The Approach of Moments for Polynomial Equations

Monique Laurent[1] and Philipp Rostalski[2]

[1] Centrum voor Wiskunde en Informatica, Science Park 123, 1098 XG Amsterdam, and Tilburg University, P.O. Box 90153, 5000 LE Tilburg, Netherlands. M.Laurent@cwi.nl
[2] Department of Mathematics, UC Berkeley, 1067 Evans Hall, Berkeley, CA 94720-3840, USA. philipp@math.berkeley.edu

Summary. In this article we present the moment based approach for computing all real solutions of a given system of polynomial equations. This approach builds upon a lifting method for constructing semidefinite relaxations of several nonconvex optimization problems, using sums of squares of polynomials and the dual theory of moments. A crucial ingredient is a semidefinite characterization of the real radical ideal, consisting of all polynomials with the same real zero set as the system of polynomials to be solved. Combining this characterization with ideas from commutative algebra, (numerical) linear algebra and semidefinite optimization yields a new class of real algebraic algorithms. This article sheds some light on the underlying theory as well as on various extensions and research directions.

1 Introduction

Computing all points $x \in \mathbb{K}^n$ ($\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$) at which a given system of polynomials in n variables

$$h_1, \ldots, h_m \in \mathbb{R}[x_1, \ldots, x_n] = \mathbb{R}[x]$$

vanishes simultaneously, is an old problem arising in many mathematical models in science and engineering, with numerous applications in different areas ranging from control, cryptography, computational geometry, coding theory and computational biology to optimization, robotics, statistics and many others (see, e.g., [43]). In this article we will focus on the characterization and the (numerical) computation of all real roots or, more generally, of all roots lying in some given basic semi-algebraic set, i.e. satisfying some prescribed polynomial inequalities. Solving this kind of problems is a widely studied area with a long tradition and numerous applications; a variety of methods has been proposed to tackle such problems, some of which will be briefly recalled in the next section. In this article we will focus on a new approach for computing the

real roots of polynomial equations. This approach is based on sums of squares of polynomials and the dual theory of moments, combined with semidefinite programming, which is the tool permitting to distinguish algorithmically between *real* and *complex nonreal* elements.

1.1 Existing methods

There is a vast literature on the subject of solving polynomial equations; for information and further references see e.g. the monographs of Basu, Pollack and Roy [2], Dickenstein and Emiris [9], Mora [27, 28], Elkadi and Mourrain [10], Stetter [42], Sturmfels [43]. We do not attempt a complete description of all existing methods, but instead we only try to give a coarse classification. Most existing algorithms can be roughly categorized according to the following criteria: local vs. global search, numerical vs. exact/symbolic computation, and solving over the complex numbers vs. solving over the real numbers.

Over the complex numbers

Symbolic methods. Gröbner bases, resultants or, more generally, border bases and generalized normal form algorithms are typical representatives of this class of methods. The main idea is to compute the structure of the quotient algebra $\mathbb{R}[x]/I$ (where I is the ideal generated by the given polynomials h_i) and to use this information to characterize the roots, e.g., using the shape lemma, or Stickelberger's theorem (viz. the eigenvalue method), or the rational univariate representation (RUR, see e.g. [28] for details).

The following basic fact plays a crucial role: The system of polynomial equations $h_1 = \cdots = h_m = 0$ has finitely many roots if and only if the quotient ring $\mathbb{R}[x]/I$ of the underlying ideal $I = \langle h_1, \ldots, h_m \rangle$ is finite dimensional as a vector space. This in turn enables to *reduce the computation of all complex roots to tasks of finite dimensional linear algebra* (like eigenvalue computations). Roughly speaking, the basic idea is to replace the given system $h_i = 0$ by a new equivalent system $g_j = 0$ with the same set of complex roots, but with a much easier structure facilitating the extraction of the roots.

For instance, one may find an equivalent system comprising polynomials in triangular form $g_1 \in \mathbb{R}[x_1], g_2 \in \mathbb{R}[x_1, x_2], \ldots, g_n \in \mathbb{R}[x_1, \ldots, x_n]$, which can be solved by solving a sequence of *univariate* root finding problems. Such an approach suffers however from the propagation of numerical errors and triangular representations are difficult to compute, typically involving lexicographic Gröbner bases. A more efficient approach is the RUR approach (cf. Rouillier [37]) where the new system has a parametric representation of the form:

$$x_1 = h_1(t)/h(t), \ldots, x_n = h_n(t)/h(t), f(t) = 0 \quad (h_i, h, f \in \mathbb{R}[t]),$$

which requires the solution of *a single univariate* polynomial: $f(t) = 0$ only.

Symbolic-numeric methods. Motivated by the great success of numerical linear algebra, a new trend in applied mathematics is to carefully combine symbolic methods (mostly border bases methods) with numerical calculations, such as singular value decomposition, LU-factorization and other workhorses of numerical linear algebra in order to derive powerful algorithms for large scale problems (see e.g. [30] for details). As mentioned above, symbolic methods are able to transform the given system $h_i = 0$ into a new, better structured system $g_j = 0$. Then the task of computing the complex roots is reduced to (numerical) linear algebra, like computing the eigenvalues/eigenvectors of companion matrices (cf. Section 2.2 below), or univariate root finding.

Numerical methods. The most successful approach in this class of methods is *homotopy continuation*. Such methods rely on Bertini's theorem allowing to deform an easier instance with known solutions of the class of problems to be solved into the original system, without encountering singularities along the path (cf. [39] for details). Keeping track of the roots during this deformation allows to compute the desired roots.

Over the real numbers

While the task of solving polynomial equations over the complex numbers is relatively well understood, computing *only the real roots* is still largely open. The need for methods tailored to real root finding is mainly motivated by applications, where often only the real roots are meaningful, and whose number is typically much smaller than the total number of complex solutions. As an illustration, just consider the simple equation $x_1^2 + x_2^2 = 0$, where not even the dimensions of the real and complex solution sets agree!

So far, real solving methods were mostly build upon local methods combined with a bisection search strategy. More recently, two new global approaches have been considered which can be seen as refinements of complex root finding methods mentioned above: the SDP based moment approach (which is the focus of this article), and a new homotopy continuation method tuned to real roots. The three main classes of methods for real roots are:

Subdivision methods. Combining exclusion criteria to remove parts of the search space not containing any real root and identify regions containing isolated real roots, with local search strategies such as Newton-Raphson or higher order search methods are the basis for the class of subdivision methods. The search space is subdivided until it contains only a single root and Newton's method converges (cf. e.g. [31] for a recent account). Exclusion criteria include real root counting techniques based e.g. on Sturm-Habicht sequences, Descartes' rule of signs (for univariate polynomials), or signatures of Hermite forms (in the multivariate case). Such techniques, combined with deformation techniques using Puiseux series, are also extended to the problem of computing at least one point in each connected component of an algebraic variety (possibly of positive dimension) (cf. [2] for a detailed account).

Khovanskii-Rolle continuation. This method is a recent extension of curve following methods (like homotopy continuation for complex roots) tailored to real roots. It exploits the fact that there are sharp bounds for the number of real roots of systems of equations with few monomials, combined with Gale duality. The approach allows to track significantly fewer paths of an auxiliary system leading to all nondegenerate real solutions of the original system. It is still under investigation, but has the potential to become an efficient algorithm for real root finding (see [3, 40] for details).

Moment methods. This class of methods was first proposed in [18] with extensions in [19, 20], and is the focus of this exposition. The basic idea is to compute the real roots by working in a smaller quotient space, obtained by taking the quotient by the *real radical ideal* $\sqrt[\mathbb{R}]{I}$ of the original ideal I, consisting of all polynomials that vanish at the set of common real roots of the original system $h_i = 0$. In this way, computing the real roots is again reduced to a task of numerical linear algebra, now in the finite dimensional vector space $\mathbb{R}[x]/\sqrt[\mathbb{R}]{I}$ (assuming only that the number of real roots is finite, while the total number of complex roots could be infinite). Finding the real radical ideal is achieved by computing the kernel of a generic moment matrix obtained by solving iteratively certain semidefinite programming problems.

1.2 The basic idea of the moment method

Most symbolic and symbolic/numeric algorithms for solving a system of polynomials decompose the structure of the polynomial ring into its ideal structure (namely, the ideal I generated by the equations to be solved) and its vector space structure (corresponding to the quotient of the polynomial ring by this ideal). While the former is treated with symbolic methods one can use efficient linear algebra for the latter. We start with an elementary introduction. Let

$$h_1(x) = \cdots = h_m(x) = 0 \tag{1}$$

be the system of polynomial equations to be solved. Denote by $D \in \mathbb{N}$ the maximum degree of the polynomials h_i and let $I = \langle h_1, \ldots, h_m \rangle$ be the ideal generated by these polynomials, i.e., the set of all polynomials $\sum_i u_i h_i$ with $u_i \in \mathbb{R}[x]$. If we form the matrix H whose rows are the coefficient vectors of the polynomials h_i, then the roots of the system (1) are precisely the elements $x \in \mathbb{C}^n$ satisfying $H[x]_D = 0$, where for any integer $t \in \mathbb{N}$,

$$[x]_t = (1, x_1, \ldots, x_n, x_1^2, x_1x_2, \ldots, x_n^t)$$

denotes the vector of all monomials of degree at most t. Augmenting the system (1) with new polynomials obtained by multiplying the h_i's by monomials does not change its set of common roots. We iterate degree by degree and add all possible multiples of the h_i's with degree at most some given $t \in \mathbb{N}$ (of course more complicated augmentations would be possible too). In other

words, we generate all 'valid' equations: $x^\alpha h_i = 0$ where $|\alpha| \leq t - \deg(h_i)$. This yields a new, larger system of polynomials whose coefficient vectors make the rows of a matrix $\tilde{H}_t$ (known as Sylvester or Macaulay-like matrix). Again, the roots of (1) are those elements $x \in \mathbb{C}^n$ satisfying $\tilde{H}_t[x]_t = 0$.

The basic idea is to *linearize* this system of equations by introducing variables $y = (y_\alpha)$ for the monomials x^α and to solve instead a *linear system*:

$$\tilde{H}_t y = 0. \tag{2}$$

The kernel of the matrix $\tilde{H}_t$ is a vector space, which contains the vectors $[x]_t$ for all roots x of the system (1) and thus also their linear span. When the system (1) has finitely many complex roots, it turns out that, for t large enough, (some projection of) the kernel of $\tilde{H}_t$ coincides with the linear span of the monomial vectors corresponding to the roots of (1), which opens the way to extracting the roots. More precisely, the central observation (dating back to [24]) is that for t large enough a Gaussian elimination on the Sylvester matrix $\tilde{H}_t$ will reveal a Gröbner basis for the ideal I and thus the desired quotient ring structure $\mathbb{R}[x]/I$. This in turn can be used to reduce the multivariate root finding problem to a simple eigenvalue calculation (as recalled in Section 2.2).

If we want to compute the real roots only, we need a mechanism to cancel out all (or as many as possible) nonreal solutions among the complex ones. This cancellation can be done by augmenting the original system (1) with additional polynomials derived from sums of squares of polynomials in the ideal I. We introduce this idea by means of a simple example.

Example 1. Consider the ideal $I \subseteq \mathbb{R}[x_1, x_2]$ generated by the polynomial $h = x_1^2 + x_2^2$. The complex variety is positive dimensional, since it consists of infinitely many complex roots: $x_2 = \pm \mathbf{i} x_1$ ($x_1 \in \mathbb{C}$), while the origin $(0, 0)$ is the only real root. If we add the two polynomials $p_1 = x_1, p_2 = x_2$ to I the real variety remains unchanged, but none of the complex nonreal roots survives this intersection. Note that p_1, p_2 have the property that the polynomial $p_1^2 + p_2^2 = h$ is a sum of squares of polynomials belonging to I.

This example illustrates the following fact: If the p_i's are polynomials for which $\sum_i p_i^2 \in I$, then each p_i vanishes at all the real roots of the ideal I (but not necessarily at its complex nonreal roots!). Thus we can add the p_i's to the original system (1) without altering its set of real roots. The formal tool behind this augmentation is the Real Nullstellensatz (see Theorem 1), which states that the set of real solutions to the system (1) remains unchanged if we add to it any polynomial appearing with an even degree in a sum of squares polynomial that belongs to I. The set of all such polynomials is known as the *real radical ideal* of I, denoted as $\sqrt[\mathbb{R}]{I}$ (see Section 2 for definitions). A main feature of the moment matrix method is that it permits to generate the polynomials in the real radical ideal in a systematic way, using duality.

Let us first look directly at the additional properties that are satisfied by a vector $y = [x]_t \in \operatorname{Ker} \tilde{H}_t$, when x is a *real* root of (1). Observe that the matrix

$[x]_s[x]_s^T$ is obviously positive semidefinite for any integer s. Again, we can 'linearize' it (replacing x^α by y_α) and then we obtain a matrix of generalized Hankel type: $M_s(y) = (y_{\alpha+\beta})_{\alpha,\beta\in\mathbb{N}^n_s}$. As an illustration we display $M_s(y)$ for the case $n = 2$:

$$[x]_s[x]_s^T = \begin{pmatrix} 1 & x_1 & x_2 & x_1^2 & \dots \\ x_1 & x_1^2 & x_1x_2 & x_1^3 & \dots \\ x_2 & x_1x_2 & x_2^2 & x_1^2x_2 & \dots \\ x_1^2 & x_1^3 & x_1^2x_2 & x_1^4 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \rightsquigarrow M_s(y) \begin{pmatrix} 1 & y_{10} & y_{01} & y_{20} & \dots \\ y_{10} & y_{20} & y_{11} & y_{30} & \dots \\ y_{01} & y_{11} & y_{02} & y_{21} & \dots \\ y_{20} & y_{30} & y_{21} & y_{40} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Matrices with this particular generalized Hankel structure are also known as *moment matrices*, see Section 4.3. Therefore, we can restrict the search in the kernel of the Sylvester matrix $\tilde{H}_t$ to the vectors y satisfying the additional positive semidefiniteness condition: $M_s(y) \succeq 0$ for all $s \leq t/2$. This condition captures precisely the 'real algebraic' nature of real numbers vs. complex numbers, as it would not be valid for vectors y corresponding to complex nonreal roots. Let us illustrate this approach on the preceding simple example.

Example 2. (Example 1 cont.) Say we wish to compute the real roots of the polynomial $h = x_1^2 + x_2^2$. After linearization, the constraint $Hy = 0$ reads: $y_{20} + y_{02} = 0$. Positive semidefiniteness requires $y_{20} \geq 0, \quad y_{02} \geq 0$ which, combined with $y_{20} + y_{02} = 0$ implies $y_{20} = y_{02} = 0$ and thus $y_{10} = y_{01} = y_{11} = 0$ (using again $M_1(y) \succeq 0$). Therefore, we find $y = (1, 0, 0, 0, 0, 0)$ as the unique solution, so that $y = [x]_2$ corresponds to the unique real root $x = (0, 0)$ of h. The kernel of $M_1(y)$ contains the vectors $(0, 1, 0)$ and $(0, 0, 1)$, which can be seen as the coefficient vectors of the two polynomials $p_1 = x_1$ and $p_2 = x_2$ in the monomial basis $\{1, x_1, x_2\}$ of $\mathbb{R}[x]_1$. In other words the kernel of $M_1(y)$ already contains a basis of the real radical ideal $\sqrt[\mathbb{R}]{I}$.

Although the above example is extremely simplistic, it conveys the main idea: The kernel of $M_s(y)$ characterizes (for s large enough) the real radical ideal and plays the role of the range space of H in standard normal form algorithms.

1.3 Organization of the article

First we recall some basic material from polynomial algebra in Section 2. This material can be found in most standard textbooks and is used throughout the article. The relation between moment matrices and real radical ideals as well as the moment method for real root finding is discussed in Section 3. This section and in particular the semidefinite characterization of the real radical ideal form the heart of the present publication. We also discuss how several other complex root finding methods may be tailored for the task of real solving by using this semidefinite characterization. Section 4 briefly discusses several related research directions ranging from polynomial optimization and the study of semi-algebraic sets to colon ideals and positive dimensional varieties. Throughout the article we illustrate the results with various examples.

2 Preliminaries of polynomial algebra

2.1 Polynomial ideals and varieties

The polynomial ring and its dual. For the sake of simplicity we deal with polynomials with real coefficients only although some results remain valid for polynomials with complex coefficients. Throughout $\mathbb{R}[x] := \mathbb{R}[x_1, \ldots, x_n]$ denotes the ring of multivariate polynomials in n variables (with roots in $\mathbb{C}^n$). For $\alpha \in \mathbb{N}^n$, x^α denotes the monomial $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, with degree $|\alpha| := \sum_{i=1}^n \alpha_i$. Set $\mathbb{N}^n_t := \{\alpha \in \mathbb{N}^n \mid |\alpha| \leq t\}$ and let

$$[x]_\infty = (x^\alpha)_{\alpha \in \mathbb{N}^n}, \quad [x]_t = (x^\alpha)_{\alpha \in \mathbb{N}^n_t}$$

denote the vectors comprising all monomials (resp., all monomials of degree at most t) in n variables. A polynomial $p \in \mathbb{R}[x]$ can be written as $p = \sum_{\alpha \in \mathbb{N}^n} p_\alpha x^\alpha$ with finitely many nonzero p_α's; its support is the set of monomials appearing with a nonzero coefficient, its (total) degree $\deg(p)$ is the largest degree of a monomial in the support of p, and $\mathrm{vec}(p) = (p_\alpha)$ denotes the vector of coefficients of p. The set $\mathbb{R}[x]_t$ consists of all polynomials with degree at most t.

Given a vector space A on $\mathbb{R}$, its dual space A^* consists of all linear functionals from A to $\mathbb{R}$. The orthogonal complement of a subset $B \subseteq A$ is

$$B^\perp := \{L \in A^* \mid L(b) = 0 \ \forall b \in B\}$$

and $\mathrm{Span}_{\mathbb{R}}(B)$ denotes the linear span of B. Then, $\mathrm{Span}_{\mathbb{R}}(B) \subseteq (B^\perp)^\perp$, with equality when A is finite dimensional. We consider here the case $A = \mathbb{R}[x]$ and $A = \mathbb{R}[x]_t$. Examples of linear functionals on $\mathbb{R}[x]$ are the *evaluation*

$$\Lambda_v : p \in \mathbb{R}[x] \mapsto \Lambda_v(p) = p(v) \tag{3}$$

at a point $v \in \mathbb{R}^n$ and, more generally, the differential functional

$$\partial_v^\alpha : p \in \mathbb{R}[x] \mapsto \partial_v^\alpha(p) = \frac{1}{\prod_{i=1}^n \alpha_i!} \left(\frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}} p \right)(v), \tag{4}$$

which evaluates at $v \in \mathbb{R}^n$ the (scaled) α-th derivative of p (where $\alpha \in \mathbb{N}$). For $\alpha = 0$, ∂_v^α coincides with the evaluation at v, i.e., $\partial_v^0 = \Lambda_v$. For $\alpha, \beta \in \mathbb{N}^n$,

$$\partial_0^\alpha(x^\beta) = 1 \ \text{ if } \alpha = \beta, \quad \text{and 0 otherwise.}$$

Therefore, any linear form $\Lambda \in \mathbb{R}[x]^*$ can be written in the form:

$$\Lambda = \sum_{\alpha \in \mathbb{N}^n} \Lambda(x^\alpha) \partial_0^\alpha.$$

This is in fact a formal power series as in general infinitely many $\Lambda(x^\alpha)$ are nonzero. Let $y = (y_\alpha)$ denote the coefficient series of Λ in (∂_0^α) i.e. $y_\alpha = \Lambda(x^\alpha)$,

such that $\Lambda(p) = y^T \operatorname{vec}(p)$ for all $p \in \mathbb{R}[x]$. For instance, the evaluation at $v \in \mathbb{R}^n$ reads $\Lambda_v = \sum_\alpha v^\alpha \partial_0^\alpha$, with coefficient series $[v]_\infty = (v^\alpha)_{\alpha \in \mathbb{N}^n}$ in (∂_0^α).

Ideals and varieties. A linear subspace $I \subseteq \mathbb{R}[x]$ is an *ideal* if $p \in I$, $q \in \mathbb{R}[x]$ implies $pq \in I$. The *ideal generated* by $h_1, \ldots, h_m \in \mathbb{R}[x]$ is defined as

$$I = \langle h_1, \ldots h_m \rangle := \Big\{ \sum_{j=1}^m u_j h_j \mid \quad u_1, \ldots, u_m \in \mathbb{R}[x] \Big\}$$

and the set $\{h_1, \ldots, h_m\}$ is then called a *basis* of I. By the finite basis theorem [6, §2.5, Thm. 4], every ideal in $\mathbb{R}[x]$ admits a finite basis. Given an ideal $I \subseteq \mathbb{R}[x]$, the *algebraic variety* of I is the set

$$V_{\mathbb{C}}(I) = \{v \in \mathbb{C}^n \mid h_j(v) = 0 \; \forall j = 1, \ldots, m\}$$

of common complex zeros to all polynomials in I and its *real variety* is

$$V_{\mathbb{R}}(I) := V_{\mathbb{C}}(I) \cap \mathbb{R}^n.$$

The ideal I is said to be *zero-dimensional* when its complex variety $V_{\mathbb{C}}(I)$ is finite. Conversely, the *vanishing ideal* of a subset $V \subseteq \mathbb{C}^n$ is the ideal

$$\mathcal{I}(V) := \{f \in \mathbb{R}[x] \mid f(v) = 0 \; \forall v \in V\}.$$

For an ideal $I \subseteq \mathbb{R}[x]$, we may also define the ideal

$$\sqrt{I} := \Big\{ f \in \mathbb{R}[x] \,\Big|\quad f^m \in I \;\text{ for some } m \in \mathbb{N} \setminus \{0\} \Big\},$$

called the *radical ideal* of I, and the *real radical ideal* (or *real ideal*)

$$\sqrt[\mathbb{R}]{I} := \Big\{ p \in \mathbb{R}[x] \,\Big|\, p^{2m} + \sum_j q_j^2 \in I \;\text{ for some } q_j \in \mathbb{R}[x], m \in \mathbb{N} \setminus \{0\} \Big\}.$$

An ideal I is said to be *radical* (resp., *real radical*) if $I = \sqrt{I}$ (resp., $I = \sqrt[\mathbb{R}]{I}$). For instance, the ideal $I = \langle x_1^2 + x_2^2 \rangle$ is not real radical since $x_1, x_2 \in \sqrt[\mathbb{R}]{I} \setminus I$. As can be easily verified, I is radical if and only if $p^2 \in I$ implies $p \in I$, and I is real radical if and only if $\sum_i p_i^2 \in I$ implies $p_i \in I \;\; \forall i$. We have the following chains of inclusion:

$$I \subseteq \sqrt{I} \subseteq \mathcal{I}(V_{\mathbb{C}}(I)), \quad I \subseteq \sqrt[\mathbb{R}]{I} \subseteq \mathcal{I}(V_{\mathbb{R}}(I)).$$

The relation between vanishing and (real) radical ideals is stated in the following two famous theorems:

Theorem 1. *Let $I \subseteq \mathbb{R}[x]$ be an ideal.*

(i) Hilbert's Nullstellensatz *(see, e.g., [6, §4.1])* $\sqrt{I} = \mathcal{I}(V_{\mathbb{C}}(I))$.
(ii) Real Nullstellensatz *(see, e.g., [4, §4.1])* $\sqrt[\mathbb{R}]{I} = \mathcal{I}(V_{\mathbb{R}}(I))$.

2.2 The eigenvalue method for complex roots

The quotient space $\mathbb{R}[x]/I$. The quotient set $\mathbb{R}[x]/I$ consists of all cosets $[f] := f + I = \{f + q \mid q \in I\}$ for $f \in \mathbb{R}[x]$, i.e. all equivalent classes of polynomials in $\mathbb{R}[x]$ modulo I. This quotient set $\mathbb{R}[x]/I$ is an algebra with addition $[f]+[g] := [f+g]$, scalar multiplication $\lambda[f] := [\lambda f]$ and multiplication $[f][g] := [fg]$, for $\lambda \in \mathbb{R}$, $f, g \in \mathbb{R}[x]$. The following classical result relates the dimension of $\mathbb{R}[x]/I$ and the cardinality of the variety $V_{\mathbb{C}}(I)$ (see e.g. [6, 42]).

Theorem 2. *Let I be an ideal in $\mathbb{R}[x]$. Then,*

$$|V_{\mathbb{C}}(I)| < \infty \iff \dim \mathbb{R}[x]/I < \infty.$$

Moreover, $|V_{\mathbb{C}}(I)| \leq \dim\ \mathbb{R}[x]/I$, with equality if and only if I is radical.

Assume that the number of complex roots is finite and set $N := \dim \mathbb{R}[x]/I$, so that $|V_{\mathbb{C}}(I)| \leq N < \infty$. Consider a set $\mathcal{B} := \{b_1, \ldots, b_N\} \subseteq \mathbb{R}[x]$ for which the cosets $[b_1], \ldots, [b_N]$ are pairwise distinct and $\{[b_1], \ldots, [b_N]\}$ is a (linear) basis of $\mathbb{R}[x]/I$. By abuse of language we also say that $\mathcal{B}$ itself is a basis of $\mathbb{R}[x]/I$. Then every $f \in \mathbb{R}[x]$ can be written in a unique way as $f = \sum_{i=1}^{N} c_i b_i + p$, where $c_i \in \mathbb{R}$ and $p \in I$. The polynomial

$$\mathcal{N}_{\mathcal{B}}(f) := \sum_{i=1}^{N} c_i b_i$$

is called the *normal form* of f modulo I with respect to the basis $\mathcal{B}$. In other words, we have the direct sum decomposition:

$$\mathbb{R}[x] = \mathrm{Span}_{\mathbb{R}}(\mathcal{B}) \oplus I,$$

and $\mathrm{Span}_{\mathbb{R}}(\mathcal{B})$ and $\mathbb{R}[x]/I$ are isomorphic vector spaces. We now introduce the eigenvalue method for computing all roots of a zero-dimensional ideal, which we first describe in the univariate case.

Computing roots with companion matrices. Consider first a univariate polynomial $p = x^d - a_{d-1}x^{d-1} - \ldots - a_1 x - a_0$ and the ideal $I = \langle p \rangle$. Then the set $\mathcal{B} = \{1, x, \ldots, x^{d-1}\}$ is a basis of $\mathbb{R}[x]/I$. The following matrix

$$\mathcal{X} := \begin{pmatrix} 0 & \ldots & 0 & a_0 \\ 1 & & & a_1 \\ & \ddots & & \vdots \\ & & 1 & a_{d-1} \end{pmatrix}$$

is known as the *companion matrix* of the polynomial p. One can easily verify that $\det(\mathcal{X} - xI) = (-1)^d p(x)$, so that the eigenvalues of $\mathcal{X}$ are precisely the roots of the polynomials p. Therefore the roots of a univariate polynomial can be found with an eigenvalue computation. Moreover, the columns of the

companion matrix $\mathcal{X}$ correspond to the normal forms of the monomials in $x\mathcal{B} = \{x, x^2, \ldots, x^d\}$ modulo I with respect to the basis $\mathcal{B}$. As we now see these facts extend naturally to the multivariate case.

Given $h \in \mathbb{R}[x]$, we define the *multiplication (by h) operator* in $\mathbb{R}[x]/I$ as

$$\begin{array}{rcl} \mathcal{X}_h : \mathbb{R}[x]/I & \longrightarrow & \mathbb{R}[x]/I \\ [f] & \longmapsto & \mathcal{X}_h([f]) := [hf]\,, \end{array} \tag{5}$$

which can be represented by its matrix (again denoted $\mathcal{X}_h$ for simplicity) with respect to the basis $\mathcal{B}$ of $\mathbb{R}[x]/I$. Namely, with $\mathcal{N}_{\mathcal{B}}(hb_j) := \sum_{i=1}^N a_{ij}b_i$ (for $a_{ij} \in \mathbb{R}$), the jth column of $\mathcal{X}_h$ is the vector $(a_{ij})_{i=1}^N$. Note also that, since $hb_j - \mathcal{N}_{\mathcal{B}}(hb_j) \in I$, polynomials in I can be read directly from $\mathcal{X}_h$. This fact will play an important role for border bases, see Section 2.3. In the univariate case, when $I = \langle p \rangle$ and $h = x$, the multiplication matrix $\mathcal{X}_x$ is precisely the companion matrix $\mathcal{X}$ of p introduced above. Throughout we also denote by $\mathcal{X}_i := \mathcal{X}_{x_i}$ the multiplication operator by the variable x_i in the multivariate case.

The following famous result (see e.g. [5, Chap. 2§4]) relates the eigenvalues of the multiplication operators in $\mathbb{R}[x]/I$ to the algebraic variety $V_{\mathbb{C}}(I)$. This result underlies the well known *eigenvalue method*, which plays a central role in many algorithms for complex root solving.

Theorem 3. (Stickelberger theorem) *Let I be a zero-dimensional ideal in $\mathbb{R}[x]$, let $\mathcal{B}$ be a basis of $\mathbb{R}[x]/I$, and let $h \in \mathbb{R}[x]$. The eigenvalues of the multiplication operator $\mathcal{X}_h$ are the evaluations $h(v)$ of the polynomial h at the points $v \in V_{\mathbb{C}}(I)$. Moreover, for all $v \in V_{\mathbb{C}}(I)$,*

$$(\mathcal{X}_h)^T [v]_{\mathcal{B}} = h(v)[v]_{\mathcal{B}},$$

setting $[v]_{\mathcal{B}} = (b(v))_{b\in\mathcal{B}}$, that is, the vector $[v]_{\mathcal{B}}$ is a left eigenvector of the multiplication operator with eigenvalue $h(v)$.

Therefore the eigenvalues of the matrices $\mathcal{X}_i$ are the ith coordinates of the points $v \in V_{\mathbb{C}}(I)$, which can be derived from the left eigenvectors $[v]_{\mathcal{B}}$. Practically, one can recover the roots from the left eigenvectors when the eigenspaces of $\mathcal{X}_h^T$ all have dimension one. This is the case when the values $h(v)$ $(v \in V_{\mathbb{C}}(I))$ are pairwise distinct (easy to achieve, e.g., if we choose h to be a generic linear form) and when the ideal I is radical (since the dimension of $\mathbb{R}[x]/I$ is then equal to the number of roots so that the vectors $[v]_{\mathcal{B}}$ $(v \in V_{\mathbb{C}}(I))$ form a complete basis of eigenvectors).

Summarizing, *the task of solving a system of polynomial equations is reduced to a task of numerical linear algebra once a basis of $\mathbb{R}[x]/I$ and a normal form algorithm are available,* as they permit the construction of the multiplication matrices $\mathcal{X}_i$, $\mathcal{X}_h$. Moreover, the roots $v \in V_{\mathbb{C}}(I)$ can be successfully constructed from the eigenvectors/eigenvalues of $\mathcal{X}_h$ when I is radical and h is generic. Our strategy for computing the real variety $V_{\mathbb{R}}(I)$ will be to

compute a linear basis of the quotient space $\mathbb{R}[x]/\sqrt[\mathbb{R}]{I}$ and the corresponding multiplication matrices, so that we we can apply the eigenvalue method precisely in this setting of having a radical (even real radical) ideal.

The number of (real) roots can be counted using Hermite's quadratic form:

$$\begin{array}{rl} S_h : \mathbb{R}[x]/I \times \mathbb{R}[x]/I \to & \mathbb{R} \\ ([f],[g]) & \mapsto \mathrm{Tr}(\mathcal{X}_{fgh}). \end{array}$$

Here, $\mathrm{Tr}(\mathcal{X}_{fgh})$ is the trace of the multiplication (by the polynomial fgh) matrix. As S_h is a symmetric matrix, all its eigenvalues are real. Denote by $\sigma_+(S_h)$ (resp., $\sigma_-(S_h)$) its number of positive (resp., negative) eigenvalues. The following classical result shows how to count the number of roots satisfying prescribed sign conditions (cf. e.g. [2]).

Theorem 4. *Let $I \subseteq \mathbb{R}[x]$ be a zero-dimensional ideal and $h \in \mathbb{R}[x]$. Then,*

$$\mathrm{rank}\, S_h = |\{v \in V_{\mathbb{C}}(I) \mid h(v) \neq 0\}|,$$

$$\sigma_+(S_h) - \sigma_-(S_h) = |\{v \in V_{\mathbb{R}}(I) \mid h(v) > 0\}| - |\{v \in V_{\mathbb{R}}(I) \mid h(v) < 0\}|.$$

In particular, for the constant polynomial $h = 1$,

$$\mathrm{rank}(S_1) = |V_{\mathbb{C}}(I)| \quad \textit{and} \quad \sigma_+(S_1) - \sigma_-(S_1) = |V_{\mathbb{R}}(I)|.$$

2.3 Border bases and normal forms

The eigenvalue method for solving polynomial equations (described in the preceding section) requires the knowledge of a basis of $\mathbb{R}[x]/I$ and of an algorithm to compute the normal form of a polynomial with respect to this basis.

A well known basis of $\mathbb{R}[x]/I$ is the set of standard monomials with respect to some monomial ordering. The classical way to find standard monomials is to construct a Gröbner basis of I (then the standard monomials are the monomials not divisible by any polynomial in the Gröbner basis). Moreover, once a Gröbner basis is known, the normal form of a polynomial can be found via a polynomial division algorithm (see, e.g., [6, Chap. 1] for details). Other techniques have been proposed, producing more general bases which do not depend on a specific monomial ordering and often are numerically more stable. In particular, algorithms have been proposed for constructing border bases of I leading to general (connected to 1) bases of $\mathbb{R}[x]/I$ (see [9, Chap. 4], [15], [29], [42]); these objects are introduced below. The moment matrix approach for computing real roots presented in this article leads naturally to the computation of such general bases.

Definition 1. *Given a set $\mathcal{B}$ of monomials, define the new sets of monomials*

$$\mathcal{B}^+ := \mathcal{B} \cup \bigcup_{i=1}^{n} x_i\mathcal{B} = \mathcal{B} \cup \{x_i b \mid b \in \mathcal{B},\ i = 1, \ldots, n\}, \quad \partial\mathcal{B} = \mathcal{B}^+ \setminus \mathcal{B},$$

called, respectively, the one-degree prolongation *of $\mathcal{B}$ and the* border *of $\mathcal{B}$. The set $\mathcal{B}$ is said to be* connected to 1 *if $1 \in \mathcal{B}$ and each $m \in \mathcal{B} \setminus \{1\}$ can be written as $m = x_{i_1} \dots x_{i_k}$ with $x_{i_1}, x_{i_1}x_{i_2}, \dots, x_{i_1} \cdots x_{i_k} \in \mathcal{B}$. Moreover, $\mathcal{B}$ is said to be* stable by division *if all divisors of $m \in \mathcal{B}$ also belong to $\mathcal{B}$. Obviously, $\mathcal{B}$ is connected to 1 if it is stable by division.*

Assume $\mathcal{B}$ is a set of monomials which is connected to 1. For each border monomial $m \in \partial\mathcal{B}$, consider a polynomial f_m of the form

$$f_m := m - r_m, \quad \text{where } r_m \in \mathrm{Span}_{\mathbb{R}}(\mathcal{B}). \tag{6}$$

The family $F := \{f_m \mid m \in \partial\mathcal{B}\}$ is called a *rewriting family for $\mathcal{B}$* in [30, 32]. Using F, one can express all border monomials in $\partial\mathcal{B}$ as linear combinations of monomials in $\mathcal{B}$ modulo the ideal $\langle F \rangle$. Moreover, the rewriting family F can be used in a division algorithm to rewrite any polynomial $p \in \mathbb{R}[x]$ as

$$p = r + \sum_{m \in \partial\mathcal{B}} u_m f_m, \quad \text{where } r \in \mathrm{Span}_{\mathbb{R}}(\mathcal{B}),\ u_m \in \mathbb{R}[x]. \tag{7}$$

This expression is in general not unique, as it depends on the order in which the polynomials of F are used throughout the division process.

Example 3. Let $\mathcal{B} = \{1, x_1, x_2\}$ with border set $\partial\mathcal{B} = \{x_1^2, x_1x_2, x_2^2\}$, and consider the rewriting family

$$F = \{f_{x_1^2} = x_1^2 + 1, f_{x_1x_2} = x_1x_2 - 1, f_{x_2^2} = x_2^2 + 1\}.$$

There are two possibilities to rewrite the polynomial $p = x_1^2x_2$. Either, first divide by $f_{x_1x_2}$ and obtain $p = x_1^2x_2 = x_1 f_{x_1x_2} + x_1$ with $r = x_1$, or first divide by $f_{x_1^2}$ and obtain $p = x_1^2x_2 = x_2 f_{x_1^2} - x_2$ with $r = -x_2$.

In view of (7), the set $\mathcal{B}$ spans the vector space $\mathbb{R}[x]/\langle F \rangle$, but is in general not linearly independent. Linear independence guaranties uniqueness of the decomposition (7) and, as Theorem 5 below shows, is equivalent to the commutativity of certain formal multiplication operators.

Consider the linear operator $\mathcal{X}_i : \mathrm{Span}_{\mathbb{R}}(\mathcal{B}) \to \mathrm{Span}_{\mathbb{R}}(\mathcal{B})$ defined using the rewriting family F, namely, for $b \in \mathcal{B}$,

$$\mathcal{X}_i(b) = \begin{cases} x_ib & \text{if } x_ib \in \mathcal{B}, \\ x_ib - f_{x_ib} = r_{x_ib} & \text{otherwise,} \end{cases}$$

and extend $\mathcal{X}_i$ to $\mathrm{Span}_{\mathbb{R}}(\mathcal{B})$ by linearity. Denote also by $\mathcal{X}_i$ the matrix of this linear operator, which can be seen as a *formal multiplication (by x_i) matrix*.

Theorem 5. *[29] Let F be a rewriting family for a set $\mathcal{B}$ of monomials connected to 1, and consider the ideal $J := \langle F \rangle$. The following conditions are equivalent:*

(i) *The formal multiplication matrices $\mathcal{X}_1, \dots, \mathcal{X}_n$ commute pairwise.*

(ii) *The set $\mathcal{B}$ is a (linear) basis of $\mathbb{R}[x]/J$, i.e., $\mathbb{R}[x] = \mathrm{Span}_{\mathbb{R}}(\mathcal{B}) \oplus J$.*

Then, the set F is said to be a border basis *of the ideal J, and the matrix $\mathcal{X}_i$ represents the multiplication operator by x_i in $\mathbb{R}[x]/J$ with respect to $\mathcal{B}$.*

This theorem is the crucial tool for efficient root finding algorithms based on normal form reductions, which iteratively construct a system of polynomial equations giving a rewriting family corresponding to a commuting family of multiplication matrices (thus reducing the root finding problem to an eigenvalue computation, see [30]). We illustrate Theorem 5 on a small example.

Example 4. Let $\mathcal{B} = \{1, x_1\}$ with border set $\partial\mathcal{B} = \{x_2, x_1x_2, x_1^2\}$, and consider the rewriting family

$$F = \{f_{x_1^2} = x_1^2 + 1, f_{x_1x_2} = x_1x_2 - 1, f_{x_2} = x_2 + x_1\}.$$

As $x_1 \in \mathcal{B}$, $x_1^2 = f_{x_1^2} - 1$, $x_2 = f_{x_2} - x_1$, and $x_2x_1 = f_{x_1x_2} + 1$, we have

$$\mathcal{X}_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \mathcal{X}_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

As the formal multiplication matrices $\mathcal{X}_1$, $\mathcal{X}_2$ commute, we can conclude that F is a border basis of $\langle F\rangle$ and $\mathbb{R}[x] = \mathrm{Span}_{\mathbb{R}}(\mathcal{B}) \oplus \langle F\rangle$.

3 The moment method for real root finding

We just saw that computing the complex roots of an ideal can be reduced to an eigenvalue computation. This technique applies only when the number of complex roots is finite, and the eigenvalue computation is on matrices whose size is at least the number of complex roots. However, in most applications, one is only interested in the real roots, whose number can be a very small fraction of the total number of roots. Therefore one needs a tool to isolate the real roots from the complex nonreal ones. As we briefly mentioned in the Introduction, a possible strategy is to add new polynomials from the real radical ideal to the original system to be solved. To find these polynomials in a systematic way we propose to work on the 'dual side', i.e. to consider linear forms Λ on the polynomial ring $\mathbb{R}[x]$ or its subspaces $\mathbb{R}[x]_t$ of bounded degree. Indeed, it turns out that the kernel of such linear forms carries all information about the real radical ideal and the real variety when the linear form is assumed to satisfy some positivity condition. In this section we explain the method in detail and illustrate it on a few examples.

3.1 Positive linear forms and real radical ideals

Given a linear form $\Lambda \in \mathbb{R}[x]^*$, consider the quadratic form on $\mathbb{R}[x]$

$$Q_\Lambda : f \in \mathbb{R}[x] \mapsto Q_\Lambda(f) = \Lambda(f^2) \in \mathbb{R},$$

with kernel $\mathrm{Ker}\, Q_\Lambda := \{f \in \mathbb{R}[x] \mid \Lambda(fg) = 0 \ \forall g \in \mathbb{R}[x]\}$.

Definition 2. (Positivity) *$\Lambda \in \mathbb{R}[x]^*$ is said to be* positive *if $\Lambda(f^2) \geq 0$ for all $f \in \mathbb{R}[x]$, i.e., if the quadratic form Q_Λ is positive semidefinite.*

The following simple lemma provides the link to real radical polynomial ideals.

Lemma 1. *[21, 26] Let $\Lambda \in \mathbb{R}[x]^*$. Then $\operatorname{Ker} Q_\Lambda$ is an ideal in $\mathbb{R}[x]$, which is real radical when Λ is positive.*

Proof. $\operatorname{Ker} Q_\Lambda$ is obviously an ideal, from its definition. Assume Λ is positive. First we show that, for $p \in \mathbb{R}[x]$, $\Lambda(p^2) = 0 \Longrightarrow \Lambda(p) = 0$. Indeed, if $\Lambda(p^2) = 0$ then, for any scalar $t \in \mathbb{R}$, we have:

$$0 \leq \Lambda((p+t)^2) = \Lambda(p^2) + 2t\Lambda(p) + t^2\Lambda(1) = t(2\Lambda(p) + t\Lambda(1)),$$

which implies $\Lambda(p) = 0$. Assume now $\sum_i p_i^2 \in \operatorname{Ker} Q_\Lambda$ for some $p_i \in \mathbb{R}[x]$; we show $p_i \in \operatorname{Ker} Q_\Lambda$. For any $g \in \mathbb{R}[x]$, we have $0 = \Lambda(g^2(\sum_i p_i^2)) = \sum_i \Lambda(p_i^2 g^2)$ which, as $\Lambda(p_i^2 g^2) \geq 0$, implies $\Lambda(p_i^2 g^2) = 0$. By the above, this in turn implies $\Lambda(p_i g) = 0$, thus showing $p_i \in \operatorname{Ker} Q_\Lambda$. Therefore, $\operatorname{Ker} Q_\Lambda$ is real radical. □

We now introduce moment matrices, which permit to reformulate positivity of Λ in terms of positive semidefiniteness of an associated matrix $M(\Lambda)$.

Definition 3. (Moment matrix) *A symmetric matrix $M = (M_{\alpha,\beta})$ indexed by $\mathbb{N}^n$ is said to be a* moment matrix *(or a* generalized Hankel matrix*) if its (α, β)-entry depends only on the sum $\alpha + \beta$ of the indices. Given $\Lambda \in \mathbb{R}[x]^*$, the matrix*

$$M(\Lambda) := (\Lambda(x^\alpha x^\beta))_{\alpha,\beta \in \mathbb{N}^n}$$

is called the moment matrix[3] *of Λ.*

As $Q_\Lambda(p) = \Lambda(p^2) = \operatorname{vec}(p)^T M(\Lambda) \operatorname{vec}(p) \ \ \forall p \in \mathbb{R}[x]$, $M(\Lambda)$ is the matrix of the quadratic form Q_Λ in the monomial base, and Λ is positive if and only if $M(\Lambda) \succeq 0$.

Moreover, $p \in \operatorname{Ker} Q_\Lambda \Longleftrightarrow \operatorname{vec}(p) \in \operatorname{Ker} M(\Lambda) \ \ \forall p \in \mathbb{R}[x]$. Throughout we identify polynomials $p = \sum_\alpha p_\alpha x^\alpha$ with their coefficient vectors $\operatorname{vec}(p) = (p_\alpha)_\alpha$ and thus $\operatorname{Ker} Q_\Lambda$ with $\operatorname{Ker} M(\Lambda)$. Hence we view $\operatorname{Ker} M(\Lambda)$ as a set of polynomials. By Lemma 1, $\operatorname{Ker} M(\Lambda)$ is an ideal of $\mathbb{R}[x]$, which is real radical when $M(\Lambda) \succeq 0$.

Example 5. For $n = 2$, consider the linear form $\Lambda \in \mathbb{R}[x]^*$ defined by $\Lambda(1) = \Lambda(x_1^2) = 1$ and $\Lambda(x_1^{\alpha_1} x_2^{\alpha_2}) = 0$ for all other monomials. Then Λ is positive and the kernel of $M(\Lambda)$ is the ideal $\langle x_2, 1 - x_1^2 \rangle$.

The kernel of a moment matrix $M(\Lambda)$ turns out to be a zero-dimensional ideal precisely when the matrix $M(\Lambda)$ has finite rank.

[3] In the literature the moment matrix is often denoted by $M(y) = (y_{\alpha+\beta})_{\alpha,\beta \in \mathbb{N}^n}$, where y is the coefficient series of Λ, i.e., $\Lambda = \sum_\alpha y_\alpha \partial_0^\alpha$.

Lemma 2. *Let $\Lambda \in \mathbb{R}[x]^*$ and let $\mathcal{B}$ be a set of monomials. Then, $\mathcal{B}$ indexes a maximal linearly independent set of columns of $M(\Lambda)$ if and only if $\mathcal{B}$ corresponds to a basis of $\mathbb{R}[x]/\operatorname{Ker} M(\Lambda)$. That is,*

$$\operatorname{rank} M(\Lambda) = \dim \mathbb{R}[x]/\operatorname{Ker} M(\Lambda).$$

Lemma 3. *If $\Lambda = \Lambda_v$ is the evaluation at $v \in \mathbb{R}^n$, then $M(\Lambda_v) = [v]_\infty [v]_\infty^T$ has rank 1 and its kernel is $\mathcal{I}(v)$, the vanishing ideal of v. More generally, if Λ is a conic combination of evaluations at real points, say $\Lambda = \sum_{i=1}^r \lambda_i \Lambda_{v_i}$ where $\lambda_i > 0$ and $v_i \in \mathbb{R}^n$ are pairwise distinct, then $M(\Lambda) = \sum_{i=1}^r \lambda_i [v_i]_\infty [v_i]_\infty^T$ has rank r and its kernel is $\mathcal{I}(v_1, \ldots, v_r)$, the vanishing ideal of the v_i's.*

The following theorem of Curto and Fialkow [7] shows that any positive linear form Λ with a finite rank moment matrix is a conic combination of evaluations at real points, thus implying that the implication of Lemma 3 holds as an equivalence. This result will play a crucial role in our approach. We give a proof, based on [21], although some details are simplified.

Theorem 6. (Finite rank moment matrix theorem) *[7] Assume that $\Lambda \in \mathbb{R}[x]^*$ is positive with $\operatorname{rank} M(\Lambda) =: r < \infty$. Then, $\Lambda = \sum_{i=1}^r \lambda_i \Lambda_{v_i}$ for some distinct $v_1, \ldots, v_r \in \mathbb{R}^n$ and some scalars $\lambda_i > 0$. Moreover, $\{v_1, \ldots, v_r\} = V_{\mathbb{C}}(\operatorname{Ker} M(\Lambda))$.*

Proof. By Lemma 1, $J := \operatorname{Ker} M(\Lambda)$ is a real radical ideal and, by Lemma 2 (combined with Theorem 2), J is zero-dimensional with $\dim \mathbb{R}[x]/J = r$. Therefore, $|V_{\mathbb{C}}(J)| = r$ and $V_{\mathbb{C}}(J) \subseteq \mathbb{R}^n$. Say,

$$V_{\mathbb{C}}(J) = \{v_1, \ldots, v_r\} \subseteq \mathbb{R}^n$$

so that $J = \mathcal{I}(v_1, \ldots, v_r)$ is the vanishing ideal of the v_i's. Let $p_1, \ldots, p_r$ be interpolation polynomials at $v_1, \ldots, v_r$, respectively, that is, $p_i(v_j) = 1$ if $i = j$ and 0 otherwise. We first claim:

$$\text{The set } \{p_1, \ldots, p_r\} \text{ forms a basis of the quotient space } \mathbb{R}[x]/J.$$

Indeed if, for some scalars λ_i, the polynomial $\sum_{i=1}^r \lambda_i p_i$ vanishes at all v_i's, then $\lambda_i = 0 \ \forall i$. Hence $\{p_1, \ldots, p_r\}$ is linearly independent in $\mathbb{R}[x]/J$ and thus a basis, since $r = \dim \mathbb{R}[x]/J$. Consider the linear form

$$\Lambda' := \sum_{i=1}^r \Lambda(p_i^2) \Lambda_{v_i}.$$

We claim that $\Lambda = \Lambda'$. As both Λ and Λ' vanish on the ideal J, it suffices to show that Λ and Λ' take the same values at all members of the basis $\{p_1, \ldots, p_r\}$ of $\mathbb{R}[x]/J$. Indeed, $\Lambda'(p_j) = \Lambda(p_j^2)$, and $\Lambda(p_j) = \Lambda(p_j^2)$ as well since $p_j - p_j^2 \in J$ belongs to J. $\square$

Example 6. Consider the linear form $\Lambda = \frac{1}{2}\Lambda_{(0,0)} + \frac{1}{2}\Lambda_{(1,2)} \in \mathbb{R}[x]^*$, with moment matrix (indexed by $1, x_1, x_2, x_1^2, \dots$):

$$M(\Lambda) = \begin{pmatrix} 1 & \frac{1}{2} & 1 & \frac{1}{2} & \cdots \\ \frac{1}{2} & \frac{1}{2} & 1 & \frac{1}{2} & \cdots \\ 1 & 1 & 2 & 1 & \cdots \\ \frac{1}{2} & \frac{1}{2} & 1 & \frac{1}{2} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} = \frac{1}{2}[v_1]_\infty [v_1]_\infty^T \Big|_{v_1=(0,0)} + \frac{1}{2}[v_2]_\infty [v_2]_\infty^T \Big|_{v_2=(1,2)}$$

Note e.g. that the 2nd and 4th columns of $M(\Lambda)$ coincide, yielding the polynomial $g_1 = -x_1 + x_1^2$ in the kernel of $M(\Lambda)$. In fact, the polynomials g_1, $g_2 = -2x_1 + x_2$, $g_3 = -2x_1 + x_1x_2$ provide a basis of the real radical ideal $\operatorname{Ker} M(\Lambda)$, whose variety is $V_{\mathbb{C}}(\operatorname{Ker} M(\Lambda)) = \{(0,0),(1,2)\} \subseteq \mathbb{R}^2$.

As background information we mention (without proof) the following characterization for the linear forms $\Lambda \in \mathbb{R}[x]^*$ with a finite rank moment matrix. When positivity is dropped, the evaluations at points $v \in V_{\mathbb{C}}(\Lambda)$ do not suffice, one also needs the more general differential operators ∂_v^α (defined in (4)).

Theorem 7. *(see [9, Thm 2.2.7], [10, Chap. 7]) Let $\Lambda \in \mathbb{R}[x]^*$ satisfying $\operatorname{rank} M(\Lambda) < \infty$. Say, $V_{\mathbb{C}}(\operatorname{Ker} M(\Lambda)) = \{v_1, \dots, v_r\}$, so that $r \le \operatorname{rank} M(\Lambda)$. Then,*

$$\Lambda = \sum_{i=1}^{r} \sum_{\alpha \in A_i} a_{\alpha,i} \partial_{v_i}^\alpha,$$

where $A_i \subseteq \mathbb{N}^n$ are finite and $a_{\alpha,i} \in \mathbb{R} \setminus \{0\}$. Moreover, $\operatorname{Ker} M(\Lambda)$ is radical if and only if

$$\Lambda = \sum_{i=1}^{r} a_i \Lambda_{v_i},$$

where $a_i \neq 0$ (i.e., $A_i = \{0\}$ $\forall i$). Moreover, $\operatorname{Ker} M(\Lambda)$ is real radical precisely when $\{v_1, \dots, v_r\} \subseteq \mathbb{R}^n$.

Excursion: Why is $M(\Lambda)$ called a moment matrix? We briefly recall how the matrices $M(\Lambda)$ arise naturally in the context of the classical moment problem in mathematics (cf. e.g. [1]). Given a finite positive Borel measure μ on $\mathbb{R}^n$, the quantity

$$\int_{\mathbb{R}^n} x^\alpha d\mu$$

is called its *moment of order* $\alpha \in \mathbb{N}^n$, and the sequence $y_\mu = (\int x^\alpha d\mu)_{\alpha \in \mathbb{N}^n}$ is called its *moment sequence*. The *moment problem* asks to characterize the sequences $y \in \mathbb{R}^{\mathbb{N}^n}$ that are the sequence of moments of some finite positive Borel measure on (some subset of) $\mathbb{R}^n$ or, equivalently, to characterize the linear forms $\Lambda \in \mathbb{R}[x]^*$ of the form

$$\Lambda = \Lambda_\mu(p) := \int p(x)d\mu \quad \text{for } p \in \mathbb{R}[x], \tag{8}$$

(then we also say that μ is a *representing measure* for Λ). A well known result of Haviland [12] claims that $\Lambda = \Lambda_\mu$ for some finite positive measure μ if and only if $\Lambda(p) \geq 0$ for all polynomials p that are nonnegative on $\mathbb{R}^n$. However, except in some exceptional cases[4] no characterization is known for the nonnegative polynomials on $\mathbb{R}^n$. Yet we find the following well known necessary condition: If Λ has a representing measure, then Λ is positive (recall Definition 2), i.e., $M(\Lambda) \succeq 0$.

Positivity of Λ is in general only a necessary condition for existence of a representing measure. However, the above result of Curto and Fialkow (Theorem 6) shows equivalence in the case when $M(\Lambda)$ has finite rank, in which case the measure μ is finite atomic with support $V_{\mathbb{C}}(\operatorname{Ker} M(\Lambda))$.

When $\mu = \delta_v$ is the Dirac measure at a point $v \in \mathbb{R}^n$, its moment sequence is $y_\mu = [v]_\infty$ with corresponding linear form $\Lambda_\mu = \Lambda_v$, the evaluation at v. More generally, when μ is finitely atomic, i.e., of the form $\mu = \sum_{i=1}^r \lambda_i \delta_{v_i}$ with finite support $\{v_1, \ldots, v_r\} \subseteq \mathbb{R}^n$, then its moment sequence is $y_\mu = \sum_{i=1}^r \lambda_i [v_i]_\infty$ with corresponding linear form $\Lambda_\mu = \sum_{i=1}^r \lambda_i \Lambda_{v_i}$.

Characterizing real radical ideals using positive linear forms on $\mathbb{R}[x]$. We now combine the above results to obtain a semidefinite characterization of real radical ideals using positive linear forms. For this define the convex set

$$\mathcal{K} = \{\Lambda \in \mathbb{R}[x]^* \mid \Lambda(1) = 1, M(\Lambda) \succeq 0 \text{ and } \Lambda(p) = 0 \ \forall p \in I\}. \tag{9}$$

For any $\Lambda \in \mathcal{K}$, $\operatorname{Ker} M(\Lambda)$ is a real radical ideal, which contains I and thus its real radical $\sqrt[\mathbb{R}]{I}$. This implies:

$$\dim \mathbb{R}[x]/\operatorname{Ker} M(\Lambda) \leq \dim \mathbb{R}[x]/\sqrt[\mathbb{R}]{I}.$$

When the real variety $V_{\mathbb{R}}(I)$ is finite, $\mathbb{R}[x]/\sqrt[\mathbb{R}]{I}$ has finite dimension as a vector space, equal to $|V_{\mathbb{R}}(I)|$, and thus $\operatorname{Ker} M(\Lambda)$ is zero-dimensional with

$$\operatorname{rank} M(\Lambda) = \dim \mathbb{R}[x]/\operatorname{Ker} M(\Lambda) \leq \dim \mathbb{R}[x]/\sqrt[\mathbb{R}]{I} = |V_{\mathbb{R}}(I)|.$$

(Recall Lemma 2). For the following element $\Lambda = \frac{1}{|V_{\mathbb{R}}(I)|} \sum_{v \in V_{\mathbb{R}}(I)} \Lambda_v$ in $\mathcal{K}$ we have $\operatorname{rank} M(\Lambda) = |V_{\mathbb{R}}(I)|$, motivating the following definition:

[4] A celebrated result of Hilbert (cf. e.g. [2]) shows that there are three sets of parameters (n, d) for which the following equivalence holds: For any polynomial p in n variables and degree $2d$, p is nonnegative on $\mathbb{R}^n$ if and only if p can be written as a sum of squares of polynomials. These parameters are $(n = 1, d)$ (univariate polynomials), $(n, d = 1)$ (quadratic polynomials), and $(n = 3, d = 2)$ (ternary quartic polynomials). In all other cases there are polynomials that are nonnegative on $\mathbb{R}^n$ but cannot be written as a sum of squares of polynomials.

Definition 4. (Generic linear forms) *Let $\mathcal{K}$ be defined as in (9) and assume $|V_{\mathbb{R}}(I)| < \infty$. A linear form $\Lambda \in \mathcal{K}$ is said to be* generic *if $M(\Lambda)$ has maximum rank, i.e., if* $\operatorname{rank} M(\Lambda) = |V_{\mathbb{R}}(I)|$.

A simple geometric property of positive semidefinite matrices yields the following equivalent definition for generic elements of $\mathcal{K}$. This is in fact the key tool used in [18] for computing the real radical ideal $\sqrt[\mathbb{R}]{I}$.

Lemma 4. *Assume $|V_{\mathbb{R}}(I)| < \infty$. An element $\Lambda \in \mathcal{K}$ is generic if and only if* $\operatorname{Ker} M(\Lambda) \subseteq \operatorname{Ker} M(\Lambda')$ *for all $\Lambda' \in \mathcal{K}$. Moreover,* $\operatorname{Ker} M(\Lambda) = \sqrt[\mathbb{R}]{I}$ *for all generic $\Lambda \in \mathcal{K}$.*

Proof. Say $r = |V_{\mathbb{R}}(I)|$, $V_{\mathbb{R}}(I) = \{v_1, \ldots, v_r\}$. Assume first $\operatorname{rank} M(\Lambda) = r$. Let $\Lambda' \in \mathcal{K}$. As $\Lambda + \Lambda' \in \mathcal{K}$, we have

$$\operatorname{Ker} M(\Lambda + \Lambda') = \operatorname{Ker} M(\Lambda) \cap \operatorname{Ker} M(\Lambda') \subseteq \operatorname{Ker} M(\Lambda),$$

implying $r \geq \operatorname{rank} M(\Lambda + \Lambda') \geq \operatorname{rank} M(\Lambda)$. Hence equality holds throughout which implies $\operatorname{Ker} M(\Lambda) = \operatorname{Ker} M(\Lambda) \cap \operatorname{Ker} M(\Lambda') \subseteq \operatorname{Ker} M(\Lambda')$.

Conversely, assume $\operatorname{Ker} M(\Lambda) \subseteq \operatorname{Ker} M(\Lambda')$ for all $\Lambda' \in \mathcal{K}$. Consider $\Lambda' = \sum_{i=1}^{r} \Lambda_{v_i} \in \mathcal{K}$ whose kernel is $\mathcal{I}(v_1, \ldots, v_r)$. This implies $\operatorname{Ker} M(\Lambda) \subseteq \mathcal{I}(v_1, \ldots, v_r)$ and thus

$$\operatorname{rank} M(\Lambda) = \dim \mathbb{R}[x]/\operatorname{Ker} M(\Lambda) \geq \dim \mathbb{R}[x]/\mathcal{I}(v_1, \ldots, v_r) = r.$$

Hence, $\operatorname{rank} M(\Lambda) = r$ and $\operatorname{Ker} M(\Lambda) = \mathcal{I}(v_1, \ldots, v_r) = \sqrt[\mathbb{R}]{I}$ (using the Real Nullstellensatz, Theorem 1 (ii), for the last equality). □

Example 7 (Example 6 cont.). Consider the set $\mathcal{K}$ corresponding to the ideal $I = \langle h_1, h_2, h_3 \rangle \subseteq \mathbb{R}[x_1, x_2]$, where

$$\begin{aligned} h_1 &= x_2^4 x_1 + 3x_1^3 - x_2^4 - 3x_1^2, \\ h_2 &= x_1^2 x_2 - 2x_1^2, \\ h_3 &= 2x_2^4 x_1 - x_1^3 - 2x_2^4 + x_1^2. \end{aligned}$$

Then, $\Lambda = \frac{1}{2}\Lambda_{(0,0)} + \frac{1}{2}\Lambda_{(1,2)}$ is a generic element of $\mathcal{K}$. Thus the real radical ideal of I is $\sqrt[\mathbb{R}]{I} = \operatorname{Ker} M(\Lambda) = \langle g_1, g_2, g_3 \rangle$, with g_1, g_2, g_3 as in Example 6.

3.2 Truncated positive linear forms and real radical ideals

In view of the results in the previous section (in particular, Lemmas 2 and 4), the task of finding the real radical ideal $\sqrt[\mathbb{R}]{I}$ as well as a linear basis of the quotient space $\mathbb{R}[x]/\sqrt[\mathbb{R}]{I}$ can be reduced to finding a generic linear form Λ in the set $\mathcal{K}$ (defined in (9)). In order to be able to deal with such linear forms computationally, we will work with linear forms on finite dimensional truncations $\mathbb{R}[x]_s$ of the polynomial ring. Given $\Lambda \in (\mathbb{R}[x]_{2s})^*$, we can define the quadratic form:

$$Q_\Lambda : f \in \mathbb{R}[x]_s \mapsto Q_\Lambda(f) = \Lambda(f^2),$$

whose matrix

$$M_s(\Lambda) = (\Lambda(x^\alpha x^\beta))_{\alpha,\beta\in\mathbb{N}^n_s}$$

in the monomial basis of $\mathbb{R}[x]_s$ is called the *truncated moment matrix of order s of* Λ. Thus Λ is positive (i.e., $\Lambda(f^2) \geq 0 \ \ \forall f \in \mathbb{R}[x]_s$) if and only if $M_s(\Lambda) \succeq 0$. Again we identify the kernels of Q_Λ and of $M_s(\Lambda)$ (by identifying polynomials with their coefficient sequences) and view $\operatorname{Ker} M_s(\Lambda)$ as a subset of $\mathbb{R}[x]_s$.

Flat extensions of moment matrices. We now present the following crucial result of Curto and Fialkow [7] for flat extensions of moment matrices.

Theorem 8. *[7]* **(Flat extension theorem)** *Let* $\Lambda \in (\mathbb{R}[x]_{2s})^*$ *and assume that* $M_s(\Lambda)$ *is a* flat extension *of* $M_{s-1}(\Lambda)$*, i.e.,*

$$\operatorname{rank} M_s(\Lambda) = \operatorname{rank} M_{s-1}(\Lambda). \tag{10}$$

Then one can extend (uniquely) Λ *to* $\tilde{\Lambda} \in (\mathbb{R}[x]_{2s+2})^*$ *in such a way that* $M_{s+1}(\tilde{\Lambda})$ *is a flat extension of* $M_s(\Lambda)$*; thus* $\operatorname{rank} M_{s+1}(\tilde{\Lambda}) = \operatorname{rank} M_s(\Lambda)$.

The proof is elementary and relies on the following basic lemma showing that the kernel of a truncated moment matrix behaves like a 'truncated ideal'.

Lemma 5. *Let* $\Lambda \in (\mathbb{R}[x]_{2s})^*$ *and* $f, g \in \mathbb{R}[x]$ *with* $f \in \operatorname{Ker} M_s(\Lambda)$.

(i) *Assume* $\operatorname{rank} M_s(\Lambda) = \operatorname{rank} M_{s-1}(\Lambda)$. *Then* $\operatorname{Ker} M_{s-1}(\Lambda) \subseteq \operatorname{Ker} M_s(\Lambda)$ *and* $fg \in \operatorname{Ker} M_s(\Lambda)$ *if* $\deg(fg) \leq s$.

(ii) *Assume* $M_s(\Lambda) \succeq 0$. *Then* $\operatorname{Ker} M_{s-1}(\Lambda) \subseteq \operatorname{Ker} M_s(\Lambda)$ *and* $fg \in \operatorname{Ker} M_s(\Lambda)$ *if* $\deg(fg) \leq s-1$.

Indeed, using property (10) and Lemma 5 (i), we see that for every monomial m of degree s, there exists a polynomial of the form $f_m = m + r_m \in \operatorname{Ker} M_s(\Lambda)$, where $r_m \in \mathbb{R}[x]_{s-1}$. If an extension $\tilde{\Lambda}$ exists, then all the polynomials $f_m, x_i f_m$ must lie in the kernel of $M(\tilde{\Lambda})$ and can be used to determine the values of $\tilde{\Lambda}$ on the monomials of degree $2s+1$ and $2s+2$. The main work consists of verifying the consistency of this construction. We also refer to [22] for a detailed exposition. This result has been generalized in [23] to moment matrices indexed by an arbitrary (connected to 1) set of monomials (instead of all monomials up to a given degree s). These flat extension theorems play a crucial role in the moment matrix approach as they allow to deduce information about the infinite moment matrix $M(\Lambda)$ from its finite section $M_s(\Lambda)$.

Theorem 9. *[18] Let* $\Lambda \in (\mathbb{R}[x]_{2s})^*$ *and assume that (10) holds. Then one can extend* Λ *to* $\tilde{\Lambda} \in \mathbb{R}[x]^*$ *in such a way that* $M(\tilde{\Lambda})$ *is a flat extension of* $M_s(\Lambda)$*, and the ideal* $\operatorname{Ker} M(\tilde{\Lambda})$ *is generated by* $\operatorname{Ker} M_s(\Lambda)$*, i.e.,*

$$\operatorname{rank} M(\tilde{\Lambda}) = \operatorname{rank} M_s(\Lambda) \quad \textit{and} \quad \operatorname{Ker} M(\tilde{\Lambda}) = \langle \operatorname{Ker} M_s(\Lambda) \rangle.$$

Moreover, any monomial set $\mathcal{B}$ *indexing a basis of the column space of* $M_{s-1}(\Lambda)$ *is a basis of the quotient space* $\mathbb{R}[x]/\operatorname{Ker} M(\tilde{\Lambda})$. *If, moreover,* $M_s(\Lambda) \succeq 0$*, then the ideal* $\langle \operatorname{Ker} M_s(\Lambda) \rangle$ *is real radical and* Λ *is of the form* $\Lambda = \sum_{i=1}^r \lambda_i \Lambda_{v_i}$*, where* $\lambda_i > 0$ *and* $\{v_1, \ldots, v_r\} = V_{\mathbb{C}}(\operatorname{Ker} M_s(\Lambda)) \subseteq \mathbb{R}^n$.

Proof. The existence of $\tilde{\Lambda}$ follows by applying iteratively Theorem 8 and the inclusion $\langle \operatorname{Ker} M_s(\Lambda)\rangle \subseteq \operatorname{Ker} M(\tilde{\Lambda})$ follows using Lemma 5 (i).

If $\mathcal{B}$ be a set of monomials indexing a column basis of $M_{s-1}(\Lambda)$, then $\mathcal{B}$ is also a column basis of $M(\tilde{\Lambda})$ and thus a basis of $\mathbb{R}[x]/\operatorname{Ker} M(\tilde{\Lambda})$ (by Lemma 2). One can verify the direct sum decomposition $\mathbb{R}[x] = \operatorname{Span}_{\mathbb{R}}(\mathcal{B}) \oplus \langle \operatorname{Ker} M_s(\Lambda)\rangle$, which implies $\operatorname{Ker} M(\tilde{\Lambda}) = \langle \operatorname{Ker} M_s(\Lambda)\rangle$. Finally, $M_s(\Lambda) \succeq 0$ implies that $M(\tilde{\Lambda}) \succeq 0$, so that $\langle \operatorname{Ker} M_s(\Lambda)\rangle = \operatorname{Ker} M(\tilde{\Lambda})$ is real radical (by Lemma 1). The final statement follows directly applying Theorem 6 to $\tilde{\Lambda}$. □

Example 8 (Example 6 cont.). Consider the linear form $\Lambda \in \mathbb{R}[x]^*$ in Example 6. Recall that $\operatorname{Ker} M(\Lambda)$ is generated by $g_1, g_2, g_3 \in \mathbb{R}[x]_2$. First note that these polynomials imply the rank condition: $\operatorname{rank} M_2(\Lambda) = \operatorname{rank} M_1(\Lambda)$ and thus permit to construct $M_2(\Lambda)$ from $M_1(\Lambda)$. Moreover, they permit to recover the infinite matrix $M(\Lambda)$ from its submatrix $M_1(\Lambda)$. For instance, since $x_1^2x_2 = x_2(x_1+g_1) = 2x_1+g_3+g_1x_2$ and $g_1, g_2, g_3 \in \operatorname{Ker} M_2(\Lambda) \subseteq \operatorname{Ker} M(\Lambda)$, we deduce that the column of $M(\Lambda)$ indexed by $x_1^2x_2$ is equal to twice its column indexed by x_1. Using the fact that $\operatorname{Ker} M(\Lambda) = \langle \operatorname{Ker} M_2(\Lambda)\rangle$, we can analogously define iteratively all columns of $M(\Lambda)$.

Computing real radical ideals using truncated positive linear forms on $\mathbb{R}[x]_t$. We saw above how to use positive linear forms on $\mathbb{R}[x]$ to characterize the real radical ideal $\sqrt[\mathbb{R}]{I}$. We now combine this characterization with the above results about flat extensions of truncated moment matrices to obtain a practical algorithm for computing $\sqrt[\mathbb{R}]{I}$ operating on finite dimensional subspaces $\mathbb{R}[x]_t \subseteq \mathbb{R}[x]$ only. As before $I = \langle h_1, \ldots, h_m\rangle$ is the ideal generated by the polynomial equations h_i to be solved. For $t \in \mathbb{N}$, define the set

$$\mathcal{H}_t = \{h_i x^\alpha \mid i = 1, \ldots, m,\ |\alpha| \le t - \deg(h_i)\} \tag{11}$$

of prolongations up to degree t of the polynomials h_i, and the truncated analogue of the set $\mathcal{K}$:

$$\mathcal{K}_t = \{\Lambda \in (\mathbb{R}[x]_t)^* \mid \Lambda(1) = 1,\ M_{\lfloor t/2\rfloor}(\Lambda) \succeq 0 \text{ and } \Lambda(f) = 0\ \forall f \in \mathcal{H}_t\}. \tag{12}$$

Note that the constraint: $\Lambda(f) = 0\ \forall f \in \mathcal{H}_t$ (i.e., $\Lambda \in \mathcal{H}_t^\perp$) corresponds to the constraint (2) of Section 1.2. As the convex set $\mathcal{K}_t$ is described by the positive semidefiniteness of an affinely parametrized matrix, it is an instance of a *spectrahedron*. The following lemma is the truncated analogue of Lemma 4.

Lemma 6. (Generic truncated linear forms) *The following assertions are equivalent for $\Lambda \in (\mathbb{R}[x]_t)^*$:*

(i) $\operatorname{rank} M_{\lfloor t/2\rfloor}(\Lambda) \ge \operatorname{rank} M_{\lfloor t/2\rfloor}(\Lambda')$ *for all* $\Lambda' \in \mathcal{K}_t$.
(ii) $\operatorname{Ker} M_{\lfloor t/2\rfloor}(\Lambda) \subseteq \operatorname{Ker} M_{\lfloor t/2\rfloor}(\Lambda')$ *for all* $\Lambda' \in \mathcal{K}_t$.
(iii) *The linear form Λ lies in the relative interior of the convex set $\mathcal{K}_t$.*

Then Λ is called a generic *element of $\mathcal{K}_t$ and the kernel $\mathcal{N}_t = \operatorname{Ker} M_{\lfloor t/2\rfloor}(\Lambda)$ is independent of the particular choice of the generic element $\Lambda \in \mathcal{K}_t$.*

Theorem 10. *We have:* $\mathcal{N}_t \subseteq \mathcal{N}_{t+1} \subseteq \ldots \subseteq \sqrt[\mathbb{R}]{I}$, *with equality* $\sqrt[\mathbb{R}]{I} = \langle \mathcal{N}_t \rangle$ *for* t *large enough.*

Proof. Let $\Lambda \in \mathcal{K}_{t+1}$ be generic. Its restriction to $(\mathbb{R}[x]_t)^*$ lies in $\mathcal{K}_t$, implying

$$\mathcal{N}_{t+1} = \operatorname{Ker} M_{\lfloor (t+1)/2 \rfloor}(\Lambda) \supseteq \operatorname{Ker} M_{\lfloor t/2 \rfloor}(\Lambda) \supseteq \mathcal{N}_t.$$

Now let Λ be a generic element of $\mathcal{K}_t$ so that $\mathcal{N}_t = \operatorname{Ker} M_{\lfloor t/2 \rfloor}(\Lambda)$. The inclusion: $\mathcal{N}_t \subseteq \mathcal{I}(V_{\mathbb{R}}(I))$ follows using Lemma 6 (ii). Indeed, $\Lambda_v \in \mathcal{K}_t$ for all $v \in V_{\mathbb{R}}(I)$, which implies that $\operatorname{Ker} M_{\lfloor t/2 \rfloor}(\Lambda) \subseteq \operatorname{Ker} M_{\lfloor t/2 \rfloor}\Lambda_v \subseteq \mathcal{I}(v)$ and thus $\operatorname{Ker} M_{\lfloor t/2 \rfloor}(\Lambda) \subseteq \bigcap_{v \in V_{\mathbb{R}}(I)} \mathcal{I}(v) = \mathcal{I}(V_{\mathbb{R}}(I)) = \sqrt[\mathbb{R}]{I}$ (by the Real Nullstellensatz).

We now show equality: $\sqrt[\mathbb{R}]{I} = \langle \mathcal{N}_t \rangle$ for t large enough. For this, let $\{g_1, \ldots, g_L\}$ be a basis of the ideal $\sqrt[\mathbb{R}]{I}$; we show that $g_l \in \mathcal{N}_t \ \ \forall l$. We have:

$$g_l^{2k} + \sum_j s_j^2 = \sum_{i=1}^{m} u_i h_i \ \text{ for some } k \in \mathbb{N} \text{ and } s_j, u_i \in \mathbb{R}[x].$$

Since $\Lambda \in \mathcal{H}_t^{\perp}$, we have $h_i \in \mathcal{N}_t$ if $t \geq 2\deg(h_i)$. Using Lemma 5 (ii), this implies that, for t large enough, $\mathcal{N}_t$ contains each $u_i h_i$ and thus $g_l^{2k} + \sum_j s_j^2$. In particular, $\Lambda(g_l^{2k} + \sum_j s_j^2) = 0$. On the other hand, $\Lambda(g_l^{2k}), \Lambda(s_j^2) \geq 0$ (since $M_{\lfloor t/2 \rfloor}(\Lambda) \succeq 0$), thus implying $\Lambda(g_l^{2k}) = 0$. An easy induction on k now permits to conclude that $g_l \in \mathcal{N}_t$. $\square$

When $V_{\mathbb{R}}(I)$ is finite, one can guaranty the equality $\sqrt[\mathbb{R}]{I} = \langle \mathcal{N}_t \rangle$ using the rank condition (10). The next results provide all the ingredients of the moment matrix algorithm for real roots, whose description is given in Section 3.3: Theorem 11 will provide a stopping criterion (when $|V_{\mathbb{R}}(I)| < \infty$) and Theorem 12 below will imply its termination, as well as provide a criterion permitting to check the (non-)existence of real roots.

Theorem 11. *[18] Let* $I = \langle h_1, \ldots, h_m \rangle$ *be an ideal in* $\mathbb{R}[x]$, $D = \max_i \deg(h_i)$, *and* $d = \lceil D/2 \rceil$. *Let* $\Lambda \in \mathcal{K}_t$ *be a generic element and assume that at least one of the following two conditions holds:*

$$\operatorname{rank} M_s(\Lambda) = \operatorname{rank} M_{s-1}(\Lambda) \qquad \text{for some } D \leq s \leq \lfloor t/2 \rfloor, \tag{13}$$

$$\operatorname{rank} M_s(\Lambda) = \operatorname{rank} M_{s-d}(\Lambda) \qquad \text{for some } d \leq s \leq \lfloor t/2 \rfloor. \tag{14}$$

Then, $\sqrt[\mathbb{R}]{I} = \langle \operatorname{Ker} M_s(\Lambda) \rangle$, *and any basis of the column space of* $M_{s-1}(\Lambda)$ *is a basis of the quotient space* $\mathbb{R}[x]/\sqrt[\mathbb{R}]{I}$.

Proof. The ideal $J := \langle \operatorname{Ker} M_s(\Lambda) \rangle$ is real radical (by Theorem 9). Moreover,

$$\operatorname{Ker} M_s(\Lambda) \subseteq \operatorname{Ker} M_{\lfloor t/2 \rfloor}(\Lambda) \subseteq \sqrt[\mathbb{R}]{I}$$

(since Λ is generic and using Theorem 10) and thus $J \subseteq \sqrt[\mathbb{R}]{I}$. Remains to show $\sqrt[\mathbb{R}]{I} \subseteq J$. Suppose first that (13) holds. The condition $\Lambda \in \mathcal{H}_t^{\perp}$ implies that

$h_i \in \operatorname{Ker} M_s(\Lambda)$ (since $s + \deg(h_i) \leq t$ as $t \geq 2D$). Thus $I \subseteq J$, implying $\sqrt[\mathbb{R}]{I} \subseteq J$ as J is real radical.

Suppose now that (14) holds. Again from Theorem 9 we know that $V_{\mathbb{C}}(\operatorname{Ker} M_s(\Lambda)) = \{v_1, \ldots, v_r\} \subseteq \mathbb{R}^n$ and $\Lambda = \sum_{i=1}^r \lambda_i \Lambda_{v_i}$ where $\lambda_i > 0$. Let $p_1, \ldots, p_r$ be interpolation polynomials at the v_i's, i.e., such that $p_j(v_i) = \delta_{i,j}$. An easy but crucial observation (made in [21]) is that we may assume that each p_j has degree at most $s - d$. Indeed, we can replace each interpolation polynomial p_j by its normal form modulo J with respect to a basis of $\mathbb{R}[x]/J$; such a basis can be obtained by picking a column basis of $M_{s-d}(\Lambda)$, its members are thus monomials of degree at most $s - d$, and their resulting normal forms are again interpolation polynomials at the v_i's. As $\deg(p_j^2) \leq 2(s-d) \leq t - 2d \leq t - \deg(h_i)$, we can claim that $\Lambda(p_i^2 h_i) = 0$ and in turn $0 = \Lambda(p_j^2 h_i) = \sum_{l=1}^r \lambda_l p_j^2(v_l) h_i(v_l) = \lambda_j h_i(v_j)$. Since $h_i(v_j) = 0$ for all i, j, we conclude that $\{v_1, \ldots, v_r\} \subseteq V_{\mathbb{R}}(I)$, implying the desired inclusion $\sqrt[\mathbb{R}]{I} = \mathcal{I}(V_{\mathbb{R}}(I)) \subseteq \mathcal{I}(v_1, \ldots, v_r) = J$. □

Theorem 12. *[18] Let I be an ideal in $\mathbb{R}[x]$.*

(i) *If $V_{\mathbb{R}}(I) = \emptyset$, then $\mathcal{K}_t = \emptyset$ for t large enough.*
(ii) *If $1 \leq |V_{\mathbb{R}}(I)| < \infty$ then, for t large enough, there exists an integer s for which (14) holds for all $\Lambda \in \mathcal{K}_t$.*

Proof. Let $\{g_1, \ldots, g_L\}$ be a Gröbner basis of $\sqrt[\mathbb{R}]{I}$ with respect to a total degree monomial ordering, and let $\mathcal{B}$ be the corresponding set of standard monomials, forming a basis of $\mathbb{R}[x]/\sqrt[\mathbb{R}]{I}$. The argument used in the proof of Theorem 10 shows the existence of $t_0 \in \mathbb{N}$ for which $\{g_1, \ldots, g_L\} \subseteq \operatorname{Ker} M_{\lfloor t/2 \rfloor}(\Lambda)$ for all $t \geq t_0$ and $\Lambda \in \mathcal{K}_t$.
(i) If $V_{\mathbb{R}}(I) = \emptyset$, then $\{1\}$ is a basis of $\sqrt[\mathbb{R}]{I} = \mathbb{R}[x]$. Thus $1 \in \operatorname{Ker} M_{\lfloor t/2 \rfloor}(\Lambda)$, implying $\Lambda(1) = 0$ if $\Lambda \in \mathcal{K}_t$, contradicting $\Lambda(1) = 1$ and thus showing $\mathcal{K}_t = \emptyset$.
(ii) As $V_{\mathbb{R}}(I)$ is finite, $s := 1 + \max_{b \in \mathcal{B}} \deg(b)$ is well defined. Choose $t \geq t_0$ and $s < \lfloor t/2 \rfloor$. For $\alpha \in \mathbb{N}^n_s$, decompose x^α as

$$x^\alpha = \sum_{b \in \mathcal{B}} \lambda_b b + \sum_{l=1}^L u_l g_l \in \operatorname{Span}_{\mathbb{R}}(\mathcal{B}) \oplus \sqrt[\mathbb{R}]{I},$$

where $\lambda_b \in \mathbb{R}$, $u_l \in \mathbb{R}[x]$, $\deg(\sum_b \lambda_b b) \leq s - 1$, and $\deg(u_l g_l) \leq s < \lfloor t/2 \rfloor$ (as the g_l's form a Gröbner basis for a *total degree ordering*, we can claim $\deg(u_l g_l) \leq s$). As $g_l \in \operatorname{Ker} M_{\lfloor t/2 \rfloor}(\Lambda)$ and $\deg(u_l g_l) < \lfloor t/2 \rfloor$, we also have that $u_l g_l \in \operatorname{Ker} M_{\lfloor t/2 \rfloor}(\Lambda)$ (recall Lemma 5). Hence, $x^\alpha - \sum_{b \in \mathcal{B}} \lambda_b b \in \operatorname{Ker} M_{\lfloor t/2 \rfloor}(\Lambda)$, which shows that $M_s(\Lambda)$ is a flat extension of $M_{s-d}(\Lambda)$. □

Example 9 (Example 7 cont.). Consider again the ideal $I = (h_1, h_2, h_3)$ from Example 7. Then, $D = 5$, $\dim \mathbb{R}[x]/I = 9$, and the variety $V_{\mathbb{C}}(I)$ consists of two real points, one of them with multiplicity eight. Table 1 shows the ranks of the moment matrix $M_s(\Lambda)$ for generic $\Lambda \in \mathcal{K}_t$. The rank condition holds at order $(t, s) = (6, 2)$. Then we can extract the two roots $v_1 = (0, 0)$ and

$v_2 = (1, 2)$ as well as the (border) basis $\{g_1, g_2, g_3\}$ of $\sqrt[\mathbb{R}]{I}$ (already discussed in Example 7). This is possible although here $s = 2$ is strictly smaller than $d = 3$ and $D = 5$; indeed, in view of Theorem 5, we can simply check whether the formal multiplication matrices commute and whether $h_i(v) = 0$ for all $i = 1, \ldots, m$ and $v \in V_{\mathbb{C}}(\operatorname{Ker} M_s(\Lambda))$.

$s =$	0	1	2	3
$t = 5$	1	3	5	—
$t = 6$	1	**2**	**2**	4

Table 1. Ranks of $M_s(\Lambda)$ for generic $\Lambda \in \mathcal{K}_t$ in Example 9.

As a byproduct of Theorem 11, we see that the rank conditions (13) or (14) also imply a full description of the convex hull of the variety $V_{\mathbb{R}}(I)$. Indeed, under (13), we can apply Theorem 9 to deduce that, for any $\Lambda \in \mathcal{K}_t$, its restriction $\pi_{2s}(\Lambda)$ can be written as a conic combination of evaluations at points of $V_{\mathbb{R}}(I)$. Combining with Theorem 12 (and representing linear forms $\Lambda \in (\mathbb{R}[x]_t)^*$ by their coefficients $y \in \mathbb{R}^{\mathbb{N}^n_t}$ as $\Lambda = \sum_\alpha y_\alpha \partial_0^\alpha$), we obtain:

Corollary 1. *Assume $|V_{\mathbb{R}}(I)| < \infty$. For some integers $1 \le s \le \lfloor t/2 \rfloor$, the projection onto $\mathbb{R}^{\mathbb{N}^n_{2s}}$ of the set*

$$\{y \in \mathbb{R}^{\mathbb{N}^n_t} \mid M_{\lfloor t/2 \rfloor}(y) \succeq 0,\ y_0 = 1,\ \operatorname{vec}(h)^T y = 0\ \forall h \in \mathcal{H}_t\}$$

is equal to the convex hull of the set $\{[v]_{2s} \mid v \in V_{\mathbb{R}}(I)\}$.

Gouveia et al. [11] consider in detail the problem of characterizing the convex hull of a real variety $V_{\mathbb{R}}(I)$. There are several connections with the characterization above. Roughly speaking, this idea can be cast within the more general realm of polynomial optimization (see Section 4.1); however, while we work here with truncated sections of the ideal I, [11] deals with linear forms on the full quotient space $\mathbb{R}[x]/I$. Moreover the points of view and emphasis are different in both publications: we focus here on algorithms to compute real roots, while [11] focuses on relaxation methods for the convex hull of a real variety $V_{\mathbb{R}}(I)$.

3.3 The moment matrix algorithm for computing real roots

We now describe the moment matrix algorithm for computing real roots, summarized in Algorithm 1 below.

Theorem 11 shows the correctness of the algorithm (i.e., equality $J = \sqrt[\mathbb{R}]{I}$) and Theorem 12 shows its termination. Algorithm 1 consists of four main parts, which we now briefly discuss; we refer to [18, 36] for additional details.

Algorithm 1 *The moment matrix algorithm for $V_{\mathbb{R}}(I)$*

Input: Generators $h_1, \ldots, h_m$ of some ideal $I = \langle h_1, \ldots, h_m \rangle$ with $|V_{\mathbb{R}}(I)| < \infty$.
Output: A basis of the ideal $\sqrt[\mathbb{R}]{I}$, a basis of $\mathbb{R}[x]/\sqrt[\mathbb{R}]{I}$, and the set $V_{\mathbb{R}}(I)$.
1: Set $t = D$.
2: Find a generic element $\Lambda \in \mathcal{K}_t$.
3: Check if (13) holds for some $D \leq s \leq \lfloor t/2 \rfloor$,
or if (14) holds for some $d \leq s \leq \lfloor t/2 \rfloor$.
4: **if** yes **then**
5: Set $J = \langle \operatorname{Ker} M_s(\Lambda) \rangle$.
6: Compute a basis $\mathcal{B} \subseteq \mathbb{R}[x]_{s-1}$ of the column space of $M_{s-1}(\Lambda)$.
7: Compute the multiplication matrices $\mathcal{X}_i$ in $\mathbb{R}[x]/J$.
8: Compute a basis $g_1, \ldots, g_l \in \mathbb{R}[x]_s$ of the ideal J.
9: **return** the basis $\mathcal{B}$ of $\mathbb{R}[x]/J$ and the generators $g_1, \ldots, g_l$ of J.
10: **else**
11: Iterate (go to Step 2) replacing t by $t + 1$.
12: **end if**
13: Compute $V_{\mathbb{R}}(I) = V_{\mathbb{C}}(J)$ (via the eigenvalues/eigenvectors of the multiplication matrices $\mathcal{X}_i$).

(i) Finding a generic element in $\mathcal{K}_t$. The set $\mathcal{K}_t$ can be represented as the feasible region of a semidefinite program and we have to find a point lying in its relative interior. Such a point can be found by solving several semidefinite programs with an arbitrary SDP solver (cf. [18, Remark 4.15]), or by solving a single semidefinite program with an interior-point algorithm using a self-dual embedding technique (see, e.g., [8], [44]). Indeed consider the semidefinite program:

$$\min_{\Lambda \in (\mathbb{R}[x]_t)^*} 1 \text{ such that } \begin{array}{l} \Lambda(1) = 1,\ M_{\lfloor t/2 \rfloor}(\Lambda) \succeq 0, \\ \Lambda(h_i x^\alpha) = 0\ \forall i\ \forall |\alpha| \leq t - \deg(h_i), \end{array} \tag{15}$$

whose dual reads:

$$\max \lambda \text{ such that } \begin{array}{l} 1 - \lambda = s + \sum_{i=1}^m u_i h_i \text{ where } s, u_i \in \mathbb{R}[x], \\ s \text{ is a sum of squares, } \deg(s), \deg(u_i h_i) \leq t. \end{array} \tag{16}$$

The feasible region of (15) is the set $\mathcal{K}_t$, as well as its set of optimal solutions, since we minimize a constant objective function over $\mathcal{K}_t$. There is no duality gap, as $\lambda = 1$ is obviously feasible for (16). Solving the program (15) with an interior-point algorithm using a self-dual embedding technique yields[5] either a solution Λ lying in the relative interior of the optimal face (i.e., a generic element of $\mathcal{K}_t$), or a certificate that (15) is infeasible thus showing $V_{\mathbb{R}}(I) = \emptyset$.

(ii) Computing the ranks of submatrices of $M_t(\Lambda)$. In order to check whether one of the conditions (13) or (14) holds we need to compute the ranks

[5] This follows under certain technical conditions on the semidefinite program, which are satisfied for (15); see [18] for details.

of matrices consisting of numerical values. This computationally challenging task may be done by detecting zero singular values and/or a large decay between two subsequent values.

(iii) Computing a basis $\mathcal{B}$ for the column space of $M_{s-1}(\Lambda)$. The set of monomials $\mathcal{B}$ indexing a maximum nonsingular principle submatrix of $M_s(\Lambda)$ directly reveals a basis of the quotient space $\mathbb{R}[x]/J$ (by Theorem 9). The choice of this basis may influence the numerical stability of the extracted set of solutions and the properties of the border basis of J as well. The options range from a monomial basis obtained using a greedy algorithm or more sophisticated polynomial bases, see [36].

(iv) Computing a basis of J and the formal multiplication matrices. Say $\mathcal{B}$ is the monomial basis (connected to 1) of the column space of $M_{s-1}(\Lambda)$ constructed at the previous step (iii). Under the rank condition (13) or (14), for any $b \in \mathcal{B}$, the monomial $x_i b$ can be written as $x_i b = r_{i,b} + q$, where $r_{i,b} \in \mathrm{Span}_{\mathbb{R}}(\mathcal{B})$ and $q \in \mathrm{Ker}\, M_s(\Lambda)$. These polynomials directly give a (border) basis of J, consisting of the polynomials $\{x_i b - r_{i,b} \mid i \leq n,\ b \in \mathcal{B}\}$ (recall Theorem 5) and thus permit the construction of multiplication matrices and the computation of $V_{\mathbb{C}}(J)$ $(= V_{\mathbb{R}}(I))$.

3.4 Adding positivity to other root finding methods

There are several variations of the moment matrix algorithm which can be applied to compute the real variety $V_{\mathbb{R}}(I)$ under milder assumptions than the rank conditions (13) or (14). First we briefly discuss a variation where the rank condition involves moment matrices indexed by arbitrary monomial sets, and then we relate the rank condition used in the moment matrix algorithm to some more general stopping criterion used in an algorithm of [35] for complex roots (cf. Theorem 14 below). In doing so we see again that positivity (i.e. the positive semidefiniteness of the matrix $M(\Lambda)$) is the crucial tool permitting to distinguish between real and complex elements.

First we observe that we may work with moment matrices indexed by arbitrary monomial sets, instead of all monomials at some truncated degree, leading to a generalization of Theorem 11. More precisely, let Λ be a generic element in $\mathcal{K}_t$. Assume that we can find a monomial set $\mathcal{B}$, connected to 1, indexing a linearly independent set of columns of the moment matrix $M_{\lfloor t/2 \rfloor}(\Lambda)$, and for which the submatrices $M_{\mathcal{B}}(\Lambda)$ and $M_{\mathcal{B}^+}$ (indexed respectively by $\mathcal{B}$ and $\mathcal{B}^+$) satisfy the rank condition:

$$\mathrm{rank}\, M_{\mathcal{B}^+}(\Lambda) = \mathrm{rank}\, M_{\mathcal{B}}(\Lambda).$$

Then we can use the generalized flat extension theorem [23, Thm. 1.4] to show that the ideal $J = \langle \mathrm{Ker}\, M_{\mathcal{B}^+}(\Lambda) \rangle$ is real radical, zero-dimensional, contained in $\sqrt[\mathbb{R}]{I}$, and thus $V_{\mathbb{R}}(I) \subseteq V_{\mathbb{C}}(J)$. Hence, one can compute the variety $V_{\mathbb{C}}(J) \subseteq \mathbb{R}^n$, and select from it the desired real variety $V_{\mathbb{R}}(I)$.

Let us stress again that the positivity condition on the linear form Λ is the essential ingredient that permits to focus solely on the real roots among the complex ones. This is best illustrated by observing (following [19]) that, if we delete the positivity condition in the moment matrix algorithm (Algorithm 1), then the *same* algorithm permits to compute all *complex* roots (assuming their number is finite). In other words, consider the analogue of the set $\mathcal{K}_t$:

$$\mathcal{K}_t^{\mathbb{C}} = \mathcal{H}_t^{\perp} = \{\Lambda \in (\mathbb{R}[x]_t)^* \mid \Lambda(1) = 1 \text{ and } \Lambda(f) = 0 \ \forall f \in \mathcal{H}_t\}, \tag{17}$$

where $\mathcal{H}_t$ is as in (11). Call an element $\Lambda \in \mathcal{K}_t^{\mathbb{C}}$ *generic*[6] if $\operatorname{rank} M_s(\Lambda)$ is maximum for all $s \leq \lfloor t/2 \rfloor$. The moment matrix algorithm for complex roots is analogous to Algorithm 1, with the following small twist: Instead of computing a generic element in the convex set $\mathcal{K}_t$, we have to compute a generic (aka random) element in the linear space $\mathcal{K}_t^{\mathbb{C}}$, thus replacing the semidefinite feasibility problem by a linear algebra computation. We refer to [19] for details on correctness and termination of this new algorithm.

Alternatively one can describe the above situation as follows: the complex analogue of Algorithm 1 is a method for complex root finding, which can be turned into an algorithm for real roots simply by adding the positivity condition on Λ. This suggests that the same recipe could be applied to other algorithms for complex root finding. This is indeed the case, for instance, for the prolongation-projection algorithm of [35], as shown in [20] and as recalled below. Ongoing work [33] studies how to extend the border basis algorithm of [32] to real root finding by incorporating positivity conditions.

We illustrate this idea on the prolongation-projection algorithm for complex roots proposed in [35]. This algorithm relies on Theorem 13 below, for which a detailed elementary treatment can be found in [19, 20]. Let π_s denote the projection from $(\mathbb{R}[x]_t)^*$ onto $(\mathbb{R}[x]_s)^*$, where $\pi_s(\Lambda)$ is the restriction of $\Lambda \in (\mathbb{R}[x]_t)^*$ to $\mathbb{R}[x]_s$, with $s \leq t$, and let $\mathcal{H}^+ = \mathcal{H} \cup x_1\mathcal{H} \cup \ldots \cup x_n\mathcal{H}$ denote the one-degree prolongation of $\mathcal{H} \subseteq \mathbb{R}[x]$.

Theorem 13. *[35] Let $I = \langle h_1, \ldots, h_m \rangle$ be a zero-dimensional ideal in $\mathbb{R}[x]$, $D = \max_i \deg(h_i)$, and $1 \leq s \leq t$. If $\dim \pi_s(\mathcal{H}_t^{\perp}) = 0$ then $V_{\mathbb{C}}(I) = \emptyset$. Assume that $D \leq s \leq t$ and the following condition holds:*

$$\dim \pi_s(\mathcal{H}_t^{\perp}) = \dim \pi_{s-1}(\mathcal{H}_t^{\perp}) = \dim \pi_s((\mathcal{H}_t^+)^{\perp}). \tag{18}$$

Then one can find a monomial basis $\mathcal{B} \subseteq \mathbb{R}[x]_{s-1}$ satisfying

$$\mathbb{R}[x]_{s-1} = \operatorname{Span}_{\mathbb{R}}(\mathcal{B}) \oplus (\mathbb{R}[x]_{s-1} \cap \operatorname{Span}_{\mathbb{R}}(\mathcal{H}_t))$$

and which is a basis of $\mathbb{R}[x]/I$. This in turn allows to construct the multiplication matrices in $\mathbb{R}[x]/I$ and thus to extract $V_{\mathbb{C}}(I)$.

[6] When Λ is positive, the maximality condition on the rank of $M_{\lfloor t/2 \rfloor}(\Lambda)$ implies that the rank of $M_s(\Lambda)$ is maximum for all $s \leq \lfloor t/2 \rfloor$. This is not true for Λ non-positive.

The new condition (18) is weaker than the rank condition (13) (see [19]). Namely, if (13) holds for generic $\Lambda \in \mathcal{K}_t^{\mathbb{C}}$, then (18) holds for $(t, 2s)$, i.e.,

$$\dim \pi_{2s}(\mathcal{H}_t^{\perp}) = \dim \pi_{2s-1}(\mathcal{H}_t^{\perp}) = \dim \pi_{2s}((\mathcal{H}_t^{+})^{\perp}).$$

Following [20] we now indicate how to obtain a prolongation-projection algorithm for real roots by adding positivity. Set

$$\mathcal{G}_t := \mathcal{H}_t \cup \{x^{\alpha} g \mid g \in \mathcal{N}_t, |\alpha| \leq \lfloor t/2 \rfloor\},$$

where $\mathcal{N}_t$ is as in Lemma 6. One can verify that $\mathcal{G}_t \subseteq \sqrt[\mathbb{R}]{I}$ and $\mathcal{K}_t \subseteq \mathcal{G}_t^{\perp}$. Moreover, both sets $\mathcal{K}_t$ and $\mathcal{G}_t^{\perp}$ have in fact the same dimension and thus the condition (19) below can be seen as the real analogue of (18). The next result is the real analogue of Theorem 13.

Theorem 14. *[20] Let $I = \langle h_1, \ldots, h_m \rangle$ be an ideal in $\mathbb{R}[x]$ with $|V_{\mathbb{R}}(I)| < \infty$, $D = \max_i \deg(h_i)$, and $1 \leq s \leq t$. If $\dim \pi_s(\mathcal{G}_t^{\perp}) = 0$ then $V_{\mathbb{R}}(I) = \emptyset$. Assume now $D \leq s \leq t$ and the following condition holds:*

$$\dim \pi_s(\mathcal{G}_t^{\perp}) = \dim \pi_{s-1}(\mathcal{G}_t^{\perp}) = \dim \pi_s((\mathcal{G}_t^{+})^{\perp}). \tag{19}$$

Then one can find a monomial basis $\mathcal{B} \subseteq \mathbb{R}[x]_{s-1}$ satisfying

$$\mathbb{R}[x]_{s-1} = \mathrm{Span}_{\mathbb{R}}(\mathcal{B}) \oplus (\mathbb{R}[x]_{s-1} \cap \mathrm{Span}_{\mathbb{R}}(\mathcal{G}_t))$$

and which is a basis of $\mathbb{R}[x]/J$, where the ideal J satisfies $I \subseteq J \subseteq \sqrt[\mathbb{R}]{I}$ and $V_{\mathbb{C}}(J) \cap \mathbb{R}^n = V_{\mathbb{R}}(I)$. Moreover, $J = \sqrt[\mathbb{R}]{I}$ if $\dim \pi_s(\mathcal{G}_t^{\perp}) = |V_{\mathbb{R}}(I)|$.

The following link with the rank condition is shown in [19]: (13) holds for generic $\Lambda \in \mathcal{K}_t$ if and only if

$$\dim \pi_{2s}(\mathcal{G}_t^{\perp}) = \dim \pi_{s-1}(\mathcal{G}_t^{\perp}) = \dim \pi_{2s}((\mathcal{G}_t^{+})^{\perp}).$$

Example 10. We apply the various algorithms to the ideal $I = \langle h_1, h_2, h_3 \rangle$ (taken from [5, Ex. 4, p.57]), where

$$\begin{aligned} h_1 &= x_1^2 - 2x_1x_3 + 5\,, \\ h_2 &= x_1x_2^2 + x_2x_3 + 1\,, \\ h_3 &= 3x_2^2 - 8x_1x_3\,, \end{aligned}$$

with $D = 3$, $|V_{\mathbb{C}}(I)| = 8$ and $|V_{\mathbb{R}}(I)| = 2$. Table 2 shows the ranks of the generic moment matrices when applying the real vs. complex versions of the moment matrix algorithm; we see that the algorithm terminates earlier in the real case, namely at order $t = 6$, compared to $t = 9$ in the complex case.

We indicate the behavior of the complex/real prolongation-projection algorithms in Table 3. Again in the real case the algorithm terminates earlier (at order $t = 5$) than in the complex case (at order $t = 6$). Note also that the

dimension condition holds at a smaller order than the rank condition in both real/complex cases.

If we replace each polynomial h_i by $h_i \cdot (1 + \sum_i x_1^2)$, we obtain an example with a positive dimensional complex variety, while the real variety is unchanged. Both proposed real root finding algorithms still converge (now at order $t = 7$) and allow the extraction of the two real roots.

(a) Generic $\Lambda \in \mathcal{K}_t$

$s =$	0	1	2	3
$t = 2$	1	4	—	—
$t = 3$	1	4	—	—
$t = 4$	1	4	8	—
$t = 5$	1	2	8	—
$t = 6$	1	**2**	**2**	10

(b) Generic $\Lambda \in \mathcal{H}_t^\perp$

$s =$	0	1	2	3	4
$t = 2$	1	4	—	—	—
$t = 3$	1	4	—	—	—
$t = 4$	1	4	8	—	—
$t = 5$	1	4	8	—	—
$t = 6$	1	4	8	11	—
$t = 7$	1	4	8	10	—
$t = 8$	1	4	8	9	10
$t = 9$	1	4	**8**	**8**	10

Table 2. Ranks of $M_s(\Lambda)$ in Example 10.

(a) For $\pi_s(\mathcal{H}_t^\perp)$

$s =$	0	1	2	3	4	5	6	7
$\dim \pi_s(\mathcal{H}_3^\perp)$	1	4	8	11	—	—	—	—
$\dim \pi_s(\mathcal{H}_4^\perp)$	1	4	8	10	12	—	—	—
$\dim \pi_s(\mathcal{H}_5^\perp)$	1	4	8	9	10	12	—	—
$\dim \pi_s(\mathcal{H}_6^\perp)$	1	4	**8**	**8**	9	10	12	—
$\dim \pi_s(\mathcal{H}_7^\perp)$	1	4	8	**8**	8	9	10	12

(b) For $\pi_s(\mathcal{G}_t^\perp)$ and $\pi_s((\mathcal{G}_t^+)^\perp)$

$s =$	0	1	2	3	4	5	6
$\dim \pi_s(\mathcal{G}_3^\perp)$	1	4	8	11	—	—	—
$\dim \pi_s((\mathcal{G}_3^+)^\perp)$	1	4	8	10	12	—	—
$\dim \pi_s(\mathcal{G}_4^\perp)$	1	4	8	10	12	—	—
$\dim \pi_s((\mathcal{G}_4^+)^\perp)$	1	4	8	9	10	12	—
$\dim \pi_s(\mathcal{G}_5^\perp)$	1	**2**	**2**	2	3	5	—
$\dim \pi_s((\mathcal{G}_5^+)^\perp)$	1	2	**2**	2	3	4	6

Table 3. Dimension tables in Example 10.

4 Further directions and connections

The moment approach which we presented for real solving polynomial equations can be extended and applied in various directions. As we can not discuss all connections and active streams of research, we now only try to give a flavor of some selected extensions that are most directly related to our topic. We briefly mentioned (at the end of Section 3.2) the link to the approach of [11] for approximating the convex hull of a real variety. Here we will briefly touch the following topics: polynomial optimization, emptyness certificates for real varieties, positive dimensional case, and quotient ideals.

4.1 Optimization and polynomial inequalities

The research field of polynomial optimization, which roots, in particular, in work of Lasserre [16], Parrilo [34], Shor [38], has recently undergone a spectacular development. We refer e.g. to the monograph [17] or the survey [22] for overview and further references. The moment approach was originally proposed in [16] for solving general nonlinear optimization problems of the form

$$f^* = \min_x f(x) \text{ such that } \begin{array}{l} h_1(x) = 0, \ldots, h_m(x) = 0, \\ g_1(x) \geq 0, \ldots, g_p(x) \geq 0, \end{array} \tag{20}$$

where $f, h_i, g_j \in \mathbb{R}[x]$. Let $I = \langle h_1, \ldots, h_m \rangle$ be the ideal generated by the h_i's, and set

$$S = \{x \in \mathbb{R}^n \mid g_1(x) \geq 0, \ldots, g_p(x) \geq 0\}, \tag{21}$$

so that (20) asks to minimize f over the semi-algebraic set $V_{\mathbb{R}}(I) \cap S$. The basic observation in [16] is that the problem (20) can be reformulated as

$$\min_{\mu} \Lambda_\mu(f) \text{ such that } \mu \text{ is a probability measure on } V_{\mathbb{R}}(I) \cap S,$$

where Λ_μ is as in (8). Such a linear forms satisfies: $\Lambda(h) = 0$ for all $h \in I$, as well as the positivity condition: $\Lambda(g_j p^2) \geq 0$ for all $p \in \mathbb{R}[x]$ and $j = 1, \ldots, p$. The following notation is useful to reformulate this positivity condition.

Definition 5. (Localizing moment matrix) *Given $\Lambda \in (\mathbb{R}[x]_t)^*$ and $g \in \mathbb{R}[x]_t$, $g\Lambda$ is the linear form on $\mathbb{R}[x]_{t-\deg(g)}$ defined by $(g\Lambda)(p) = \Lambda(pg)$. Its moment matrix is called a* localizing moment matrix *and*

$$\Lambda(gp^2) \geq 0 \ \forall p \in \mathbb{R}[x]_{\frac{t-\deg(g)}{2}} \iff M_{\lfloor \frac{t-\deg(g)}{2} \rfloor}(g\Lambda) \succeq 0.$$

The semidefinite program (15) can be modified in the following way to yield a relaxation of (20):

$$f_t^* = \min_{\Lambda \in (\mathbb{R}[x]_t)^*} \Lambda(f) \text{ such that } \begin{array}{l} \Lambda(1) = 1, \ \Lambda(h) = 0 \ \forall h \in \mathcal{H}_t, \\ M_{\lfloor \frac{t-\deg(g_j)}{2} \rfloor}(g_j\Lambda) \succeq 0 \ (j = 0, 1, \ldots, p) \end{array} \tag{22}$$

(setting $g_0 = 1$). The dual semidefinite program reads:

$$\max \lambda \text{ such that } f - \lambda = \sum_{j=0}^{p} \sigma_j g_j + \sum_{i=1}^{m} u_i h_i \tag{23}$$

where $u_i \in \mathbb{R}[x]$, σ_j are sums of squares of polynomials with $\deg(u_i h_i)$, $\deg(\sigma_j g_j) \leq t$. Then, $f_t^* \leq f^*$ for all t. Moreover, asymptotic convergence of (22) and (23) to the minimum f^* of (20) can be shown when the feasible region of (20) is compact (and satisfies some additional technical condition, see [16]). We now group some results showing finite convergence under certain rank condition, which can be seen as extensions of Theorems 11 and 12.

Theorem 15. *[13, 18, 22] Let* $D := \max_{i,j}(\deg(h_i), \deg(g_j))$, $d := \lceil D/2 \rceil$, $t \geq \max(\deg(f), D)$, *and let* Λ *be a generic optimal solution to (22) (i.e., for which* $\operatorname{rank} M_{\lfloor t/2 \rfloor}(\Lambda)$ *is maximum), provided it exists.*

(i) *If the rank condition (14) holds with* $2s \geq \deg(f)$, *then* $f_t^* = f^*$ *and* $V_{\mathbb{C}}(\operatorname{Ker} M_s(\Lambda))$ *is equal to the set of global minimizers of the program (20).*
(ii) *If* $V_{\mathbb{R}}(I)$ *is nonempty and finite, then (14) holds with* $2s \geq \deg(f)$.
(iii) *If* $V_{\mathbb{R}}(I) = \emptyset$, *then the program (22) is infeasible for* t *large enough.*

In other words, under the rank condition (14), one can compute all global minimizers of the program (20), since, as before, one can compute a basis of the space $\mathbb{R}[x]/\langle \operatorname{Ker} M_s(\Lambda) \rangle$ from the moment matrix and thus apply the eigenvalue method. Moreover, when the equations $h_i = 0$ have finitely many real roots, the rank condition is guaranteed to hold after finitely many steps.

By choosing the constant objective function $f = 1$ in (20), we can also compute the S-radical ideal:

$$\sqrt[S]{I} := \mathcal{I}(V_{\mathbb{R}}(I) \cap S).$$

When $|V_{\mathbb{R}}(I)|$ is nonempty and finite, one can show that

$$\mathcal{I}(V_{\mathbb{R}}(I) \cap S) = \langle \operatorname{Ker} M_s(\Lambda) \rangle$$

for a generic optimal solution Λ of (22) and s, t large enough. An analogous result holds under the weaker assumption that $|V_{\mathbb{R}}(I) \cap S|$ is nonempty and finite. In this case Λ needs to be a generic feasible solution of the modified semidefinite program obtained by adding to (22) the positivity conditions:

$$M_{\lfloor \frac{t - \deg(g)}{2} \rfloor}(g\Lambda) \succeq 0 \quad \text{for } g = g_1^{e_1} \cdots g_p^{e_p} \ \forall e \in \{0,1\}^p.$$

The key ingredient in the proof is to use the Positivstellensatz (to characterize the polynomials in $\mathcal{I}(V_{\mathbb{R}}(I) \cap S)$, see [41]) instead of the Real Nullstellensatz (used in Theorem 10 to characterize the polynomials in $\mathcal{I}(V_{\mathbb{R}}(I))$).

Let us illustrate on an example how to 'zoom in' on selected roots, by incorporating semi-algebraic constraints or suitably selecting the cost function.

Example 11. Consider the system (from [14]), known as Katsura 5:

$$\begin{aligned}
h_1 &= 2x_6^2 + 2x_5^2 + 2x_4^2 + 2x_3^2 + 2x_2^2 + x_1^2 - x_1\,,\\
h_2 &= x_6x_5 + x_5x_4 + 2x_4x_3 + 2x_3x_2 + 2x_2x_1 - x_2\,,\\
h_3 &= 2x_6x_4 + 2x_5x_3 + 2x_4x_2 + x_2^2 + 2x_3x_1 - x_3\,,\\
h_4 &= 2x_6x_3 + 2x_5x_2 + 2x_3x_2 + 2x_4x_1 - x_4\,,\\
h_5 &= x_3^2 + 2x_6x_1 + 2x_5x_1 + 2x_4x_1 - x_5\,,\\
h_6 &= 2x_6 + 2x_5 + 2x_4 + 2x_3 + 2x_2 + x_1 - 1,
\end{aligned}$$

with $D = 2$, $|V_{\mathbb{C}}(I)| = 32$, and $|V_{\mathbb{R}}(I)| = 12$. Table 4(a) shows the ranks of the generic moment matrices for the moment matrix algorithm to compute $V_{\mathbb{R}}(I)$. At order $(t, s) = (6, 3)$, the algorithm finds all twelve real roots.

Next we apply the moment matrix algorithm to compute the real roots in $S = \{x \in \mathbb{R}^6 \mid g(x) = x_1 - 0.5 \geq 0\}$; the ranks are shown in Table 4(b) and all five elements of $V_{\mathbb{R}}(I) \cap S$ can be computed at order $(t, s) = (4, 2)$.

If we are interested e.g. only in the roots in $V_{\mathbb{R}}(I) \cap S$ with the smallest x_2-coordinate then we minimize the polynomial x_2 (instead of the constant one polynomial). The moment matrix algorithm now terminates at order $(t, s) = (2, 1)$ and finds the unique element of $V_{\mathbb{R}}(I) \cap S$ with the smallest x_2-coordinate.

(a) Generic $\Lambda \in \mathcal{K}_t$

$s =$	0	1	2	3
$t = 2$	1	6	—	—
$t = 3$	1	6	—	—
$t = 4$	1	6	16	—
$t = 5$	1	6	16	—
$t = 6$	1	6	**12**	**12**

(b) Generic $\Lambda \in \mathcal{K}_t$ with $M_{\lfloor \frac{t-1}{2} \rfloor}(g\Lambda) \succeq 0$

$s =$	0	1	2
$t = 2$	1	6	—
$t = 3$	1	6	—
$t = 4$	1	**5**	**5**

Table 4. Ranks of $M_s(\Lambda)$ in Example 11.

4.2 Exact certificates of emptiness

If the moment method is applied to an empty real variety $V_{\mathbb{R}}(I)$ (or subset $V_{\mathbb{R}}(I) \cap S$), then the underlying semidefinite optimization problem is infeasible for t large enough, which thus can be thought of as a *numerical certificate* of emptiness, see Theorems 12 and 15. If we solve the semidefinite program (15) with a primal-dual interior point solver and infeasibility is detected, an improving ray is returned, i.e. a solution to the dual problem (16) of the form:

$$1 - \lambda^* = \sigma + \sum_{i=1}^{m} u_i h_i \quad \text{where } \sigma, u_i \in \mathbb{R}[x] \text{ and } \sigma \text{ is a sum of squares,} \tag{24}$$

with $\lambda^* > 1$. By scaling both sides with an arbitrary positive number, one can generate a feasible solution of the dual problem (16) with an arbitrary high cost function value, thus certifying infeasibility of the primal problem.

On the other hand, by the Real Nullstellensatz, we know that an algebraic certificate for emptyness of $V_{\mathbb{R}}(I)$ is that $1 \in \sqrt[\mathbb{R}]{I}$, i.e.,

$$1 + \sigma = \sum_{i=1}^{m} u_i h_i \text{ for some } \sigma, u_i \in \mathbb{R}[x] \text{ where } \sigma \text{ is a sum of squares.} \tag{25}$$

In principle, such a certificate can be directly derived from an improving ray such as (24). The difficulty, however, arise from numerical imprecisions

and the certificate computed using semidefinite programming does not hold exactly when all computations are done in floating point arithmetics.

We may thus only derive polynomials u_i, σ satisfying

$$1 + \sigma + \epsilon = \sum_{i=1}^{m} u_i h_i, \tag{26}$$

where $\epsilon \in \mathbb{R}[x]_t$ represents the cumulated error term. This approximate certificate can still be used to produce an *exact* certificate for the nonexistence of roots in some ball around the origin.

Proposition 1. *[36, Prop. 7.38] Suppose (26) holds, where $\epsilon \in \mathbb{R}[x]_t$ satisfies $|\epsilon(0)| \ll 1$. Then there exists some $\delta > 0$ such that $V_{\mathbb{R}}(I) \cap B_\delta = \emptyset$, where B_δ is the Euclidean ball of radius δ.*

Example 12. Consider the ideal $I = (h_1, h_2, h_3)$ generated by

$$\begin{aligned} h_1 &= x_1^4 + x_2^4 + x_3^4 - 4\,, \\ h_2 &= x_1^5 + x_2^5 + x_3^5 - 5\,, \\ h_3 &= x_1^6 + x_2^6 + x_3^6 - 6 \end{aligned}$$

with $D = 6$, $|V_{\mathbb{C}}(I)| = 120$, and $V_{\mathbb{R}}(I) = \emptyset$.

Already at the first relaxation order $t = 6$, the primal (moment) problem becomes infeasible, the solver returns an improving direction for the dual (SOS) problem and we obtain a numerical certificate of the form (26). The error polynomial $\epsilon \in \mathbb{R}[x]$ is a dense polynomial with $\deg(\epsilon) = 6$, with $k = 84$ coefficients, each of which is smaller than $\epsilon_{\max} = 4.1\text{e-}11$, and a constant term $\epsilon_0 < 8.53\text{e-}14$. Using the conservative estimate of [36, §7.8.2] one can rigorously certify the emptiness of the set $V_{\mathbb{R}}(I) \cap B_\delta$ with $\delta = 38.8$. In other words, even if we only solved the problem numerically with a rather low accuracy, we still obtain a proof that the ideal I does not have any real root $v \in V_{\mathbb{R}}(I)$ with $\|v\|_2 < 38.8$. By increasing the accuracy of the SDP solver the radius of the ball B_δ can be further increased. This example illustrates that it is sometimes possible to draw exact conclusions from numerical computations.

4.3 Positive dimensional varieties

The algorithms presented so far for computing the real variety $V_{\mathbb{R}}(I)$ and the real radical ideal $\sqrt[\mathbb{R}]{I}$ work under the assumption that $V_{\mathbb{R}}(I)$ is finite. Indeed, the rank condition (13) or (14) will never hold in the case of a positive dimensional variety $V_{\mathbb{R}}(I)$ (since (13) implies $\dim \mathbb{R}[x]/\sqrt[\mathbb{R}]{I} = \operatorname{rank} M_{s-1}(\Lambda)$, by Theorem 11). Nevertheless, the moment method can in principle be applied to find a basis of $\sqrt[\mathbb{R}]{I}$ also in the positive dimensional case. Indeed Theorem 10 shows that, for t large enough, $\langle \operatorname{Ker} M_{\lfloor t/2 \rfloor}(\Lambda) \rangle = \sqrt[\mathbb{R}]{I}$ for generic $\Lambda \in \mathcal{K}_t$. However, it is not clear yet how to recognize equality $\sqrt[\mathbb{R}]{I} = \langle \operatorname{Ker} M_{\lfloor t/2 \rfloor}(\Lambda) \rangle$

in the positive dimensional case. As we now briefly mention, these questions relate e.g. to the study of the Hilbert function of the real radical ideal $\sqrt[\mathbb{R}]{I}$.

Recall that, for an ideal $J \subseteq \mathbb{R}[x]$, its *Hilbert function* HF_J is defined by

$$\mathrm{HF}_J(t) = \dim \mathbb{R}[x]_t/(J \cap \mathbb{R}[x]_t) \quad \text{for } t \in \mathbb{N}.$$

For any large t, $\mathrm{HF}_J(t)$ is equal to a polynomial, the *Hilbert polynomial* of J (see e.g. [5] for details). If $\mathcal{B}$ is the set of standard monomials for the ideal J with respect to a total degree monomial ordering, then $|\mathcal{B} \cap \mathbb{R}[x]_t| = \mathrm{HF}_J(t)$.

Proposition 2. *[36] There exists an integer t_0 for which the following properties hold for all $t \geq t_0$: For any generic $\Lambda \in \mathcal{K}_t$, $\langle \mathrm{Ker}\, M_{\lfloor t/2 \rfloor}(\Lambda)\rangle = \sqrt[\mathbb{R}]{I}$ and* $\mathrm{rank}\, M_s(\Lambda) = \mathrm{HF}_{\sqrt[\mathbb{R}]{I}}(s)$ *for all $1 \leq s < \lfloor t/2 \rfloor$.*

It is still ongoing research how to use this information to systematically get hold of points in the positive dimensional real variety $V_{\mathbb{R}}(I)$, or how to formulate a solid termination criterion, permitting e.g. to recognize that $\sqrt[\mathbb{R}]{I}$ has been found. In the following example we illustrate a possible first step in this direction by using the moment matrix algorithm to compute *parametric* multiplication matrices.

Example 13. Consider the ideal $I = \langle h_1, h_2, h_3\rangle$, where

$$\begin{aligned} h_1 &= x_1^2 + x_1x_2 - x_1x_3 - x_1 - x_2 + x_3,\\ h_2 &= x_1x_2 + 2x_2^2 - x_2x_3 - x_1 - cx_2 + x_3,\\ h_3 &= x_1x_3 + x_2x_3 - x_3^2 - x_1 - x_2 + x_3 \end{aligned}$$

(taken from [42, p.397, Eq.(9.60)], where a solution strategy based on Gröbner bases is proposed). The variety $V_{\mathbb{C}}(I)$ consists of three components: two one-dimensional zero-manifolds, the straight line $(t, 0, t)$ for all $t \in \mathbb{C}$, another straight line $(t, 1, t{+}1)$ for $t \in \mathbb{C}$, and an isolated zero at $(1, 1, 1)$. Furthermore, the first positive dimensional manifold has an embedded point of multiplicity two at $(1, 0, 1)$. Table 5 shows the ranks of the generic moment matrices for the (real) moment matrix algorithm.

s =	0	1	2
$t = 2$	1	4	—
$t = 3$	1	4	—
$t = 4$	1	**4**	**6**

Table 5. Ranks of $M_s(\Lambda)$ for generic $\Lambda \in \mathcal{K}_t$ in Example 13.

Order $t = 4$ is the first time when a generic moment matrix has a nontrivial kernel and we can compute the following polynomials in $\mathrm{Ker}\, M_2(\Lambda)$:

$$
\begin{aligned}
f_1 &= x_1x_3 + x_1 + x_2 - x_3 - x_1^2 - x_1x_2\,,\\
f_2 &= x_2^2 - x_2\,,\\
f_3 &= x_2x_3 + x_1 - x_3 - x_1x_2\,,\\
f_4 &= x_3^2 + 3x_1 + 2x_2 - 3x_3 - x_1^2 - 2x_1x_2\,.
\end{aligned}
$$

From these polynomials we can derive the *parametric multiplication matrices* $\mathcal{X}_i(x_1)$ (indexed by $\mathcal{B} = \{1, x_2, x_3\}$) in the ring $\mathbb{K}[x_2, x_3]$ with $\mathbb{K} = \mathbb{R}(x_1)$, i.e., we consider x_1 as a parameter. These matrices read:

$$
\mathcal{X}_1(x_1) = \begin{pmatrix} x_1 & 0 & x_1^2 - x_1 \\ 0 & x_1 & x_1 - 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathcal{X}_2(x_1) = \begin{pmatrix} 0 & 0 & -x_1 \\ 1 & 1 & x_1 \\ 0 & 0 & 1 \end{pmatrix},
$$

and

$$
\mathcal{X}_3(x_1) = \begin{pmatrix} 0 & -x_1 & x_1^2 - 3x_1 \\ 0 & x_1 & 2x_1 - 2 \\ 1 & 1 & 3 \end{pmatrix}.
$$

As the matrices $\mathcal{X}_2(x_1)$ and $\mathcal{X}_3(x_1)$ commute, we deduce that $\{f_2, f_3, f_4\}$ forms a border basis of the ideal

$$
J = \langle \operatorname{Ker} M_2(\Lambda) \rangle \subseteq \mathbb{K}[x] = \mathbb{R}(x_1)[x_2, x_3].
$$

In this simple example we are able to compute the eigenvectors of $\mathcal{X}_3(x_1)^T$ parametrically and we can employ Theorem 3 over $\mathbb{K}[x]/J$ to compute all components containing real solutions, i.e. characterizing $V_{\mathbb{R}}(I)$.

These eigenvectors are $v_1^{(3)}(x_1) = (1, 1, x_1 + 1)$ with eigenvalue $\lambda_1^{(3)}(x_1) = x_1 + 1$, $v_2^{(3)}(x_1) = (1, 0, x_1)$ with eigenvalue $\lambda_2^{(3)}(x_1) = x_1$ and $v_3^{(3)}(x_1) = (1, 1, -x_1 + 2)$ with eigenvalue $\lambda_3^{(3)}(x_1) = 2 - x_1$ for all values of $x_1 \in \mathbb{C}$. All three eigenvectors $v_1^{(3)}(x_1), \ldots, v_3^{(3)}(x_1)$ are also eigenvectors for $\mathcal{X}_2(x_1)^T$ with corresponding eigenvalues $\lambda_1^{(2)}(x_1) = 1$, $\lambda_2^{(2)}(x_1) = 0$ and $\lambda_3^{(2)}(x_1) = 1$. The first two eigenvectors $v_1^{(3)}(x_1), v_2^{(3)}(x_1)$ represent respectively the two positive dimensional zero-manifolds $(t, 1, t + 1)$ and $(t, 0, t)$ (just replace x_1 by t in the eigenvectors/eigenvalues) and are also (parametric) eigenvectors for $\mathcal{X}_1(x_1)^T$ with $\lambda_1^{(1)}(x_1) = x_1$ and $\lambda_2^{(1)}(x_1) = x_1$.

The third eigenvector $v_3^{(3)}(1)$ is not an eigenvector for $\mathcal{X}_1(x_1)^T$ for all values of $x_1 \in \mathbb{C}$. For $x_1 = 1$, it is indeed an eigenvector of $\mathcal{X}_1(1)^T$ with $\lambda_3^{(1)}(1) = 1$ and, from $\lambda_3^{(2)}(1) = \lambda_3^{(3)}(1) = 1$, we see that this eigenvector characterizes the isolated solution $(1, 1, 1) \in V_{\mathbb{C}}(I)$. Compared to Stetter [42, p.397, Eq.(9.60)], we obtain 3×3 multiplication matrices (as opposed to 4×4 matrices in [42]), since the moment matrix approach computed the (real) radical ideal $J = \sqrt[\mathbb{R}]{I}$ and thus removed the multiplicity at the embedded double point $(1, 0, 1)$.

In general eigenvectors cannot be computed parametrically, but we may, e.g., still try to explore the variety $V_{\mathbb{C}}(I)$ by sampling values for the parameter vector (e.g. x_1 in the above example). The general question of how to best characterize positive dimensional manifolds is largely open. See e.g. [25] for a discussion in the complex case. Even the task of computing finitely many points on a positive dimensional real variety is not completely trivial, since random linear sections are likely to miss the real locus of the variety altogether. One possibility is to optimize a generic linear objective function over the real variety using the ideas discussed above, but this approach usually reveals only a single point per semidefinite optimization problem.

4.4 Ideal quotients

The *ideal quotient* (also known as *colon ideal*) $I : J$ of two ideals I and J is defined as

$$I : J = \{p \in \mathbb{R}[x] \,|\, pJ \subseteq I\}\,.$$

If the ideal J is given by a system of generators, say $J = \langle g_1, \ldots, g_m\rangle$, then the computation of the ideal quotient of I by J can be done element by element, using the identity:

$$I : J = \bigcap_{i=1}^{m} I : \langle g_i\rangle.$$

Hence it suffices to study the ideal quotient $I : \langle g\rangle$, also denoted as $I : g$, of I by a principal ideal $\langle g\rangle$. If $\{f_1, \ldots, f_p\}$ is a basis of the ideal $I \cap \langle g\rangle$, then $\{f_1/g, \ldots, f_p/g\}$ is a basis of $I : g$. (See e.g. [5, §4] for details.)

Say I is an ideal given by its generators $h_1, \ldots, h_m$. We now see that the moment approach can be easily adapted to find a semidefinite characterization of the ideal

$$\sqrt[\mathbb{R}]{I} : g = \mathcal{I}(V_{\mathbb{R}}(I) \setminus V_{\mathbb{R}}(g)).$$

For this we use again the notion of localizing matrix (recall Definition 5).

Proposition 3. *Let $g \in \mathbb{R}[x]_k$, set $\rho = 1 + \lceil k/2\rceil$ and $D = \max_i \deg(h_i)$. Let Λ be a generic element in $\mathcal{K}_{t+k}$.*

(i) $\langle \operatorname{Ker} M_{\lfloor t/2\rfloor}(g\Lambda)\rangle \subseteq \sqrt[\mathbb{R}]{I} : g$, *with equality for t large enough.*
(ii) *If the rank condition:* $\operatorname{rank} M_s(\Lambda) = \operatorname{rank} M_{s-\rho}(\Lambda)$ *holds for some s with* $\max(D, \rho) \le s \le \lfloor t/2\rfloor$, *then* $\sqrt[\mathbb{R}]{I} : g = \langle \operatorname{Ker} M_{s-\rho+1}(g\Lambda)\rangle$.
(iii) *If $V_{\mathbb{R}}(I)$ is nonempty finite, then the rank condition in (ii) holds at some order (t, s).*

Proof. We use the fact that $p \in \operatorname{Ker} M_{\lfloor t/2\rfloor}(g\Lambda) \iff pg \in \operatorname{Ker} M_{\lfloor t/2\rfloor}(\Lambda)$, for $p \in \mathbb{R}[x]_{\lfloor t/2\rfloor - k}$. As $\Lambda \in \mathcal{K}_{t+k}$ is generic, $\operatorname{Ker} M_s(\Lambda) \subseteq \sqrt[\mathbb{R}]{I}$, which implies $\langle \operatorname{Ker} M_s(g\Lambda)\rangle \subseteq \sqrt[\mathbb{R}]{I} : g$, for any $s \le \lfloor t/2\rfloor$.
(i) To show equality: $\langle \operatorname{Ker} M_{\lfloor t/2\rfloor}(g\Lambda)\rangle = \sqrt[\mathbb{R}]{I} : g$ for t large enough, we proceed

as in the proof of Theorem 10: Pick a basis $\{g_1, \ldots, g_L\}$ of the ideal $\sqrt[\mathbb{R}]{I} : g$, so that each $g_l g$ belongs to $\sqrt[\mathbb{R}]{I}$; apply the Real Nullstellensatz to $g_l g$ to conclude that, for t large, $g_l g \in \operatorname{Ker} M_{\lfloor t/2 \rfloor}(\Lambda)$ and thus $g_l \in \operatorname{Ker} M_{\lfloor t/2 \rfloor}(g\Lambda)$.
(ii) Assume now $\operatorname{rank} M_s(\Lambda) = \operatorname{rank} M_{s-\rho}(\Lambda)$ for $D, \rho \leq s \leq \lfloor t/2 \rfloor$. Then there exists $\tilde{\Lambda} \in \mathbb{R}[x]^*$ for which $M(\tilde{\Lambda})$ is a flat extension of $M_s(\tilde{\Lambda})$ and $\sqrt[\mathbb{R}]{I} = \operatorname{Ker} M(\tilde{\Lambda}) = \langle \operatorname{Ker} M_{s-\rho+1}(\Lambda) \rangle$ (use Theorems 9 and 11). Therefore,

$$\sqrt[\mathbb{R}]{I} : g = \operatorname{Ker} M(\tilde{\Lambda}) : g = \operatorname{Ker} M(g\tilde{\Lambda}).$$

One can verify that $M(g\tilde{\Lambda})$ is a flat extension of $M_{s-\rho}(g\tilde{\Lambda})$, which implies that $\operatorname{Ker} M(g\tilde{\Lambda}) = \langle \operatorname{Ker} M_{s-\rho+1}(g\tilde{\Lambda}) \rangle$ (using Theorem 9) is thus equal to $\langle \operatorname{Ker} M_{s-\rho+1}(g\Lambda) \rangle$ (since $\tilde{\Lambda}$ and Λ coincide on $\mathbb{R}[x]_{2s}$).
(iii) follows from an easy modification of the proof of Theorem 12. □

It would be interesting to see whether the above stopping criterion can be modified in order to deal with zero-dimensional (or even empty) colon ideals resulting from a positive dimensional ideal I, and also whether the moment approach can deal with the ideal $\sqrt[\mathbb{R}]{I : g}$.

Example 14. Consider the zero-dimensional ideal $I = \langle x_2 - x_1^2, x_2^2 - x_2 \rangle$ with a double root at $v = (0, 0)$ and two simple roots at $v = (\pm 1, 1)$. The corresponding real radical ideal is $\sqrt[\mathbb{R}]{I} = \langle x_2 - x_1^2, x_2^2 - x_2, x_1 x_2 - x_1 \rangle$. The colon ideal computation with $g = x_1$ converges at order $(t, s) = (6, 3)$ and we obtain $\sqrt[\mathbb{R}]{I} : g = (x_2 - x_1^2, x_2 - 1)$ with variety $V_{\mathbb{R}}(I) \setminus V_{\mathbb{R}}(g) = \{(-1, 1), (1, 1)\}$. The corresponding ranks of $M_s(\Lambda)$ and $M_s(g\Lambda)$ for generic $\Lambda \in \mathcal{K}_t$ are shown in Table 6.

(a) $\operatorname{rank} M_s(\Lambda)$

$s =$	0	1	2	3
$t = 4$	1	3	3	—
$t = 5$	1	3	3	—
$t = 6$	1	**3**	3	**3**

(b) $\operatorname{rank} M_s(g\Lambda)$

$s =$	0	1	2	3
$t = 4$	1	2	—	—
$t = 5$	1	2	—	—
$t = 6$	1	**2**	**2**	—

Table 6. Rank sequences for generic $\Lambda \in \mathcal{K}_t$ in Example 14.

References

1. N.I. Akhiezer, *The Classical Moment Problem*, Hafner, 1965.
2. S. Basu, R. Pollack, and M.F. Roy, *Algorithms in Real Algebraic Geometry*, Springer, 2003.
3. D. Bates and F. Sottile, *Khovanskii-Rolle continuation for real solutions*, arXiv:0908.4579, 2009.
4. J. Bochnak, M. Coste, and M.-F. Roy, *Real Algebraic Geometry*, Springer, 1998.

5. D.A. Cox, J.B. Little, and D.B. O'Shea, *Using Algebraic Geometry*, Springer, 1998.
6. ———, *Ideals, Varieties, and Algorithms : An Introduction to Computational Algebraic Geometry and Commutative Algebra*, Springer, 2005.
7. R.E. Curto and L. Fialkow, *Solution of the truncated complex moment problem for flat data*, Memoirs of the American Mathematical Society **119** (1996), no. 568, 1–62.
8. E. de Klerk, *Aspects of Semidefinite Programming - Interior Point Algorithms and Selected Applications*, Kluwer, 2002.
9. A. Dickenstein and I. Z. Emiris (eds.), *Solving Polynomial Equations: Foundations, Algorithms, and Applications*, Algorithms and Computation in Mathematics, vol. 14, Springer, 2005.
10. M. Elkadi and B. Mourrain, *Introduction à la Résolution des Systèmes d'Equations Polynomiaux*, Mathématiques et Applications, vol. 59, Springer, 2007.
11. J. Gouveia, P.A. Parrilo and R.R. Thomas, *Theta Bodies for Polynomial Ideals*, SIAM Journal of Optimization, **20**(4) (2010), pp. 2097–2118.
12. E.K. Haviland, *On the momentum problem for distribution functions in more than one dimension. II*, American Journal of Mathematics **58** (1936), no. 1, 164–168.
13. D. Henrion and J.B. Lasserre, *Detecting global optimality and extracting solutions in GloptiPoly*, In *Positive Polynomials in Control*, Lecture Notes in Control and Information Sciences, pp. 293–310, Springer, 2005.
14. `http://www.mat.univie.ac.at/~neum/glopt/coconut/Benchmark/Library3/katsura5.mod`
15. A. Kehrein and M. Kreuzer, *Characterizations of border bases*, Journal of Pure and Applied Algebra **196** (2005), 251–270.
16. J.B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM Journal on Optimization **11** (2001), 796–817.
17. ———, *Moments, Positive Polynomials, and their Applications*, Imperial College Press, 2009.
18. J.B. Lasserre, M. Laurent, and P. Rostalski, *Semidefinite characterization and computation of real radical ideals*, Foundations of Computational Mathematics **8** (2008), no. 5, 607–647.
19. ———, *A unified approach for real and complex zeros of zero-dimensional ideals.*, In *Emerging Applications of Algebraic Geometry* (M. Putinar and S. Sullivant, eds.), vol. 149, Springer, 2009, pp. 125–156.
20. ———, *A prolongation-projection algorithm for computing the finite real variety of an ideal*, Theoretical Computer Science **410** (2009), no. 27-29, 2685–2700.
21. M. Laurent, *Revisiting two theorems of Curto and Fialkow*, Proceedings of the American Mathematical Society **133** (2005), no. 10, 2965–2976.
22. ———, *Sums of squares, moment matrices and optimization over polynomials*, In *Emerging Applications of Algebraic Geometry* (M. Putinar and S. Sullivant, eds.), Springer, 2009, pp. 157–270.
23. M. Laurent and B. Mourrain, *A generalized flat extension theorem for moment matrices*, Archiv der Mathematik **93** (2009), no. 1, 87–98.
24. D. Lazard, *Resolution des systemes d'equations algebriques*, Theoretical Computer Science **15** (1981), 77–110.
25. ———, *Thirty years of polynomial system solving, and now?* Journal of Symbolic Computation **44** (2009), 222-231.

26. H.M. Möller, *An inverse problem for cubature formulae*, Computational Technologies **9** (2004), 13–20.
27. T. Mora, *Solving Polynomial Equation Systems I: The Kronecker-Duval Philosophy*, Encyclopedia of Mathematics and its Applications, no. 88, Cambridge University Press, 2003.
28. ———, *Solving Polynomial Equation Systems II: Macaulay's Paradigm and Gröbner Technology*, Encyclopedia of Mathematics and its Applications (v. 2), no. 88, Cambridge University Press, 2005.
29. B. Mourrain, *A new criterion for normal form algorithms*, AAECC, Lecture Notes in Computer Science, pp. 430–443, 1999.
30. ———, *Pythagore's dilemma, symbolic - numeric computation and the border basis method*, In *Symbolic-numeric computation*, Trends in Mathematics, pp. 223–243, Birkhäuser, 2007.
31. B. Mourrain and J. P. Pavone, *Subdivision methods for solving polynomial equations*, Journal of Symbolic Computation **44** (2009), no. 3, 292–306.
32. B. Mourrain and P. Trebuchet, *Generalized normal forms and polynomials system solving*, In Proc. Intern. Symp. on Symbolic and Algebraic Computation (M. Kauers, ed.), pp. 253-260, ACM Press, 2005.
33. B. Mourrain, J.B. Lasserre, M. Laurent, P. Rostalski and P. Trebuchet, *Moment matrices and border bases*, In preparation.
34. P.A. Parrilo, *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*, Ph.D. thesis, California Institute of Technology, 2000.
35. G. Reid and L. Zhi, *Solving polynomial systems via symbolic-numeric reduction to geometric involutive form*, Journal of Symbolic Computation **44** (2009), no. 3, 280 – 291.
36. P. Rostalski, *Algebraic Moments – Real Root Finding and Related Topics*, Ph.D. thesis, ETH Zürich, 2009.
37. F. Rouillier, *Solving zero-dimensional systems through the rational univariate representation*, Journal of Applicable Algebra in Engineering, Communication and Computing **9** (1999), no. 5, 433–461.
38. N.Z. Shor, *An approach to obtaining global extremums in polynomial mathematical programming problems*, Kibernetica **5** (1987), 102–106.
39. A.J. Sommese and C.W. Wampler, *The Numerical Solution of Systems of Polynomials Arising in Engineering and Science*, World Scientific Press, 2005.
40. F. Sottile, *Numerical real algebraic geometry*, In preparation, 2009.
41. G. Stengle, *A Nullstellensatz and a Positivstellensatz in semialgebraic geometry*, Mathematische Annalen **207** (1974), 87–97.
42. H.J. Stetter, *Numerical Polynomial Algebra*, SIAM, 2004.
43. B. Sturmfels, *Solving Systems of Polynomial Equations*, American Mathematical Society, 2002.
44. H. Wolkowicz, R. Saigal, and L. Vandenberghe (eds.), *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*, vol. 27, Kluwer Academic Publishers, Dordrecht, The Netherlands / Boston, MA, 2000.