Computing sparse polynomials via witness sets

Jonathan D. Hauenstein ∗ Laura Matusevich † Chris Peterson ‡ Samantha N. Sherman §

February 8, 2020

Abstract

Sparse polynomials that vanish on algebraic sets are preferred in many computations since they are easy to evaluate and often arise from underlying structure. For example, a monomial vanishes on an algebraic set if and only if the algebraic set is contained in the union of the coordinate hyperplanes. This paper considers computing sparse polynomials that vanish on an algebraic set which is represented by a witness set. Our approach relies on using numerical homotopy methods to sample points on the algebraic set along with incorporating multiplicity information using Macaulay dual spaces. If the algebraic set is defined by polynomials with rational coefficients, exactness recovery such as lattice based methods can be used to find exact representations of the sparse polynomials and Gröbner basis techniques can be used to certify correctness. Several examples are presented demonstrating the approach.

1 Introduction

The number of terms in a polynomial is one measure of its complexity. A polynomial with relatively few terms is said to be sparse. Sparse polynomials can be used to identify structure. In terms of sparsity, monomials (polynomials with one term) and binomials (polynomials with at most two terms) are special. Monomials show that the corresponding solution set is contained in the union of the coordinate hyperplanes. Also, sparsity-preserving operations can be performed on ideals generated by binomials [12]. Our aim is to describe a general approach to compute sparse polynomials that is not limited to just binomials.

---

*Department of Applied and Computational Mathematics and Statistics, University of Notre Dame, Notre Dame, IN 46556 (hauenstein@nd.edu, www.nd.edu/~jhauenst).
†Department of Mathematics, Texas A&M University, College Station, TX 77843-3368 (laura@math.tamu.edu, www.math.tamu.edu/~laura).
‡Department of Mathematics, Colorado State University, Fort Collins CO 80523-1874 (peterson@math.colostate.edu, www.math.colostate.edu/~peterson).
§Department of Applied and Computational Mathematics and Statistics, University of Notre Dame, Notre Dame, IN 46556 (ssherma1@nd.edu, sites.nd.edu/samantha-sherman).
Sparse polynomials are also useful for bounding the number of solutions. For example, over the real numbers, Descartes’ rule of sign and multivariate generalizations, e.g., [1, 3, 18, 22], utilize sparsity to bound the number of real solutions. Over the complex numbers, the Bernstein-Kouchnirenko-Khovanskii (BKK) theorem, e.g., see [22, Chaps. 3-4], provides an upper bound in terms of the Newton polyhedra of the polynomials. Since one also has a trivial bound from Bézout’s theorem based on degrees, one can play off the two complexity measures (sparseness and degree) against each other to design solving approaches in numerical algebraic geometry.

Since monomials and binomials are extremely useful to identify in many algebraic geometric computations, previous methods have been proposed to search for them. Monomials can be determined using Gröbner basis methods, e.g., see [20]. A Gröbner-free approach for deciding if a given ideal is generated by binomials using linear algebra computations is presented in [7]. For some polynomial systems arising from biological models, this approach can be more efficient than Gröbner basis methods [15]. An approach for detecting binomiality after an ambient automorphism was proposed in [17]. A theoretical approach based on tropical geometry, matrix theory, and computational number theory is presented in [16] for deciding if an ideal contains a binomial.

To extend beyond binomials, the approach of this article is to introduce a bound on the degree. That is, for a degree bound $d$ and sparsity bound $t$, this article describes a numerical algebraic geometric approach for computing all polynomials with at most $t$ terms of degree at most $d$ that vanish on an algebraic set represented by a witness set. The witness set allows sample points and multiplicity information to be computed. From this data, we propose two methods to compute sparse polynomials: checking all minors of a matrix constructed from the degree $d$ Veronese embedding of the points or using an $\ell_1$-relaxation. If the algebraic set is defined over the rational numbers, then exactness recovery techniques, such as lattice based methods [2] can be applied to obtain exact representations of the sparse polynomials. In this case, one can certify the correctness of the computed polynomials using standard Gröbner basis methods.

The rest of the article is structured as follows. Section 2 summarizes necessary background information on witness sets, sampling, multiplicity, exactness recovery, and $\ell_1$-relaxation. The method for computing sparse polynomials is described in Section 3 and demonstrated on several examples in Section 4.

2 Background

The following provides necessary background on topics from numerical algebraic geometry and sparsity. For more details regarding numerical algebraic geometry, see [5, 21].
2.1 Witness sets and sampling

For a system of polynomials

\[ f(x) = \begin{bmatrix} f_1(x_1, \ldots, x_N) \\ \vdots \\ f_n(x_1, \ldots, x_N) \end{bmatrix}, \]

the corresponding algebraic set (or variety) is

\[ V(f) = \{ x \in \mathbb{C}^N \mid f(x) = 0 \} \subset \mathbb{C}^N. \]

Each algebraic set can be decomposed uniquely (up to reordering) into a union of irreducible components creating the irreducible decomposition, say

\[ V(f) = \bigcup_{i=1}^{k} A_i \]

where each \( A_i \) is irreducible. Let \( \dim A_i \) be the dimension of \( A_i \), which is equal to the minimum of the dimension of the tangent space at each point of \( A_i \).

A witness set for an irreducible algebraic set \( X \subset \mathbb{C}^N \) is \( \{ g, L, W \} \) which provides a geometric representation of \( X \) where:

- **witness system** \( g \) is a polynomial system such that \( X \) is an irreducible component of \( V(g) \),
- **witness slice** \( L \) is a linear space with \( \text{codim } L = \dim X \) which intersects \( X \) transversely, and
- **witness point set** \( W = X \cap L \) with \( \#W = \deg X \).

One key aspect of a witness set is that one can deform the linear space \( L \) to compute other points on \( X \), i.e., sample points from \( X \). For example, if \( L' \subset \mathbb{C}^N \) is a general linear space with \( \text{codim } L' = \text{codim } L = \dim X \), then one can consider the homotopy deforming \( L \) to \( L' \) along \( X \), namely

\[ X \cap (t \cdot L + (1-t) \cdot L'). \] (1)

At \( t = 1 \), one starts with \( W = X \cap L \) which deforms to \( W' = X \cap L' \) at \( t = 0 \). By selecting various linear spaces, one is able to sample as many generic points on \( X \) as needed.

**Example 1.** To illustrate, consider \( X = V(x^2 + y^2 - 1) \subset \mathbb{C}^2 \). Clearly, \( \dim X = 1 \) and \( \deg X = 2 \). An example of a witness set for \( X \), as illustrated in Figure 1, is \( \{ g, L, W \} \) where:

- \( g = x^2 + y^2 - 1 \),
- \( L = V(x + 2y - 1) \) with \( \text{codim } L = \dim X = 1 \),
- \( W = X \cap L = \{(1, 0), (-3/5, 4/5)\} \) with \( \#W = \deg X = 2 \).

For the witness set in Ex. 1, the coefficients of the linear space \( L \) were chosen to be small integers for illustration purposes. In practice, the coefficients are chosen to be random complex numbers to ensure, with probability one, that \( L \) is transverse to \( X \).
2.2 Multiplicity

If \( g \) is a witness system for \( X \), then \( X \) could have multiplicity greater than 1 as a component of \( \mathcal{V}(g) \). There are two aspects associated to this that one needs to consider. First, one wants multiplicity one in order to numerically perform the path tracking for (1). However, one wants to capture the multiplicity structure imposed by a given polynomial system in order to compute sparse polynomials satisfying that multiplicity structure. These aspects are addressed using deflation and Macaulay dual spaces, respectively.

Suppose that \( X \subset \mathbb{C}^N \) is an irreducible algebraic set with witness set \( \{g, \mathcal{L}, W\} \). Considering (1), let \( \ell \) and \( \ell' \) be linear systems such that \( \mathcal{L} = \mathcal{V}(\ell) \) and \( \mathcal{L}' = \mathcal{V}(\ell') \). Then, one can translate the geometric deformation (1) to the algebraic homotopy

\[
H(x,t) = \begin{bmatrix} g(x) \\ t \cdot \ell(x) + (1-t) \cdot \ell'(x) \end{bmatrix} = 0.
\]

The multiplicity of \( X \) with respect to \( g \) is the multiplicity of \( w \in W \) with respect to \( \{g, \ell\} \). Hence, if \( X \) has multiplicity one with respect to \( g \), then \( \dim \ker Jg(w) = \dim X \) where \( Jg(w) \) is the Jacobian matrix of \( g \) evaluated at \( w \), showing that the solution path defined by (2) starting at \( w \) is nonsingular for \( 0 \leq t \leq 1 \).

When the multiplicity is greater than one, deflation is a process of removing multiplicity to return to the multiplicity one case, e.g., see [9, 19, 14]. The following describes using isosingular deflation [14] in order to construct a witness system \( g' \) from \( g \) such that \( X \) has multiplicity 1 with respect to \( g' \). Define \( g_0 = g \) and consider the deflation operator \( \mathcal{D} \) with

\[
(g_{i+1}, w) = \mathcal{D}(g_i, w)
\]

where \( g_{i+1} \) consists of \( g_i \) and all \( (r+1) \times (r+1) \) minors of \( Jg_i(x) \) where \( r = \text{rank } Jg_i(w) \). The sequence \( s_i = \dim \ker Jg_i(w) \) is a nonincreasing sequence of nonnegative integers that
limits to \( \dim X \), i.e., there exists \( i^* \) such that \( s_i = \dim X \) for all \( i \geq i^* \). Hence, one can take \( g' = g_{i^*} \) to be used for path tracking along \( X \).

Next, we aim to encode the multiplicity structure of \( w \) with respect to \( f = \{g, \ell\} \) using a Macaulay dual space, e.g., see [9]. For \( \alpha \in \mathbb{Z}_{\geq 0}^N \), consider the differential

\[
\partial_\alpha = \frac{1}{\alpha!} \frac{\partial^{\mid \alpha \mid}}{\partial x^\alpha}
\]

where \( \alpha! = \alpha_1! \cdots \alpha_N! \) and \( \mid \alpha \mid = \alpha_1 + \cdots + \alpha_N \). Evaluating at \( w \) yields the linear functional \( \partial_\alpha[w] \) defined by

\[
\partial_\alpha[w](p) = (\partial_\alpha p)(w).
\]

Consider the infinite dimensional complex vector space \( D_w = \text{span} \{ \partial_\alpha[w] \mid \alpha \in \mathbb{Z}_{\geq 0}^N \} \).

The Macaulay dual space of \( f \) at \( w \) is the vector space

\[
D_w[f] = \{ \partial \in D_w \mid \partial((x - w)^\beta f_j) = 0 \text{ for all } \beta \in \mathbb{Z}_{\geq 0}^N \text{ and all } j \}.
\]

Since \( w \) is an isolated solution of \( f = 0 \), \( D_w[f] \) is a finite dimensional vector space whose dimension is equal to the multiplicity of \( w \) with respect to \( f \).

**Example 2.** Consider the polynomial system

\[
g(x, y, z) = \begin{bmatrix} 2xy + 2xz - 2yz - 1 \\ x^2 + y^2 + z^2 - 1 \end{bmatrix}
\]

for which \( X = \mathcal{V}(g) \) is an irreducible curve of degree 2 which has multiplicity 2 with respect to \( g \). For simplicity of presentation, consider \( \ell = x + y + z - 1 \) and \( w = (2, 1 + \sqrt{5}, 1 - \sqrt{5})/4 \in X \cap \mathcal{V}(\ell) \). Since \( \text{rank } Jg(w) = 1 \), one iteration of isosingular deflation yields

\[
g' = \begin{bmatrix} g \\ 4(x + y)(y - x + z) \\ 4(x + z)(y - x + z) \\ 4(y - z)(y - x + z) \end{bmatrix}
\]

with \( \dim \text{null } Jg'(w) = 1 = \dim X \). Additionally, for \( f = \{g, \ell\} \), \( w \) has multiplicity 2 with respect to \( f \) with Macaulay dual space

\[
D_w[f] = \text{span} \left\{ \partial_{(0,0,0)}[w], 2\sqrt{5}\partial_{(1,0,0)}[w] - (1 + \sqrt{5})\partial_{(0,1,0)}[w] + (1 - \sqrt{5})\partial_{(0,0,1)}[w] \right\}.
\]
2.3 Exactness recovery

For algebraic sets which are defined over the rational numbers, one often aims to recover exact rational numbers from numerical data. For example, one reason to do this is to be able to certify the result using Gröbner basis methods over the rational numbers.

The exactness recovery considered in [2] is: given a numerical approximation \( \tilde{w} \) of a generic point \( w \in X \subset \mathbb{C}^N \), compute exact polynomials \( p \) of degree at most \( d \) which vanish on \( X \). Thus, one aims to use lattice based methods to compute integer vectors \( c \) such that
\[
  c \cdot \nu_d(\tilde{w}) \approx 0
\]
where \( \nu_d(z) \) is the degree \( d \) Veronese embedding of \( z \), namely
\[
  \nu_d(z) = (1, z_1, \ldots, z_N, z_1^2, z_1 z_2, \ldots, z_N^2, \ldots, z_N^d).
\]

Example 3. Consider the numerical approximation
\[
  \tilde{w} = (0.5510119638, -0.2463075497, 0.7973195135)
\]
of a point \( w \in X \cap \mathcal{V}(x + \sqrt{2}y + z - 1) \) where \( X \subset \mathbb{C}^3 \) as in Ex. 2. Using PSLQ in Maple for \( d = 1 \) yields \( c = [0, -1, 1, 1] \) corresponding with \( y + z - x \) which does indeed vanish on \( X \) and is contained in \( \sqrt{\langle g \rangle} \) where \( g \) is as in (4). To search for other linear polynomials, we can repeat the PSLQ computation without the \( x \)-coordinate (since it is dependent on \( y \) and \( z \)) yielding \([2798, 5601, -1779]\) suggesting that there are no other vanishing linears with integer coefficients. Repeating with a 50-digit numerical approximation yields
\[
  [-47223690816349078, 4761102144861194, 60698860867498949].
\]

2.4 Sparsest nonzero null vector

In Section 2.3, lattice based methods are used to search for integer null vectors. To compute sparse polynomials, one is looking to compute sparse null vectors as described in Section 3.

For a vector \( x \in \mathbb{R}^N \), define
\[
  \|x\|_0 = \# \{ i \mid x_i \neq 0 \}.
\]
For a matrix \( A \in \mathbb{R}^{m \times k} \), the sparsest nonzero null vector solves
\[
  \min \{ \|x\|_0 \mid Ax = 0, \ x \neq 0 \}.
\]
Since one can arbitrarily rescale null vectors, one can pick a general coordinate patch to fix a scaling. That is, for a general vector \( v \in \mathbb{R}^N \), one can consider
\[
  \min \{ \|x\|_0 \mid Ax = 0, \ v \cdot x = 1 \}.
\]

Following a common technique in compressed sensing, e.g., see [11, 6, 10], one replaces the nonconvex optimization problem (6) with the convex optimization problem
\[
  \min \{ \|x\|_1 \mid Ax = 0, \ v \cdot x = 1 \}
\]
where \( \|x\|_1 = |x_1| + \cdots + |x_N| \). We say that (7) is an \( \ell_1 \)-relaxation of (6). In fact, under suitable conditions on the matrix \( A \), e.g., see [11, 6, 10], the solution to (7) solves (6).
3 Computing sparse polynomials

If \( X \subset \mathbb{C}^N \) is an algebraic set and for each irreducible component on \( X \) a witness set is known, the following describes how to compute all polynomials that vanish on \( X \) having degree at most \( d \) and at most \( t \) terms. For simplicity of presentation, we assume that \( X \) is irreducible but the method trivially extends to the reducible case by considering each irreducible component.

3.1 Multiplicity one

Suppose that \( X \) is described by a witness set \( \{g, \mathcal{L}, W\} \) where \( X \) has multiplicity one with respect to \( g \). Then, the key observation is that knowing \( t \) general points on \( X \) is enough to compute vanishing polynomials with at most \( t \) terms. This is well known in the case of \( t = 1 \), i.e., monomials, by simply looking to see which coordinates of one general point vanish.

**Theorem 4.** Suppose that \( q_1, \ldots, q_t \in X \) and

\[
A = \begin{bmatrix}
\nu_d(q_1) \\
\vdots \\
\nu_d(q_t)
\end{bmatrix}.
\]

Assume that \( p(x) = \sum_{i=1}^t c_i x^{\alpha_i} \) is a nonzero polynomial with at most \( t \) terms and \( \deg p \leq d \) which vanishes on \( X \). Then, the determinant of the \( t \times t \) submatrix of \( A \) whose columns correspond to the monomials \( x^{\alpha_1}, \ldots, x^{\alpha_t} \) vanishes.

**Proof.** Let \( A_{\alpha_1, \ldots, \alpha_t} \) be the \( t \times t \) submatrix of \( A \) whose columns correspond to \( x^{\alpha_1}, \ldots, x^{\alpha_t} \) and \( c = [c_1, \ldots, c_t]^T \). Since \( p \) vanishes on \( X \), it follows that \( c \) is a nonzero null vector of \( A_{\alpha_1, \ldots, \alpha_t} \). Hence, \( A_{\alpha_1, \ldots, \alpha_t} \) is rank deficient yielding \( \det A_{\alpha_1, \ldots, \alpha_t} = 0 \).

The following provides a method for finding vanishing polynomials with precisely \( t \) terms of degree at most \( d \).

**Theorem 5.** Suppose that \( q_1, \ldots, q_t \) are general points on \( X \) and

\[
A = \begin{bmatrix}
\nu_d(q_1) \\
\vdots \\
\nu_d(q_t)
\end{bmatrix}.
\]

Let \( A_{\alpha_1, \ldots, \alpha_t} \) be the \( t \times t \) submatrix of \( A \) whose columns correspond to \( x^{\alpha_1}, \ldots, x^{\alpha_t} \). If \( \text{rank} A_{\alpha_1, \ldots, \alpha_t} = t - 1 \) and \( c \in \mathbb{C}^t \) is a nonzero null vector of \( A_{\alpha_1, \ldots, \alpha_t} \) such that every \( c_i \neq 0 \), then \( p(x) = \sum_{i=1}^t c_i x^{\alpha_i} \) has precisely \( t \) terms with \( \deg p \leq d \) and vanishes on \( X \).

**Proof.** Let \( w \in X \) and \( B = \begin{bmatrix} A \\ \nu_d(w) \end{bmatrix} \). Assume that \( B_{\alpha_1, \ldots, \alpha_t} \) is the \( (t + 1) \times t \) submatrix of \( B \) whose columns correspond to \( x^{\alpha_1}, \ldots, x^{\alpha_t} \). Since \( \text{rank} A_{\alpha_1, \ldots, \alpha_t} = t - 1 \) and \( q_1, \ldots, q_t \) are general, it immediately follows that \( \text{rank} B_{\alpha_1, \ldots, \alpha_t} = t - 1 \) so that \( \text{null} A_{\alpha_1, \ldots, \alpha_t} = \text{null} B_{\alpha_1, \ldots, \alpha_t} \). Hence, \( B_{\alpha_1, \ldots, \alpha_t} \cdot c = 0 \) yields \( p(w) = 0 \) showing that \( p \) vanishes on \( X \). \( \square \)
Example 6. Let $X = \mathcal{V}(g) \subset \mathbb{C}^3$ be the twisted cubic curve with witness system

$$g(x, y, z) = \begin{bmatrix} y - x^2 \\ xy - z \\ xz - y^2 \end{bmatrix}.$$ 

Clearly, $X$ has multiplicity one with respect to $g$. Consider searching for all polynomials of degree at most $d = 2$ with at most $t = 2$ terms that vanish on $X$. One could utilize sampling (see Section 2.1) to compute 2 general points on $X$. For illustrative purposes, consider

$$q_1 = (\sqrt[4]{2}, \sqrt[4]{4}, \sqrt[4]{8}) \quad \text{and} \quad q_2 = (-\sqrt[4]{3}, \sqrt[4]{9}, -\sqrt[4]{27}).$$

Since each coordinate is nonzero, we know that there are no monomials that vanish on $X$. Hence, to search for all binomials of degree at most 2, we consider the $2 \times 10$ matrix

$$A = \begin{bmatrix} 1 & x & y & z & x^2 & xy & xz & y^2 & yz & z^2 \\ 1 & \sqrt[4]{2} & \sqrt[4]{4} & \sqrt[4]{8} & \sqrt[4]{16} & \sqrt[4]{16} & \sqrt[4]{32} & \sqrt[4]{64} \\ 1 & -\sqrt[4]{3} & \sqrt[4]{9} & -\sqrt[4]{27} & \sqrt[4]{81} & \sqrt[4]{81} & -\sqrt[4]{243} & \sqrt[4]{729} \end{bmatrix}.$$ 

Searching over the $2 \times 2$ minors yields 3 that vanish, corresponding to columns $\{y, x^2\}, \{z, xy\}$, and $\{xz, y^2\}$. Each corresponding $2 \times 2$ submatrix has rank 1 with null vector $c = [1, -1]^T$. This shows that the 3 polynomials in $g$ are the 3 binomials of degree at most 2 that vanish on $X$.

As alluded to in the proof of Theorem 5, adding rows to $A$ arising from other points on $X$ does not change the rank of any submatrix obtained by taking all rows and at most $t$ columns since the first $t$ rows of $A$ arise from $t$ general points of $X$. However, adding rows to $A$ does increase the rank of the overall matrix until it reaches the value of the Hilbert function of $X$ in degree $d$, e.g., see [13, 8]. After stabilization of the rank, the null space of the corresponding matrix is the vector space of all polynomials of degree at most $d$ that vanish on $X$.

One can utilize a modification of the (real) $\ell_1$-relaxation in (7) to compute sparse null vectors by reducing from complex linear algebra to real linear algebra by taking real and imaginary parts as needed. If the rank of the matrix is less than the corresponding value of the Hilbert function and this process computes a corresponding polynomial with more terms than general points utilized, there is no longer a guarantee that this polynomial vanishes identically on $X$. Hence, one needs to test this resulting polynomial at one additional general point to determine if it does indeed vanish identically on $X$. If, say, a binomial is found, one can then search to find other sparse polynomials by repeating with the matrix obtained by removing one of the columns corresponding to a monomial in the binomial.

Example 7. Reconsider the setup from Ex. 6 and let $A_k$ be the $k \times 10$ matrix with rows $\nu_2(x_i)$ for $i = 1, \ldots, k$ where $x_i \in X$ are general points. Since we are searching for binomials and the Hilbert function of $X$ in degree 2 is 7, i.e., a 3 dimensional linear space of quadratics vanish on $X$, we consider $k = 2, \ldots, 7$. As the solution obtained by (7) can be sensitive to
the choice of the random patch, we performed 100 random trials for each value of $k$. Table 1 summarizes the results of this experiment using linprog in Matlab with “success” indicating that one of the three binomials was found. This table shows that the success rate increased as $k$ increased, i.e., as more points were utilized.

Table 1: Frequency of $\ell_1$-relaxation successfully computing a binomial using $k$ points out of 100 random trials for the twisted cubic.

<table>
<thead>
<tr>
<th>$k$</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>successes</td>
<td>3</td>
<td>4</td>
<td>7</td>
<td>31</td>
<td>44</td>
<td>57</td>
</tr>
</tbody>
</table>

### 3.2 Multiplicity greater than one

When $X$ has multiplicity greater than one with respect to witness system $g$, one can utilize deflation (see Section 2.2) to produce a witness system which can be used for sampling (see Section 2.1). The following describes how to modify the setup from Section 3.1 to recover sparse polynomials vanishing on $X$ with the multiplicity structure imposed by $g$ via [13]. The key piece is to extend the definition of the Veronese embedding $\nu_d$ from (5) to one that depends upon both a point and a linear functional.

Let $z \in \mathbb{C}^N$ and $\partial \in D_z$ be a corresponding linear functional (as defined in (3)). Then, define the degree $d$ Veronese embedding of $z$ with respect to $\partial$ as (by abuse of notation)

$$\nu_d(z, \partial) = \partial(\nu_d(x)).$$

(8)

Thus, one applies $\partial$ (which includes evaluating at $z$) to the vector of monomials obtained from the Veronese embedding of degree $d$ of the vector of variables $x = (x_1, \ldots, x_N)$. This is a generalization of the Veronese embedding since

$$\nu_d(z, \partial_{(0, \ldots, 0)}[z]) = \nu_d(z).$$

With this setup, all items from Section 3.1 naturally extend to the case when the multiplicity is greater than one. In particular, the following are natural extensions of Theorems 4 and 5 whose proofs follow in the same manner and are thus omitted.

**Theorem 8.** Suppose that $q_1, \ldots, q_t \in X$, $\partial_1, \ldots, \partial_t$ such that $\partial_i \in D_{q_i}[f_{q_i}]$ where $f_{q_i} = \{g, \ell_{q_i}\}$ and $\ell_{q_i}$ consists of $\dim X$ linear polynomials with $\ell_{q_i}(q_i) = 0$, and

$$A = \begin{bmatrix} \nu_d(q_1, \partial_1) \\ \vdots \\ \nu_d(q_t, \partial_t) \end{bmatrix}.$$  

Assume that $p(x) = \sum_{i=1}^t c_i x^{\alpha_i}$ is a nonzero polynomial with at most $t$ terms and $\deg p \leq d$ which vanishes on $X$ as well as on the multiplicity structure imposed by $g$. Then, the determinant of the $t \times t$ submatrix of $A$ whose columns correspond to the monomials $x^{\alpha_1}, \ldots, x^{\alpha_t}$ vanishes.
Theorem 9. Suppose that $q_1, \ldots, q_t$ are general points on $X, \partial_1, \ldots, \partial_t$ such that $\partial_i \in D_{q_i}[f_{q_i}]$ is general where $f_{q_i} = \{g_1, \ell_{q_i}\}$ and $\ell_{q_i}$ consists of $\dim X$ general linear polynomials with $\ell_{q_i}(q_i) = 0$, and

$$A = \begin{bmatrix} \nu_d(q_1, \partial_1) \\ \vdots \\ \nu_d(q_t, \partial_t) \end{bmatrix}.$$  

Let $A_{\alpha_1, \ldots, \alpha_t}$ be the $t \times t$ submatrix of $A$ whose columns correspond to $x^{\alpha_1}, \ldots, x^{\alpha_t}$. If $\text{rank} A_{\alpha_1, \ldots, \alpha_t} = t - 1$ and $c \in \mathbb{C}^t$ is a nonzero null vector of $A_{\alpha_1, \ldots, \alpha_t}$ such that every $c_i \neq 0$, then $p(x) = \sum_{i=1}^t c_i x^{\alpha_i}$ has precisely $t$ terms with $\deg p \leq d$ and vanishes on $X$ as well as the multiplicity structure imposed by $g$.

Example 10. To illustrate, consider $X = \mathcal{V}(g) \subset \mathbb{C}^2$ where $X$ has multiplicity 2 with respect to $g(x, y) = (x - y)^2$. Every point $q \in X$ is of the form $q = (a, a)$ where $a \in \mathbb{C}$. The corresponding linear function $\ell_q$ has the form $\ell_q = b(x - a) + c(y - a)$. Thus, for the polynomial system $f_q = \{g, \ell_q\}$,

$$D_{q}[f_q] = \text{span}\{\partial(0,0)[q], c\partial(1,0)[q] - b\partial(0,1)[q]\}.$$  

Hence, a general point $q \in X$ with general element $\partial \in D_{q}[f_q]$ can be written as

$$q = (a, a) \text{ and } \partial = d\partial(0,0)[q] + e(c\partial(1,0)[q] - b\partial(0,1)[q])$$

where $a, b, c, d, e \in \mathbb{C}$ are general with

$$\nu_2(q, \partial) = (d, ad + ce, ad - be, a^2d + 2ace, a^2d - abe + ace, a^2d - 2abe).$$

It is easy to verify that the only linear relation amongst the entries of $\nu_2(q, \partial)$ that vanishes for all $a, b, c, d, e \in \mathbb{C}$ corresponds with $x^2 - 2xy + y^2 = (x - y)^2$ as expected.

4 Examples

In the following examples, Bertini [4] is used to generate sample points and Matlab is used to search for sparse polynomials either exhaustively searching minors or utilizing the $\ell_1$-relaxation in (7).

4.1 Random parameterization of twisted cubic

The classical twisted cubic parameterized by $t \mapsto (t, t^2, t^3)$ was considered in Examples 6 and 7. As a demonstration that our numerical methods work with varieties defined by polynomials with arbitrary complex coefficients, consider the twisted cubic parameterized by $t \mapsto (at, bt^2, ct^3)$ where

$$a = \sqrt{2} - \sqrt{-3}, \quad b = 1 + \sqrt{-1}, \quad \text{and} \quad c = 0.1653 - 0.9302\sqrt{-1}.$$
Hence, we consider $X = V(g) \subset \mathbb{C}^3$ where
\[
g(x, y, z) = \begin{bmatrix} bx^2 - a^2 y \\ cxy - abz \\ b^2xz - acy^2 \end{bmatrix} = 0.
\]

With the goal of finding all polynomials of degree at most $d = 2$ with at most $t = 2$ terms, searching over all $2 \times 2$ submatrices as described by Theorem 5 yields the three binomials in $g$.

Similar to Ex. 7, we then utilized \texttt{linprog} in \texttt{Matlab} to solve (7), doing 100 random trials for each $k = 2, \ldots, 7$ points. Table 2 summarizes the results of this experiment with “success” indicating that one of the three binomials was found. As with Table 1, the success rate increased as $k$ increased.

Table 2: Frequency of $\ell_1$-relaxation successfully computing a binomial using $k$ points out of 100 random trials for the random twisted cubic

<table>
<thead>
<tr>
<th>$k$</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>successes</td>
<td>0</td>
<td>9</td>
<td>30</td>
<td>52</td>
<td>59</td>
<td>71</td>
</tr>
</tbody>
</table>

4.2 Multiple component

The advantage of our approach based on sample points is that it is easy to switch between using the multiplicity structure and ignoring the multiplicity structure thereby computing sparse polynomials in the corresponding radical ideal. To demonstrate, consider the polynomial system
\[
g(x, y, z) = \begin{bmatrix} 2xy - 2y^2 + z^2 + 2x - y - 1 \\ x^2 - y^2 + z^2 + 2x - y - 1 \end{bmatrix}.
\]

The algebraic set $X = V(g)$ is a curve of degree 2 that has multiplicity 2 with respect to $g$. In order to sample points on $X$, we utilize $g'$ constructed using isosingular deflation (see Section 2.2) where
\[
g' = \begin{bmatrix} g \\ 2(x - y)(2x - 2y + 1) \\ 4z(x - y) \end{bmatrix}.
\]

Thus, consider searching for all polynomials of degree at most $d = 2$ with at most $t = 3$ terms for $X$ with respect to $g$ (multiplicity 2) and $g'$ (multiplicity 1).

Since no coordinates are zero at a general point on $X$, no monomial vanishes. Utilizing the multiplicity structure with respect to $g$, searching over the $2 \times 2$ and $3 \times 3$ submatrices as described by Theorem 9 found a single trinomial, namely $x^2 - 2xy + y^2$, which is easily observed to be the difference of the polynomials in $g$.

In the multiplicity one case, we first considered linear polynomials. This produced the binomial $x - y$. Since $x$ is dependent on $y$, we then searched for trinomials of degree at most 2.
in variables \( y \) and \( z \). This search produced the trinomial \( z^2 + y - 1 \). Table 3 summarizes the results.

<table>
<thead>
<tr>
<th>( n )</th>
<th>Without multiplicity</th>
<th>With multiplicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>( x - y )</td>
<td>(-)</td>
</tr>
<tr>
<td>3</td>
<td>( z^2 + y - 1 )</td>
<td>( x^2 - 2xy + y^2 )</td>
</tr>
</tbody>
</table>

We now turn to utilizing the \( \ell_1 \)-relaxation (7) to attempt to find sparse polynomials using the degree 2 Veronese embedding with \( k \) points for \( k = 2, \ldots, 9 \). A summary of the results of this experiment using 100 random trials for each \( k \) with \texttt{linprog} in \texttt{Matlab} utilizing multiplicity structure with respect to \( g \) is provided in Table 4. This table summarizes the number of terms of the corresponding polynomial when the computation successfully computed an optimizer and the frequency of the polynomial vanishing on \( X \). For example, using \( k = 3 \) points, out of the 100 trials, a trinomial that vanished on \( X \), namely, \( x^2 - 2xy + y^2 \), was computed 11 times. A 7-term polynomial was computed 4 times, but since the number of terms, namely 7, is more than the number of points, namely 3, one needs to check to see if this polynomial vanishes on \( X \) by testing at an additional general point. In this case, 2 out of the 4 vanished on \( X \). The vanishing 7-term polynomial computed is

\[
x^2 + 2xy - 3y^2 + 2z^2 + 4x - 2y - 2
\]

which is the sum of the 6-term polynomials in \( g \).

Finally, we repeat the same experiment without multiplicity with the results summarized in Table 5. Some examples of the binomials found in this experiment are

\[
x^2 - y^2, \quad x^2 - yz, \quad xz - yz
\]

with the trinomial \( x^2 - 2xy + y^2 \). Some quadrinomials and 5-term polynomials had the form

\[
\alpha(x - y) + xy - y^2 \quad \text{and} \quad \beta(z^2 + y - 1) - xy + y^2
\]

for some \( \alpha, \beta \in \mathbb{C} \).

### 4.3 Cubic-centered 12-bar

The final example arises from the cubic-centered 12-bar spherical linkage shown in Figure 2 and first presented in [23]. This linkage consists of rotational joints at the center of a cube and its vertices. Links connect along the edges of the cube and from each vertex to the center. To remove trivial motion in space, the center point, \( p_0 \), is fixed at the origin and two adjacent vertices are fixed, say \( p_7 = (-1,1,-1) \) and \( p_8 = (-1,-1,-1) \). This results in 18 variables arising from the three coordinates in each \( p_1, \ldots, p_6 \), say \( p_i = (x_i, y_i, z_i) \), with 17
Table 4: Summary of 100 random trials using $k = 2, \ldots, 9$ points with multiplicity structure.

<table>
<thead>
<tr>
<th>$k$</th>
<th># of terms</th>
<th>total successes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>1/1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>11/11</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>25/25</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>35/35</td>
<td>14/14</td>
</tr>
<tr>
<td>6</td>
<td>43/43</td>
<td>8/8</td>
</tr>
<tr>
<td>7</td>
<td>42/42</td>
<td>4/4</td>
</tr>
<tr>
<td>8</td>
<td>28/28</td>
<td>15/15</td>
</tr>
<tr>
<td>9</td>
<td>46/46</td>
<td>3/3</td>
</tr>
</tbody>
</table>

Table 5: Summary of 100 random trials using $k = 2, \ldots, 9$ points without multiplicity structure.

<table>
<thead>
<tr>
<th>$k$</th>
<th># of terms</th>
<th>total successes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>3/3</td>
<td>7/8</td>
</tr>
<tr>
<td>3</td>
<td>10/10</td>
<td>6/6</td>
</tr>
<tr>
<td>4</td>
<td>14/14</td>
<td>17/17</td>
</tr>
<tr>
<td>5</td>
<td>29/29</td>
<td>19/19</td>
</tr>
<tr>
<td>6</td>
<td>21/21</td>
<td>25/25</td>
</tr>
<tr>
<td>7</td>
<td>21/21</td>
<td>24/24</td>
</tr>
<tr>
<td>8</td>
<td>29/29</td>
<td>20/20</td>
</tr>
<tr>
<td>9</td>
<td>29/29</td>
<td>24/24</td>
</tr>
</tbody>
</table>

polynomials constraints, say $g = 0$ with:

$$\|p_i - p_j\|^2 - 4 = 0, \quad (i, j) \in \left\{ (1, 2), (1, 3), (1, 5), (2, 4), (2, 6), (3, 4), (3, 7), (4, 8), (5, 6), (5, 7), (6, 8) \right\},$$

$$\|p_k\|^2 - 3 = 0, \quad k = 1, \ldots, 6.$$ 

There are two types of curves in $\mathcal{V}(g)$. First, there are two degree 6 curves which are complex conjugates of each other with the linkage shown in Figure 2 being a point of intersection of these curves. In the interest of locating sparse polynomials, we consider one, say $X$, of the second type being a degree 4 complex curve that contains a real curve. A real point on $X$ is shown in Figure 3. We note that the multiplicity of $X$ with respect to $g$ is one.

Since each coordinate of a general point on $X$ is nonzero, no monomial vanishes on $X$. 13
Hence, searching for linear binomials yields the following 12:

\[ x_1 - x_4, \quad y_1 - y_4, \quad z_1 - 1, \]
\[ x_2 + 1, \quad y_2 + 1, \quad z_2 + 1, \]
\[ x_3 - x_5, \quad y_3 - y_5, \quad z_3 - z_5, \]
\[ x_6 + 1, \quad y_6 - 1, \quad z_6 + 1. \]

Due to these 12 linear binomials, we only need to consider higher degree polynomials in, say, 1, \( x_4, x_5, y_4, y_5, z_4, z_5 \). The search for degree 2 binomials using Theorem 5 located one:

\[ z_4 z_5 - x_4 x_5. \]

By removing the column corresponding to \( z_4 z_5 \) and searching for trinomials using Theorem 5
found the following 9:

\[ 2x_4x_5 + y_4y_5 - 1, \quad 2x_5 + z_4 + x_5y_4, \quad x_5 + 2z_4 - y_5z_4, \]
\[ 3x_5 + 2x_5y_4 + y_5z_4, \quad 2x_4 + z_5 - x_4y_5, \quad 3z_4 - x_5y_4 - 2y_5z_4, \]
\[ 3x_4 - 2x_4y_5 - y_4z_5, \quad x_4 + 2z_5 + y_4z_5, \quad 3z_5 + x_4y_5 + 2y_4z_5. \]

The search for quadrinomials found 61, some of which are:

\[ x_5 - y_5 + z_5 - 1, \quad x_4z_5 + x_5 + z_4 + 1, \]
\[ x_5^2 + y_5z_5 + x_5 - 1, \quad x_4z_5 - x_5y_4 - x_5 - 1, \]
\[ x_4z_5 - x_4y_4 + y_5z_4 + z_4^2, \quad x_5z_4 + x_4 - x_5 + y_5. \]

5 Conclusion

This paper described an approach based on using sample points to compute sparse polynomials that vanish on an algebraic set. In particular, sparse polynomials with at most \( t \) terms can be found using \( t \) sample points by searching over \( t \times t \) submatrices of the Veronese embedding of the sample points. One advantage of using sample points is that multiplicity structure can be easily be utilized via Macaulay dual spaces or ignored thereby computing sparse polynomials in the corresponding radical ideal. When the ideal is defined over the rational numbers, exactness recovery techniques can be used to find exact representations of the sparse polynomials which can then be certified using Gröbner basis techniques.

Since the number of entries of the Veronese embedding is combinatorially in the number of variables and degree, the bottleneck in our exhaustive submatrix search is the growth in the number of submatrices that one needs to consider. One approach for overcoming this is to solve an \( \ell_1 \)-relaxation (7). Experiments showed that using more points often resulted in successfully computing sparse polynomials using the \( \ell_1 \)-relaxation.

Acknowledgments

The second and third authors thank SLAM (Southwest Local Algebra Meeting) and its organizers as SLAM provided both time and a venue for discussions which helped initiate some ideas found in this paper. The first and fourth authors were partially supported by NSF CCF-1812746. The third author was partially supported by NSF CCF-1712788 and NSF ATD-1830676, The fourth author was also supported by the Schmitt Leadership Fellowship in Science and Engineering.

References


