deduce eq. (5), and thus $f \in \operatorname{Ann}_t(\sigma)$. □

We can now prove the equivalence between the flat truncation and the zero dimensional support for the quadratic module $Q+(f-f^*)$ defining the minimizers.

Theorem 4.4. *Assume that we have MoM finite convergence. Then* $\dim \frac{\mathbb{R}[\mathbf{X}]}{\mathrm{supp}(Q+(f-f^*))} = 0$ *if and only if there exists d such that a generic $\sigma^* \in \mathcal{L}_{2d}^{\min}(\mathbf{g})$ has flat truncation.*

In particular, if $\rho = \rho(S^{\min})$, $D = \max(d_{\mathbf{g}}, \lceil \frac{\deg(f)}{2} \rceil)$ and $\delta \in \mathbb{N}$ is such that $f - f^* \in \overline{\mathcal{Q}_{2\delta}(\mathbf{g})}$, flat truncation happens for $\sigma^* \in \mathcal{L}_{2d}^{\min}(\mathbf{g})$ generic at degree $\rho - 1$ when d is such that:

(i) $(\sqrt[\mathbb{R}]{\mathrm{supp}\, Q(\mathbf{g})})_{2\delta+2\rho+2D-\deg(f)-2} \subset \overline{\mathcal{Q}_{2d}(\mathbf{g})}$;

(ii) $\mathcal{I}(S^{\min})_{2\rho+2D-2} \subset \overline{\mathcal{Q}_{2d}(\mathbf{g}) + (f-f^*)_{2d}}$;

(iii) $\delta + 2\rho + 2D - \deg(f) - 2 \leq d$.

Proof. Let us assume without loss of generality that $f^* = 0$.

We first show that flat truncation implies $\dim \frac{\mathbb{R}[\mathbf{X}]}{\mathrm{supp}(Q+(f))} = 0$. As in the proof of Theorem 4.1, if $\sigma^* \in \mathcal{L}_{2d}^{\min}(\mathbf{g})$ is generic satisfying flat truncation at degree t then $(\sigma^*)^{[2(t+d_{\mathbf{g}})]}$ is a generic element of $\mathcal{L}_{2d}(\mathbf{g}, \pm f)^{[2(t+d_{\mathbf{g}})]}$. Since the flat truncation property is satisfied, we conclude from Theorem 3.20 that $\sqrt[\mathbb{R}]{\mathrm{supp}(Q+(f))} = (\mathrm{Ann}_{t+1}(\sigma^*)) = \mathcal{I}(S^{\min})$ and finally, applying Lemma 1.2, $\dim \frac{\mathbb{R}[\mathbf{X}]}{\mathrm{supp}(Q+(f))} = \dim \frac{\mathbb{R}[\mathbf{X}]}{\mathcal{I}(S^{\min})} = 0$.

Conversely, if $\dim \frac{\mathbb{R}[\mathbf{X}]}{\mathrm{supp}(Q+(f))} = 0$, we deduce from Theorem 3.21 that the flat truncation property is satisfied for any $\sigma \in \mathcal{L}_{2d}(\mathbf{g}, \pm f)$ at degree $\rho - 1 = \rho(S(\mathbf{g}, \pm(f-f^*)) - 1 = \rho(S^{\min}) - 1$ for d such that $\mathcal{I}(S^{\min})_{2(\rho-1+D)} \subset \overline{\mathcal{Q}_{2d}(\mathbf{g}) + (f)_{2d}}$. Let $a = 2\rho - 2 + 2D$ and $\sigma^* \in \mathcal{L}_{2d}^{\min}(\mathbf{g})$ generic. We want to show that $(\sigma^*)^{[a]} \in \mathcal{L}_{2d}(\mathbf{g}, \pm f)^{[a]}$, so that we can conclude using Theorem 3.21. Since $\sigma^* \in \mathcal{L}_{2d}^{\min}(\mathbf{g}) \subset \mathcal{L}_{2d}(\mathbf{g})$, it is sufficient to prove that:

$$\langle \sigma^* | qf \rangle = 0 \text{ for all } q \text{ of degree} \leq a - \deg(f). \tag{6}$$

We prove now (6), starting from $\langle \sigma^* | f \rangle = f^* = 0$. MoM finite convergence implies that $\langle \sigma | f \rangle \geq 0$ for all $\sigma \in \mathcal{L}_{2d}(\mathbf{g})$, and therefore $f \in \mathcal{L}_{2d}(\mathbf{g})^\vee = \overline{\mathcal{Q}_{2d}(\mathbf{g})}$. Let $\delta \leq d$ be minimal such that $f \in \overline{\mathcal{Q}_{2\delta}(\mathbf{g})}$ and let $\mathbf{h} = h_1, \dots h_m$ be a graded basis of $\sqrt[\mathbb{R}]{\mathrm{supp}\, Q}$. From [Mar08, lemma 4.1.4] we deduce that $\mathcal{Q}_{2\delta}(\mathbf{g}) + (\mathbf{h})_{2\delta}$ is closed (as a subset of $\mathbb{R}[\mathbf{X}]_{2\delta}$ with the Euclidean topology), and therefore $\overline{\mathcal{Q}_{2\delta}(\mathbf{g})} \subset \mathcal{Q}_{2\delta}(\mathbf{g}) + (\mathbf{h})_{2\delta}$. Thus:

$$f = g + h = \sum_{i=0}^{s} s_i g_i + \sum_{i=1}^{m} p_i h_i \in \mathcal{Q}_{2\delta}(\mathbf{g}) + (\mathbf{h})_{2\delta},$$

where we set $g_0 = 1$ for notation convenience, $g = \sum_{i=0}^{s} s_i g_i \in \mathcal{Q}_{2\delta}(\mathbf{g})$ and $h = \sum_{i=1}^{m} p_i h_i \in (\mathbf{h})_{2\delta}$. It is then enough to prove that $\langle \sigma^* | qg \rangle = \langle \sigma^* | qh \rangle = 0$ where $\deg(qg) \leq b, \deg(qh) \leq b$ for $b = 2\delta + a - \deg(f) = 2\delta + 2\rho + 2D - \deg(f) - 2$.

We start by proving $\langle \sigma^* | qh \rangle = 0$. We deduce from lemma 3.6 that for d big enough we have $(\mathbf{h})_b \subset \overline{\mathcal{Q}_{2d}(\mathbf{g})}$ and $\mathcal{L}_{2d}(\mathbf{g})^{[b]} \subset \mathcal{L}_b(\pm \mathbf{h})$. Therefore

$$\langle \sigma^* | qh \rangle = \left\langle (\sigma^*)^{[b]} \middle| qh \right\rangle = 0.$$

Now we prove that $\langle \sigma^* | qg \rangle = 0$. Since $\delta + (a - \deg(f)) \leq d$, we can apply Lemma 4.3 with $g \in \mathcal{Q}_{2\delta}(\mathbf{g})$ and $t = a - \deg(f) \geq \deg(q)$, and conclude that $\langle \sigma^* | qg \rangle = 0$, as desired.

Therefore $\langle \sigma^* | qf \rangle = \langle \sigma^* | qg \rangle + \langle \sigma^* | qh \rangle = 0$ for all q of degree $\leq a - \deg(f)$ and (6) is satisfied. This implies that $(\sigma^*)^{[a]} \in \mathcal{L}_{2d}(\mathbf{g}, \pm f)^{[a]}$. Therefore we can apply Theorem 3.21 to conclude that the flat truncation property is satisfied for σ^*. □

Let us briefly comment the degree conditions in Theorem 4.4.

(i) If S has nonempty interior, it is not necessary to check the first condition, since in this case $\operatorname{supp} Q = 0$. More generally if the quadratic module is reduced, that is if $\sqrt[\mathbb{R}]{\operatorname{supp} Q} = \operatorname{supp} Q$, the first condition is automatically satisfied;

(ii) The second condition is the key one: it tells us that flat truncation happens when the ideal of the minimizers, truncated in the appropriate degree, can be described using the truncated quadratic module and the truncated ideal generated by $f - f^*$;

(iii) The third condition is technical, derived from Lemma 4.3. It allows to move from $\mathcal{L}_{2d}^{\min}(\mathbf{g})$ to $\mathcal{L}_{2d}(\mathbf{g}, \pm(f - f^*))$, where we can apply the results of the previous section.

We illustrate Theorem 4.4 in the following example, showing how it can help to predict the flat truncation degree.

Example 4.5. We continue Example 2.5. Notice that $f - f^* = X^2 \in Q_2 := \mathcal{Q}_2(\mathbf{g}) = \mathcal{Q}_2(1 - X^2 - Y^2, X + Y - 1)$ (i.e. the SoS relaxation is exact) and then the MoM relaxation has finite convergence. Using Theorem 4.4, we analyse if flat truncation holds at some degree. We have $\mathcal{I}(S^{\min}) = (X, Y - 1) \subset \sqrt[\mathbb{R}]{\operatorname{supp}(Q + (f - f^*))} = \sqrt[\mathbb{R}]{\operatorname{supp}(Q + (X^2))}$ where $Q := \mathcal{Q}(1 - X^2 - Y^2, X + Y - 1)$. Indeed:

$$
\begin{aligned}
X &= \frac{X^2 + (Y-1)^2}{2} + \frac{1 - X^2 - Y^2}{2} + X + Y - 1 \in Q_2 \subset \overline{Q_2 + (X^2)_2} \\
-X + \varepsilon &= \frac{\varepsilon}{2}\Big(1 - \frac{X^2}{\varepsilon^2} + (1 - \frac{X}{\varepsilon})^2\Big) \in Q_2 + (X^2)_2 \ \forall \varepsilon > 0 \Rightarrow -X \in \overline{Q_2 + (X^2)_2} \\
1 - Y &= \frac{1}{2}\Big(X^2 + (1-Y)^2 + 1 - X^2 - Y^2\Big) \in Q_2 \subset \overline{Q_2 + (X^2)_2} \\
Y - 1 &= X + Y - 1 - X \in Q_2 + \overline{Q_2 + (X^2)_2} = \overline{Q_2 + (X^2)_2}
\end{aligned}
$$

that implies $(X, Y - 1)_1 \subset \operatorname{supp}(\overline{Q_2 + (X^2)_2}) \subset \sqrt[\mathbb{R}]{\operatorname{supp}(Q + (f - f^*))}$ and thus $\dim \frac{\mathbb{R}[\mathbf{X}]}{\operatorname{supp}(Q + (X^2))} = 0$. Theorem 4.4 implies that flat truncation holds for a high enough order d of the MoM relaxation.

We investigate the degree conditions in Theorem 4.4 to prove that flat truncation happens for the MoM relaxation at order $d = 1$. We have $I(S^{\min}) = (X, Y - 1)$, $\rho = 1$, $d_{\mathbf{g}} = 1$, $\deg(f) = 2$, $D = 1$ and $\delta = 1$.

(i) As S has nonempty interior, $\operatorname{supp} Q = 0$ and the first point (i) is satisfied.

(ii) Notice that $2(\rho - 1 + D) = 2$, and therefore we have to show that $(X, Y - 1)_2 \subset \overline{Q_2 + (X^2)_2}$. Since we have shown above that $(X, Y - 1)_1 \subset \overline{Q_2 + (X^2)_2}$, it is enough to prove that $\pm X^2, \pm X(Y - 1), \pm(Y - 1)^2 \in \overline{Q_2 + (X^2)_2}$. Now, $\pm X^2, (Y - 1)^2 \in \overline{Q_2 + (X^2)_2}$ by definition. Finally:

$$
\begin{aligned}
-(Y-1)^2 &= 1 - Y^2 - X^2 + X^2 + 2(X + Y - 1) - 2X \in Q_2 + \overline{Q_2 + (X^2)_2} = \overline{Q_2 + (X^2)_2} \\
\pm X(Y-1) &= \frac{1}{2}\Big((\pm X + (Y-1))^2 - X^2 - (Y-1)^2\Big) \in \overline{Q_2 + (X^2)_2},
\end{aligned}
$$

concluding the proof of the second point (ii).

(iii) We have $1 = \delta + 2\rho + 2D - \deg(f) - 2 \leq d = 1$, and thus the third point (iii) is satisfied.

Therefore flat truncation happens at degree $\rho - 1 = 0$ for the MoM relaxation at order $d = 1$.

Related properties have been previously investigated. It is shown in [Nie13b, th. 2.2] that, under genericity assumptions, if for an order d big enough we have $f^*_{\mathrm{SoS},d} = f^*_{\mathrm{MoM},d}$ (strong duality) and sup = max in the definition of $f^*_{\mathrm{SoS},d}$, then there is finite convergence (that is $f^*_{\mathrm{MoM},d} = f^*$) if and only if flat truncation is satisfied for every $\sigma \in \mathcal{L}_{2d}^{\min}(\mathbf{g})$ (or, equivalently, if it is satisfied for $\sigma^* \in \mathcal{L}_{2d}^{\min}(\mathbf{g})$ generic). Theorem 4.4 applies for different cases, for instance when there is finite convergence but the SoS relaxation is not exact (see example 2.13). This is possible since our analysis investigates the closure of the quadratic modules we are considering. Furthermore, under genericity assumption,

as a corollary of Theorem 4.4 we will show (in Theorem 4.7 and Corollary 4.8) that we have finite convergence, the SoS relaxation and MoM relaxation are exact and the flat truncation property is satisfied.

Another improvement made is the estimation of the order d of the relaxation that is sufficient to have flat truncation, answering a question in [Nie13b]. To the best of out knowledge, this is is the first result in this direction. These conditions depends on properties of the minimizers and the quadratic module $\mathcal{Q}_{2d}(\mathbf{g})$ that might be difficult to check a priori. However they may be analyzed in some specific cases, such as optimization problems with a single minimizer, to deduce more precise bounds.

4.2 Boundary Hessian Conditions

In this section, we show that if regularity conditions, known as Boundary Hessian Conditions (BHC), are satisfied, then the flat truncation property holds. These are conditions on the minimizers of a polynomial f on a basic semialgebraic set S introduced by Marshall in [Mar06] and [Mar09], which are particular cases of the so called local-global principle. Under these conditions, global properties of polynomials (e.g. $f \in Q$) can be deduced from local properties (e.g. checking the BHC at the minimizers of f on $\mathcal{S}(Q)$). We refer to [Sch05a], [Sch06] and [Mar08, ch. 9] for more details. We introduce BHC conditions following [Nie14].

Definition 4.6 (Boundary Hessian Conditions)**.** Consider a POP with inequality constraints $\mathbf{g} = g_1, \dots, g_r$, equality constraints $\mathbf{h} = h_1, \dots, h_s$ and objective function f. Let $V = \mathcal{V}(\mathbf{h}) \subset \mathbb{R}^n$ and suppose that $Q = \mathcal{Q}(\mathbf{g}, \pm\mathbf{h})$ is Archimedean. We say that the *Boundary Hessian Conditions* hold at a minimizer point $\xi \in S(\mathbf{g}, \pm\mathbf{h})$ of f if ξ is a smooth point of V and:

(i) we can choose $g_{i_1} = t_1, \dots, g_{i_k} = t_k$ that are part of a regular system of parameters $t_1, \dots, t_m$, $m \geq k$, for V at ξ and for some neighbourhood U of ξ we have $\mathcal{S}(g_{i_1}, \dots, g_{i_k}, \pm\mathbf{h}) \cap U = \mathcal{S}(\mathbf{g}, \pm\mathbf{h}) \cap U$;

(ii) On V, locally at ξ we have that $\nabla f = a_1 \nabla t_1 + \cdots + a_m \nabla t_m$, where a_i are strictly positive real numbers;

(iii) On V, locally at ξ we have that $\mathrm{Hess}(f)(0, \dots, 0, t_{k+1}, \dots t_m)$ is positive definite in $t_{k+1}, \dots t_m$.

These conditions are related to standard conditions in optimization at a point $\xi \in S$ (see e.g. [Ber99]). Hereafter, the active constraints at $\xi \in S$ are the constraints $g_{i_1}, \dots, g_{i_m}$ such that $g_{i_j}(\xi) = 0$. To simplify the description of these conditions, we consider a constraint $\pm g(x) \geq 0$ as a single (equality) constraint. Therefore an equality constraint defining the set S is an active constraint at a point $\xi \in S$.

- *Constraint Qualification Condition* (CQC): for the active constraints $g_{i_1}, \dots, g_{i_m}$ at ξ, the gradients $\nabla g_{i_1}(\xi), \dots, \nabla g_{i_m}(\xi)$ are linearly independent.
- *Strict Complementary Condition* (SCC): for the active constraints $g_{i_1}, \dots, g_{i_m}$ at ξ, there exist $a_1, \dots, a_m \in \mathbb{R}$ with $a_j > 0$ if g_{i_j} is not an equality constraint such that $\nabla f(\xi) = a_1 \nabla g_{i_1}(\xi) + \cdots + a_m \nabla g_{i_m}(\xi)$.
- *Second Order Sufficiency Condition* (SOSC): for $L(x) = f(x) - \sum_{j=1}^m a_j g_{i_j}$ with $a_i > 0$ if $g_{i_j}(x)$ is not an equality constraint, we have $\forall v \in \langle \nabla g_{i_1}(\xi), \dots, \nabla g_{i_m}(\xi) \rangle^\perp$, $v \neq 0$, $v^t \nabla^2 L(\xi) v > 0$.

If these conditions are satisfied at every minimizer ξ, then the BHC conditions are satisfied with the active sign constraints at ξ as regular parameters $t_1 = g_{i_1}, \dots, t_k = g_{i_k}$, see [Nie14].

Notice that when BHC hold, the minimizers are non-singular, isolated points and thus finite. It is proved in [Mar06] that if BHC holds at every minimizer of f on $\mathcal{S}(\mathbf{g})$ then $f - f^* \in \mathcal{Q}(\mathbf{g})$, which implies that the SoS relaxation is exact. [Nie14] proved that the BHC at every minimizer of f, which hold generically, implies the SoS finite convergence property.

In this section, we prove that, if the BHC hold, then the flat truncation property holds.

Theorem 4.7. *Let $f \in \mathbb{R}[\mathbf{X}]$, $Q = \mathcal{Q}(\mathbf{g})$ be an Archimedean finitely generated quadratic module and assume that the BHC hold at every minimizer of f on $S = \mathcal{S}(\mathbf{g})$. Then the SoS relaxation is exact, the MoM relaxation has finite convergence and the flat truncation holds for $\sigma^* \in \mathcal{L}_{2d}^{\min}(\mathbf{g})$ generic when d is big enough. If conditions (i)-(iii) in Theorem 4.4 are satisfied for the relaxation order d, then the flat truncation property holds.*

Proof. If BHC hold at every minimizer of f on $\mathcal{S}(\mathbf{g})$ then $S^{\min}$ is finite and $f - f^* \in \mathcal{Q}(\mathbf{g})$ (see [Mar06]), which implies that the SoS relaxation is exact and thus the MoM relaxation has finite convergence. Moreover if the BHC conditions hold at every minimizer of f on S, then $\dim \frac{\mathbb{R}[\mathbf{X}]}{\operatorname{supp}(Q+(f-f^*))} = 0$ (see the proof of [Mar06, th. 2.3], where it is shown that the field of fractions of $\mathbb{R}[\mathbf{X}]$ modulo any minimal prime ideal lying over $\operatorname{supp}(Q+(f-f^*))$ is isomorphic to $\mathbb{R}$, that implies $\dim \frac{\mathbb{R}[\mathbf{X}]}{\operatorname{supp}(Q+(f-f^*))} = 0$). Then we conclude applying Theorem 4.4. □

We show now that flat truncation and moment exactness hold *generically*. For polynomials $f \in \mathbb{R}[\mathbf{X}]_d$ and $g_1 \in \mathbb{R}[\mathbf{X}]_{d_1}, \dots, g_s \in \mathbb{R}[\mathbf{X}]_{d_s}$, we say that a property holds generically (or that the property holds for generic $f, g_1, \dots, g_s$) if there exists finitely many nonzero polynomials $\phi_1, \dots, \phi_l$ in the coefficients of polynomials in $\mathbb{R}[\mathbf{X}]_d$ and $\mathbb{R}[\mathbf{X}]_{d_1}, \dots, \mathbb{R}[\mathbf{X}]_{d_s}$ such that, when $\phi_1(f, \mathbf{g}) \neq 0, \dots, \phi_l(f, \mathbf{g}) \neq 0$, the property holds.

Corollary 4.8. *For $f \in \mathbb{R}[\mathbf{X}]_d$ and $g_1 \in \mathbb{R}[\mathbf{X}]_{d_1}, \dots, g_s \in \mathbb{R}[\mathbf{X}]_{d_s}$ generic satisfying the Archimedean condition, the SoS relaxation is exact, the MoM relaxation has finite convergence and the flat truncation holds for $\sigma^* \in \mathcal{L}_{2d}^{\min}(\mathbf{g})$ generic when d is big enough. If conditions (i)-(iii) in Theorem 4.4 are satisfied for the relaxation order d, then the flat truncation holds.*

Proof. By [Nie14, th. 1.2] BHC hold generically. We apply Theorem 4.7 to conclude. □

Here is an example where BHC holds.

Example 4.9 (Robinson form)**.** We find the minimizers of Robinson form $f = x^6 + y^6 + z^6 + 3x^2y^2z^2 - x^4(y^2+z^2) - y^4(x^2+z^2) - z^4(x^2+y^2)$ on the unit sphere $h = x^2 + y^2 + z^2 - 1$. The Robinson polynomial has minimum $f^* = 0$ on the unit sphere, and the minimizers on $\mathcal{V}_{\mathbb{R}}(h)$ are:

$$\frac{\sqrt{3}}{3}(\pm1, \pm1, \pm1), \frac{\sqrt{2}}{2}(0, \pm1, \pm1), \frac{\sqrt{2}}{2}(\pm1, 0, \pm1), \frac{\sqrt{2}}{2}(\pm1, \pm1, 0).$$

BHC are satisfied at every minimizer (see [Nie14, ex. 3.2]), flat truncation holds and we can recover the minimizers from Theorem 4.7. We estimate the bounds of Theorem 4.4 and compare with the numerical experiments. It is not necessary to check (i), since $(h) = \sqrt[\mathbb{R}]{\operatorname{supp} \mathcal{Q}(\pm h)}$. For the point (ii), we estimate the regularity of the minimizers as the regularity of twenty generic points on a sphere, that is $\rho = 5$. Then $2\rho + 2D - 2 = 14$, and thus we expect flat truncation for $d \geq 7$. For the point (iii), we need to have $d \geq \delta + 2\rho + 2D - \deg(f) - 2 \geq 3 + 10 + 6 - 6 - 2 = 11$. However, in practice for this example we have flat truncation numerically at order 6 and not before (using the SDP solver SDPA). We recover a good approximation of the minimizers at this order:

```
v, M = minimize(f, [h], [], X, 6)
w, Xi = get_measure(M)
```

Here $f^*_{\mathrm{MoM},6} \approx v = -1.27211 \cdot 10^{-7}$ and the minimizers with positive coordinates are (all the twenty minimizers are found):

	ξ_1	ξ_2	ξ_3	ξ_4
x	0.577351068999	$8.812477930640\,10^{-12}$	0.707107158043	0.707107157553
y	0.577351069076	0.707107158048	$1.271729446125\,10^{-13}$	0.707107157555
z	0.577351066102	0.707107158048	0.707107158042	$2.478771201340\,10^{-9}$

4.3 Finite semialgebraic set

In this section we consider the case when $S = \mathcal{S}(\mathbf{g}) = \{\xi_1, \dots, \xi_r\} \subset \mathbb{R}^n$ is non-empty and finite.

Theorem 4.10. *Let $Q = \mathcal{Q}(\mathbf{g})$ and suppose that $\dim \frac{\mathbb{R}[\mathbf{X}]}{\operatorname{supp} Q} = 0$. Then S is finite, the MoM relaxation has finite convergence and the flat truncation holds for $\sigma^* \in \mathcal{L}_{2d}^{\min}(\mathbf{g})$ generic when d is big enough. If conditions (i)-(iii) in Theorem 4.4 are satisfied, then flat truncation holds at the relaxation order d.*

Proof. Since $\dim \frac{\mathbb{R}[\mathbf{X}]}{\operatorname{supp} Q} = 0$, we deduce that S is finite and we have MoM finite convergence from Theorem 3.21, Theorem 3.20 and Theorem 4.1. Indeed if d is big enough then flat truncation is satisfied for any $\sigma \in \mathcal{L}_{2d}(\mathbf{g})$, in particular for $\sigma^* \in \mathcal{L}_{2d}^{\min}(\mathbf{g})$ generic. We conclude applying Theorem 4.4. □

As corollaries, we have that the conclusions of Theorem 4.10 hold:

- for the moment relaxation $\mathcal{L}_{2d}(\Pi\mathbf{g})$ when $S = \mathcal{S}(\mathbf{g}) = \mathcal{S}(\Pi\mathbf{g})$ is finite, since by the real Nullstellensatz,

$$\dim \frac{\mathbb{R}[\mathbf{X}]}{\operatorname{supp} Q(\Pi\mathbf{g})} = \dim \frac{\mathbb{R}[\mathbf{X}]}{\operatorname{supp} \mathcal{O}(\mathbf{g})} = \dim \frac{\mathbb{R}[\mathbf{X}]}{\sqrt{\operatorname{supp} \mathcal{O}(\mathbf{g})}} = \dim \frac{\mathbb{R}[\mathbf{X}]}{\mathcal{I}(\mathcal{S}(\mathbf{g}))} = 0.$$

  See [Nie13c, th. 4.1] and [LLR08, rem. 4.9].

- for the moment relaxation $\mathcal{L}_d(\mathbf{g}, \pm\mathbf{h})$ when $\mathcal{V}_{\mathbb{R}}(\mathbf{h})$ is finite, since for $Q = Q(\mathbf{g}, \pm\mathbf{h})$,

$$\dim \frac{\mathbb{R}[\mathbf{X}]}{\operatorname{supp} Q} = \dim \frac{\mathbb{R}[\mathbf{X}]}{\sqrt{\operatorname{supp} Q}} = \dim \frac{\mathbb{R}[\mathbf{X}]}{\sqrt[\mathbb{R}]{\operatorname{supp} Q}} \leq \dim \frac{\mathbb{R}[\mathbf{X}]}{\sqrt[\mathbb{R}]{(\mathbf{h})}} = 0.$$

  See [Nie13c, th. 1.1] and [LLR08]. This includes Polynomial Optimization problems with binary variables and equations of the form $X_i^2 - X_i = 0$, for which MOM relaxations are of particular interest, see e.g. [Lau03].

Notice that, even if the SoS relaxation has the finite convergence property and the MoM relaxation is exact, it may not be SoS exact for a finite real variety, as shown in Example 2.12 and Example 2.13.

Example 4.11 (Gradient ideal)**.** We compute the minimizers of Example 2.13. Let $f = (X^4Y^2 + X^2Y^4 + Z^6 - 2X^2Y^2Z^2) + X^8 + Y^8 + Z^8 \in \mathbb{R}[X, Y, Z]$. We want to minimize f over the gradient variety $\mathcal{V}_{\mathbb{R}}\left(\frac{\partial f}{\partial X}, \frac{\partial f}{\partial Y}, \frac{\partial f}{\partial Z}\right)$ with $\dim \frac{\mathbb{R}[\mathbf{X}]}{(\frac{\partial f}{\partial X}, \frac{\partial f}{\partial Y}, \frac{\partial f}{\partial Z})} = 0$. By Theorem 4.10, we deduce that flat truncation holds for an order of relaxation d high enough. In this example, we have $\rho = 1$, $D = 4$, $\deg(f) = 8$, $\delta \geq 4$, so that we expect flat truncation at an order $d \geq 4$, from Theorem 4.4.

```
v, M = minimize(f, differentiate(f,X), [], X, 4)
w, Xi = get_measure(M, 2.e-2)
```

The approximation of the minimum $f^* = 0$ is $v = -1.6279 \cdot 10^{-9}$, and the decomposition with a threshold of $2 \cdot 10^{-2}$ gives the following numerical approximation of the minimizer (the origin):

$$\xi = (2.976731510689691\ 10^{-17}; -9.515032317137384\ 10^{-19}; 3.763401209219283\ 10^{-18}).$$

4.4 Gradient, KKT and Polar ideals

Another approach which has been investigated to make the relaxations exact, is to add equality constraints satisfied by the minimizers (and independent of the minimum f^*) to a Polynomial Optimization Program.

For global optimization we can consider the gradient equations (see [NDS06]): obviously $\nabla f(x^*) = \mathbf{0}$ for all the minimizers x^* of f on $S = \mathbb{R}^n$. For constrained optimization we can consider Karush–Kuhn–Tucker (KKT) constraints, adding new variables (see [DNP07]) or projecting them to the variables $\mathbf{X}$ (Jacobian equations, see [Nie13a]). We shortly describe them.

Let $g_1,\dots,g_r,h_1,\dots,h_s \in \mathbb{R}[\mathbf{X}]$ defining $S = \mathcal{S}(\mathbf{g},\pm\mathbf{h})$, and let $f \in \mathbb{R}[\mathbf{X}]$ be the objective function. Let $\mathbf{\Lambda} = (\Lambda_1,\dots,\Lambda_r)$ and $\mathbf{\Gamma} = (\Gamma_1,\dots,\Gamma_s)$ be variables representing the *Lagrange multipliers* associated with $\mathbf{g}$ and $\mathbf{h}$. The *KKT constraints* associated to the optimization problem $\min f(x)\colon x \in \mathcal{S}(\mathbf{g},\pm\mathbf{h})$ are:

$$\begin{cases} \dfrac{\partial f}{\partial X_i} - \displaystyle\sum_{k=1}^{r} \Lambda_k^2 \frac{\partial g_k}{\partial X_i} - \sum_{j=1}^{s} \Gamma_j \frac{\partial h_j}{\partial X_i} = 0 & \forall i \\ \Lambda_k g_k = 0, \quad h_j = 0, \quad g_k \geq 0 & \forall j,k, \end{cases} \tag{7}$$

where the polynomials belong to $\mathbb{R}[\mathbf{X},\mathbf{\Gamma},\mathbf{\Lambda}]$. These are sufficient but not necessary conditions for $x^* \in S$ being a minimizer.

For $x \in S$, we say that g_i is an *active constraint at* x if $g_i(x) = 0$. Let $x^* \in S$ and $g_{i_1},\dots g_{i_k}$ be the active constraints at x^*. The KKT constraints are necessary if the Constraint Qualification Condition (CQC) holds, that is, if $\nabla h_1(x^*),\dots,\nabla h_s(x^*),\nabla g_{i_1}(x^*),\dots,\nabla g_{i_k}(x^*)$ are linearly independent at the minimizer $x^* \in S$ (also called Linear Independence Constraint Qualification in [NW06, th. 12.1]). We cannot avoid the CQC hypothesis: for example if $f = X_1 \in \mathbb{R}[X_1]$ and $g_1 = X_1^3 \in \mathbb{R}[X_1]$, then $x^* = 0$ is a minimizer, but the KKT equations are not satisfied at $x^* = 0$. To avoid this problem we define the *polar ideal*. Observe from eq. (7) that, if KKT constraints are satisfied at x and

- if g_i is not an active constraint at x, then $\Lambda_i = 0$;
- if $g_{i_1},\dots g_{i_k}$ are the active constraints at x, then the gradients $\nabla f(x),\nabla h_1(x),\dots,\nabla h_s(x),\nabla g_{i_1}(x),\dots,\nabla g_{i_k}(x)$ are linearly dependent.

Definition 4.12. For $f,g_1,\dots,g_r,h_1,\dots,h_s \in \mathbb{R}[\mathbf{X}]$ as before, the *polar ideal* is defined as follows:

$$J := (\mathbf{h}) + \sum_{A=\{a_1,\dots,a_k\}\subset\{1,\dots r\}} \Big(\operatorname{rank}\operatorname{Jac}(f,\mathbf{h},g_{a_1},\dots,g_{a_k}) < s+k+1\Big) \prod_{b\notin A} g_b.$$

where $\Big(\operatorname{rank}\operatorname{Jac}(f,\mathbf{h},g_{a_1},\dots,g_{a_k}) < l\Big)$ is the ideal generated by the $l \times l$ minors of the Jacobian matrix $\operatorname{Jac}(f,\mathbf{h},g_{a_1},\dots,g_{a_k})$.

We could replace the generators of the ideal in this definition by polynomials defining the same variety. This variety, known also as Jacobian or augmented Jacobian variety, coincides with the one defined by $h_1,\dots,h_{m_1},\varphi_i,\dots,\varphi_r$ in [Nie13a].

The improvement that we make from the KKT constraints is to consider conditions that are necessary for being a minimizer, similar to Fritz John Optimality Conditions (see [Ber99, sec. 3.3.5]). Indeed we prove in the next lemma that every minimizer belongs to $\mathcal{V}_{\mathbb{R}}(J)$.

Lemma 4.13. *Let x^* be a minimizer of f on $S = \mathcal{S}(\mathbf{g},\pm\mathbf{h})$. Then $x^* \in \mathcal{V}_{\mathbb{R}}(J)$.*

Proof. Since $x^* \in S$, then $x^* \in \mathcal{V}_{\mathbb{R}}(\mathbf{h})$.

If the CQC hold at x^*, then x^* is a KKT point (see [NW06, th. 12.1]) and $\nabla f(x) = \sum_j \gamma_j \nabla \mathbf{h}_j(x) + \sum_j \lambda_j^2 \nabla \mathbf{g}_j(x)$ for some γ_j and λ_i in $\mathbb{R}$. As $\lambda_k = 0$ if g_k is not an active constraint, we have that

$$\nabla f(x^*),\nabla h_1(x^*),\dots,\nabla h_r(x^*),\nabla g_{i_1}({}^*x),\dots,\nabla g_{i_k}(x^*)$$

are linearly dependent, where $g_{i_1},\dots g_{i_k}$ are the active constraints at x^*. Thus

$$\operatorname{rank}\operatorname{Jac}(f(x^*),\mathbf{h}(x^*),g_{a_1}(x^*),\dots,g_{a_k}(x^*)) < s+k+1 \text{ if } \{i_1,\dots,i_k\} \subset \{a_1,\dots,a_k\}.$$

On the other hand, if $i_j \notin \{a_1,\dots,a_k\}$ then $g_{i_j}(x^*) = 0$. This implies $x^* \in \mathcal{V}_{\mathbb{R}}(J)$.

If the CQC do not hold at x^* and $g_{i_1},\dots,g_{i_k}$ are the active constraints, then the gradients $\nabla h_1(x^*),\dots,\nabla h_s(x^*)$ and $\nabla g_{i_1}(x^*),\dots,\nabla g_{i_k}(x^*)$ are linearly dependent. This implies that $\nabla f(x^*)$, $\nabla h_1(x^*)$, $\dots,\nabla h_s(x^*)$ and $\nabla g_{i_1}(x^*),\dots,\nabla g_{i_k}(x^*)$ are also linearly dependent, and we conclude as in the previous case. □

Theorem 4.14. *Let $Q = \mathcal{Q}(\mathbf{g}, \pm\mathbf{h})$ and $J = (\mathbf{h}')$ be the polar ideal, where $\mathbf{h}' \subset \mathbb{R}[\mathbf{X}]$ is a finite set of generators. If $\dim \frac{\mathbb{R}[\mathbf{X}]}{\mathrm{supp}(\mathcal{Q}(\mathbf{g})+(\mathbf{h}'))} = 0$ then MoM relaxation $(\mathcal{L}_{2d}(\mathbf{g}, \pm\mathbf{h}'))_{d\in\mathbb{N}}$ has finite convergence and the flat truncation holds for $\sigma^* \in \mathcal{L}_{2d}^{\min}(\mathbf{g}, \pm\mathbf{h}')$ generic when d is big enough. In particular this holds when $\mathcal{V}_{\mathbb{R}}(J)$ is finite.*

Proof. Minimizers belongs to $\mathcal{V}_{\mathbb{R}}(J)$ by Lemma 4.13. Then MoM exactness follows from Theorem 4.10. □

The assumption in [NDS06], [DNP07] and [Nie13a] for finite convergence and SoS exactness are smoothness conditions or radicality assumptions on the associated complex variety. In particular, Assumption 2.2 in [Nie13a, th. 2.3] requires the varieties defined by the active constraints to be non-singular to conclude finite convergence of the hierarchy. Our condition for finite convergence and flat truncation is of a different nature, since it is on the finiteness of the real polar variety. For instance we can apply Theorem 4.14 in Example 4.15, but Assumption 2.2 in [Nie13a] is not satisfied, since the minimizer is a singular point. Moreover notice that in our theorem we use only the defining inequalities $\mathbf{g}$ and not their products $\Pi\mathbf{g}$, as done in [Nie13a, th. 2.3] (in other words, we only need the quadratic module and not the preordering).

In the following example, BHC are not satisfied. But adding the polar constraints, we obtain an exact relaxation with the flat truncation property.

Example 4.15 (Singular minimizer)**.** We minimize $f = X$ on the compact semialgebraic set $S = \mathcal{S}(X^3 - Y^2, 1 - X^2 - Y^2)$. We have $f^* = 0$ and the only minimizer is the origin, which is a singular point of the boundary of S. Thus BHC do not hold, and we cannot apply Theorem 4.7. We have $\dim \frac{\mathbb{R}[\mathbf{X}]}{\mathrm{supp}(Q+(X))} = 0$ since $\mathrm{supp}(Q+(X)) \supset (X, Y^2)$, but we cannot apply Theorem 4.4, as we don't have finite convergence of the SoS and MoM relaxations. Indeed $X \notin Q = \mathcal{Q}(X^3 - Y^2, 1 - X^2 - Y^2)$, since $X \notin \mathcal{Q}(X^3, 1 - X^2)$. This implies that the SoS and MoM relaxations do not have finite convergence, following Example 2.8. This example also shows that we cannot remove the hypothesis of MoM finite convergence in Theorem 4.4.

To get flat truncation, we add the polar equations, that define a finite real polar variety, as we show in the following. First notice that, since $\mathcal{V}(X^3 - Y^2)$ is singular, Assumption 2.2 in [Nie13a] is not satisfied and the finite convergence of the relaxation $\mathcal{O}_{2d}(\mathbf{g}, \pm\mathbf{h}')$ using the polar variety cannot be deduced from [Nie13a, th. 2.3]. The generators of the polar variety are $\mathbf{h}' = (1 - X^2 - Y^2)(X^3 - Y^2)$, $Y(1 - X^2 - Y^2)$, $Y(X^3 - Y^2)$. The real roots are $(-1, 0)$, $(1, 0)$, $(0, 0)$ and the two real intersections of $1 - X^2 - Y^2 = 0$ and $X^3 - Y^2 = 0$. Therefore $\dim \frac{\mathbb{R}[\mathbf{X}]}{\mathrm{supp}(Q+(\mathbf{h}'))} \leq \dim \frac{\mathbb{R}[\mathbf{X}]}{\sqrt[\mathbb{R}]{(\mathbf{h}')}} = 0$, and Theorem 4.14 implies flat truncation. We recover the minimizer considering the MoM relaxation of order 5:

```
v, M = polar_minimize(f, [], [x^3-y^2,1-x^2-y^2], X, 5)
w, Xi = get_measure(M, 2.e-3)
```

The approximation of the minimum $f^* = 0$ is $v = -0.0045$, and the decomposition of the moment sequence with a threshold of $2 \cdot 10^{-3}$ gives the following approximation of the minimizer (the origin):

$$\xi = (-0.004514367348787526, 2.1341684460860045\ 10^{-21}).$$

The error of approximation on ξ is of the same order than the error on the minimum f^*.

5 Conclusion

We investigated the convex cones $\mathcal{L}_d(\mathbf{g})$ dual to the truncated quadratic modules $\mathcal{Q}_d(\mathbf{g})$ from a new perspective. We studied the kernels of moment matrices or annihilators of moment sequences in these cones and characterize the ideal they generate (Theorem 3.12). We focused on the zero dimensional case and its relationships with the flat truncation property (Theorem 3.20 and Theorem 3.21), that can be used to certify that a linear functional is coming from a measure.

The main contributions of the paper are the applications of the previous analysis to flat truncation in Lasserre's MoM relaxation for Polynomial Optimization. We studied the flat truncation property in this context (Theorem 4.1) and deduced new necessary and sufficient conditions for flat truncation (Theorem 4.4). These conditions can be used to show that, under regularity and thus genericity assumptions (Boundary Hessian Conditions), the flat truncation property is satisfied (Theorem 4.7, Corollary 4.8). We applied these results to Polynomial Optimization on finite sets (Theorem 4.10) and for singular cases, adding polar equations, to obtain flat truncation (Theorem 4.14).

Theorem 4.4 provides the first known degree bounds for the flat truncation property to hold, in terms of the inequalities $\mathbf{g}$ and the objective function f (in particular depending on the regularity of the minimizers). An interesting question would be to investigate if it is possible to improve these degree bounds. Another possible research direction is to investigate regularity conditions, simpler than Boundary Hessian Conditions, that imply flat truncation for MoM relaxation of a certain order d.

Acknowledgments. The authors thank D. Henrion and M. Laurent for useful discussions about the outer approximation of semialgebraic sets with moments of degree one and flat truncation properties, and the anonymous referees for their helpful suggestions for the improvement of the results and the presentation of the paper.

This work has been partially supported by European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Actions, grant agreement 813211 (POEMA).

References

[ABM15] Marta Abril Bucero and Bernard Mourrain. "Border Basis relaxation for polynomial optimization". *Journal of Symbolic Computation* 74 (2015), pp. 378–399.

[BCR98] Jacek Bochnak, Michel Coste, and Marie-Francoise Roy. *Real Algebraic Geometry*. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics. Berlin Heidelberg: Springer-Verlag, 1998. ISBN: 978-3-540-64663-1.

[Ber99] Dimitri P. Bertsekas. *Nonlinear Programming*. en. Athena Scientific, 1999. ISBN: 978-1-886529-00-7.

[BS87] David Bayer and Michael Stillman. "A criterion for detectingm-regularity". *Inventiones Mathematicae* 87.1 (1987), pp. 1–11.

[CF96] Raul E. Curto and Lawrence A. Fialkow. *Solution of the Truncated Complex Moment Problem for Flat Data*. Providence, R.I: Amer Mathematical Society, 1996. ISBN: 978-0-8218-0485-8.

[CF98] Raúl E. Curto and Lawrence A. Fialkow. *Flat Extensions of Positive Moment Matrices: Recursively Generated Relations*. American Mathematical Soc., 1998. 73 pp. ISBN: 978-0-8218-0869-6.

[CLO15] David A. Cox, John B. Little, and Donal O'Shea. *Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra*. Fourth edition. Undergraduate texts in mathematics. Cham Heidelberg New York Dordrecht London: Springer, 2015. ISBN: 978-3-319-16720-6 978-3-319-16721-3.

[DNP07] James Demmel, Jiawang Nie, and Victoria Powers. "Representations of positive polynomials on noncompact semialgebraic sets via KKT ideals". *Journal of Pure and Applied Algebra* 209.1 (2007), pp. 189–200.

[Eis05] David Eisenbud. *The Geometry of Syzygies: A Second Course in Algebraic Geometry and Commutative Algebra*. Graduate Texts in Mathematics. New York: Springer-Verlag, 2005. ISBN: 978-0-387-22215-8.

[HK14] Didier Henrion and Milan Korda. "Convex Computation of the Region of Attraction of Polynomial Control Systems". *IEEE Transactions on Automatic Control* 59.2 (2014), pp. 297–312.

[HL05] Didier Henrion and Jean bernard Lasserre. "Detecting global optimality and extracting solutions in GloptiPoly". *Chapter in D. Henrion, A. Garulli (Editors). Positive polynomials in control. Lecture Notes in Control and Information Sciences*. Springer Verlag, 2005.

[JH16] Cédric Josz and Didier Henrion. "Strong duality in Lasserre's hierarchy for polynomial optimization". *Optimization Letters* 10.1 (2016), pp. 3–10.

[KS19] Tom-Lukas Kriel and Markus Schweighofer. "On the Exactness of Lasserre Relaxations and Pure States Over Real Closed Fields". *Foundations of Computational Mathematics* 19.6 (2019), pp. 1223–1263.

[Las01] Jean B. Lasserre. "Global Optimization with Polynomials and the Problem of Moments". *SIAM Journal on Optimization* 11.3 (2001), pp. 796–817.

[Las10] Jean-Bernard Lasserre. *Moments, positive polynomials and their applications*. Imperial College Press optimization series v. 1. London : Signapore ; Hackensack, NJ: Imperial College Press ; Distributed by World Scientific Publishing Co, 2010. ISBN: 978-1-84816-445-1.

[Las+13] Jean-Bernard Lasserre, Monique Laurent, Bernard Mourrain, Philipp Rostalski, and Philippe Trébuchet. "Moment matrices, border bases and real radical computation". *Journal of Symbolic Computation* 51 (2013), pp. 63–85.

[Las15] Jean Bernard Lasserre. *An Introduction to Polynomial and Semi-Algebraic Optimization*. Cambridge: Cambridge University Press, 2015. ISBN: 978-1-107-44722-6.

[Lau03] Monique Laurent. "A Comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre Relaxations for 0–1 Programming". *Mathematics of Operations Research* 28.3 (2003), pp. 470–496.

[Lau07] Monique Laurent. "Semidefinite representations for finite varieties". *Mathematical Programming* 109.1 (2007), pp. 1–26.

[Lau09] Monique Laurent. "Sums of squares, moment matrices and optimization over polynomials". *Emerging applications of algebraic geometry*. Vol. 149. IMA Volumes in Mathematics and Its Applications. Springer, 2009, pp. 157–270.

[LLR08] Jean Bernard Lasserre, Monique Laurent, and Philipp Rostalski. "Semidefinite Characterization and Computation of Zero-Dimensional Real Radical Ideals". *Foundations of Computational Mathematics* 8.5 (2008), pp. 607–647.

[LM09] Monique Laurent and Bernard Mourrain. "A Generalized Flat Extension Theorem for Moment Matrices". *Archiv der Mathematik* 93.1 (2009), pp. 87–98.

[LV21] Monique Laurent and Luis Felipe Vargas. *Finite convergence of sum-of-squares hierarchies for the stability number of a graph*. 2021.

[Mar03] Murray Marshall. "Optimization of Polynomial Functions". *Canadian Mathematical Bulletin* 46.4 (2003), pp. 575–587.

[Mar06] Murray Marshall. "Representations of Non-Negative Polynomials Having Finitely Many Zeros". *Annales de la faculté des sciences de Toulouse Mathématiques* 15.3 (2006), pp. 599–609.

[Mar08] Murray Marshall. *Positive Polynomials and Sums of Squares*. American Mathematical Soc., 2008. ISBN: 978-0-8218-7527-8.

[Mar09] Murray Marshall. "Representations of Non-Negative Polynomials, Degree Bounds and Applications to Optimization". *Canadian Journal of Mathematics* 61.1 (2009), pp. 205–221.

[MH15] Daniel K. Molzahn and Ian A. Hiskens. "Sparsity-Exploiting Moment-Based Relaxations of the Optimal Power Flow Problem". *IEEE Transactions on Power Systems* 30.6 (2015), pp. 3168–3180.

[Mou18] Bernard Mourrain. "Polynomial–Exponential Decomposition From Moments". *Foundations of Computational Mathematics* 18.6 (2018), pp. 1435–1492.

[MT05] B. Mourrain and P. Trébuchet. "Generalized normal forms and polynomials system solving". *ISSAC: Proceedings of the ACM SIGSAM International Symposium on Symbolic and Algebraic Computation*. Ed. by M. Kauers. 2005, pp. 253–260.

[NDS06] Jiawang Nie, James Demmel, and Bernd Sturmfels. "Minimizing Polynomials via Sum of Squares over the Gradient Ideal". *Mathematical Programming* 106.3 (2006), pp. 587–606.

[Nie13a] Jiawang Nie. "An exact Jacobian SDP relaxation for polynomial optimization". *Mathematical Programming* 137.1-2 (2013), pp. 225–255.

[Nie13b] Jiawang Nie. "Certifying convergence of Lasserre's hierarchy via flat truncation". en. *Mathematical Programming* 142.1 (2013), pp. 485–510.

[Nie13c] Jiawang Nie. "Polynomial Optimization with Real Varieties". *SIAM Journal on Optimization* 23.3 (2013), pp. 1634–1646.

[Nie14] Jiawang Nie. "Optimality conditions and finite convergence of Lasserre's hierarchy". *Mathematical Programming* 146.1-2 (2014), pp. 97–121.

[NW06] Jorge Nocedal and S. Wright. *Numerical Optimization*. 2nd ed. Springer Series in Operations Research and Financial Engineering. New York: Springer-Verlag, 2006. ISBN: 978-0-387-30303-1.

[Par02] Pablo A. Parrilo. *An Explicit Construction of Distinguished Representations of Polynomials Nonnegative Over Finite Sets*. 2002.

[Put93] Mihai Putinar. "Positive Polynomials on Compact Semi-algebraic Sets". *Indiana University Mathematics Journal* 42.3 (1993), pp. 969–984.

[Rez96] Bruce Reznick. "Some Concrete Aspects Of Hilbert's 17th Problem". *In Contemporary Mathematics*. American Mathematical Society, 1996, pp. 251–272.

[Sch00] Claus Scheiderer. "Sums of squares of regular functions on real algebraic varieties". *Transactions of the American Mathematical Society* 352.3 (2000), pp. 1039–1069.

[Sch05a] Claus Scheiderer. "Distinguished representations of non-negative polynomials". *Journal of Algebra* 289.2 (2005), pp. 558–573.

[Sch05b] Claus Scheiderer. "Non-existence of degree bounds for weighted sums of squares representations". *Journal of Complexity* 21.6 (2005), pp. 823–844.

[Sch06] Claus Scheiderer. "Sums of squares on real algebraic surfaces". *manuscripta mathematica* 119.4 (2006), pp. 395–410.

[Sch17] Konrad Schmüdgen. *The Moment Problem*. Graduate Texts in Mathematics. Springer International Publishing, 2017. ISBN: 978-3-319-64545-2.

[Sch91] Konrad Schmüdgen. "The K-moment problem for compact semi-algebraic sets". *Mathematische Annalen* 289.1 (1991), pp. 203–206.