Computing Sparse Projection Operators

Carlos D'Andrea* Ioannis Z Emiris‡

Abstract: We produce a new family of projection operators applying linear perturbations on sparse (or toric) resultants. These operators are not identically zero but vanish on all proper components of the zero set of the input equations after specialization of the coefficients, hence generalizing the sparse resultant. Our work addresses a major drawback of sparse resultant matrices, whose determinant may be identically zero for specific coefficients. The perturbation can be easily computed using these matrices. It also respects the notion of sparsity based on the polynomial supports and Newton polytopes. Our perturbations may be viewed as a generalization of Canny's Generalized Characteristic Polynomial (GCP) construction in the homogeneous case, while they provide a simpler and more efficient alternative to Rojas' toric GCP. It suffices to add linear terms in the perturbation variable, with random coefficients, for all monomials on the diagonal of any sparse resultant matrix formulation. We examine the practical usefulness of our approach through its implementation and its application to specific examples.

Keywords: polynomial system solving, sparse resultant, degeneracy, perturbation, generalized characteristic polynomial, mixed volume, Maple implementation.

1 Introduction

Matrix methods for polynomial system solving are based on various formulations of the polynomial resultant. They lie at the crossroads of algebraic and numeric computation and witness today a strong interest in what concerns algorithmic research, software development, as well as numerous applications [GKZ94, Stu94, CLO98, Roj99]. We concentrate on the sparse resultant and its matrix, also known as Newton matrix. It generalizes the coefficient matrix of a linear system, Sylvester's matrix and Macaulay's matrix, by strongly exploiting a model of sparseness which is defined below. The matrix determinant is, for generic coefficients, a nontrivial multiple of the sparse resultant. Yet, the Achilles' heel of the approach has been the possibility that the determinant may be identically zero under some coefficient specialization. This is important in various applications, as illustrated in section 7, and is precisely the issue we address. It also provides an essential algorithmic block towards completion of the sparse elimination project.

Let $\mathcal{A} = (\mathcal{A}_0, \mathcal{A}_1, \ldots, \mathcal{A}_n)$, with $\mathcal{A}_i \subset \mathbb{Z}^n$ be finite sets such that their union spans the affine lattice $\mathbb{Z}^n$. This is a technical assumption that can be eventually removed. Consider a system of generic Laurent polynomials (i.e. with integer exponents) in n variables $x = (x_1, \ldots, x_n)$ of the following form:

$$f_i(x) = \sum_{a \in \mathcal{A}_i} c_{ia}\ x^a, \quad c_{ia} \neq 0, \quad i = 0, 1, \ldots, n, \qquad (1)$$

The *sparse (or toric) resultant* $\mathrm{Res}_{\mathcal{A}}(f_0, \ldots, f_n)$ is a polynomial in $\mathbb{Z}[c_{ia}]$ which vanishes after a specialization of the coefficients iff the specialized system (1) has a solution in a suitable compactification of the torus $(\mathbb{C}^*)^n$ named *toric variety* [GKZ94]; here $\mathbb{C}^* := \mathbb{C} \setminus \{0\}$. It is possible to extend this condition to the existence of arbitrary affine roots; see [Roj99] and the references thereof. For the sake of simplicity, we consider coefficients in a field whose algebraic closure is $\mathbb{C}$, but this field can be arbitrary.

Let us consider system (1) as one in n variables and m parameters; the latter may be the input coefficients or any smaller parameter set on which the coefficients depend. The sparse resultant then defines a *projection operator* from a space of dimension $n + m$ to one of dimension m, and for generic coefficients it does not vanish [Stu94, CLO98]. Typically, it is expressed as a matrix determinant. This matrix also serves to reduce the *solution of a well-constrained system* of n polynomials to a matrix eigenproblem or to the factorization of a univariate polynomial. However, there are three

*Departamento de Matemática, F.C.E y N., UBA (1428) Buenos Aires, Argentina, cdandrea@dm.uba.ar. Partially supported by UBACYT TX 94.
‡INRIA, B.P. 93, Sophia-Antipolis 06902, France, emiris@sophia.inria.fr, http://www.inria.fr/saga/emiris.

cases where this construction fails:

- The case where the resultant is identically zero at the specialized coefficients due to lack of genericity, usually because the coefficients depend on few parameters.
- The well-constrained system has infinite solutions in the toric variety, forming "excess components", so that the sparse resultant cannot provide information on the isolated roots.
- The sparse resultant may be nonzero, but the matrix determinant that serves to express it may be identically zero due to a vanishing extraneous factor.

Canny [Can90] settled the dense homogeneous case by means of a homogeneous GCP. In the sparse context, the only existing approach focuses on the u-resultant and requires substantial additional computation [Roj99]. Recall that the u-resultant reduces the solution of systems of n equations to a determinant factorization by introducing a new polynomial [Emi96, CLO98].

We settle the sparse case in its generality by an efficient generalization of the GCP. The main contribution of this paper is a simple perturbation scheme that produces sparse projection operators defined by a resultant matrix determinant and which, after specialization, are non identically zero but vanish on roots in the proper components of the system's variety, including all isolated roots. Thus, we avoid all shortcomings mentioned above. Moreover, our technique can be coupled with the u-resultant. Our perturbation respects the structure modeled by supports and Newton polytopes, so that

- the input supports $\mathcal{A}_i$ are not incremented (i.e., each perturbation polynomial has support in a respective $\mathcal{A}_i$), and
- complexity measures can be derived in terms of the sparsity parameters.

The main algorithm can be summarized in a single sentence: add linear terms in a perturbation variable, with random coefficients, to all monomials on the diagonal of any sparse resultant matrix.

The next section overviews the most relevant previous work. A very general context of efficient (i.e. linear) perturbations is defined in section 3 to subsume all previous constructions. Section 4 discusses the combinatorial properties of sparse systems used to define sparse projection operators. In section 5 we propose an efficient algorithm for computing them by means of determinants, trade randomization for efficiency, bound the asymptotic complexity, and derive bounds on the number of roots from these operators. Optimality and its complexity are addressed in section 6. In section 7 we examine the practical usefulness of our approach through its implementation and its application to a system from robotics. The paper concludes with directions of further research.

2 Related Work

It has been shown that the sparse resultant of $f_0, \ldots, f_n$ is a homogeneous polynomial in the coefficients of each polynomial. Its degree in the coefficients of f_i is $\mathrm{MV}(f_0, \ldots, f_{i-1}, f_{i+1}, \ldots, f_n)$, which expresses the generic number of toric roots of this subsystem. $\mathrm{MV}(\cdot)$ is the *mixed volume* of the respective Newton polytopes, where the *Newton polytope* of a polynomial is the convex hull of the exponent vectors in the polynomial's support [GKZ94, Stu94, Emi96, CLO98].

In [Can90], a linear perturbation called *Generalized Characteristic Polynomial (GCP)* allows us to compute a projection operator which is not identically zero, but vanishes on all proper components of the system $f_i = 0$ in the dense case, i.e. when $\mathcal{A}_i = \{x \in \mathbb{N}_0^n : x_1 + \ldots + x_n \leq d_i\}$. Macaulay's formula for the multivariate resultant allows computation of this projection operator as a quotient of two matrices both having a linear perturbation in the diagonal (just like the characteristic polynomial of square matrices) [Can90]. Similar constructions had been examined by Chistov and Grigoryev, and Renegar.

For the special case of u-resultants in the toric context, [Roj99] introduces a toric GCP. This permits to solve sparse systems of n polynomial equations in n unknowns, by avoiding the problematic cases enumerated above. The algorithm for computing this GCP is based on *fillings* of the Newton polytopes which, though intricate, lie at the heart of the issue of efficient perturbations. One advantage, illustrated in section 7, is that the maximum degree of the perturbed determinant in the perturbation variable is usually smaller than with our construction. Unfortunately, this computation requires time exponential in the system dimension, namely in $O(n^{2.616} k^{2n+2})$, where k is the total number of support elements.

There are different matrix formulations yielding projection operators and exploiting structure; see, e.g. [KS96]. Our main algorithm is based on the sparse resultant matrix, also known as *Newton matrix,* based on a mixed subdivision, as defined in [CE93] and generalized by Canny and Pedersen, and Sturmfels [Stu94, CLO98]. We also apply our perturbation to the incremental Newton matrix algorithm [Emi94], which yields smaller matrices, optimal for all system classes where this is possible.

3 Linear Perturbations

The *raison d'être* of linear perturbations is that they minimize the degree in the perturbation variable, thus reducing complexity. Suppose we have a family $p := (p_0(x)\ldots, p_n(x))$ of Laurent polynomials such that $\mathrm{supp}(p_i) \subset \mathcal{A}_i$, and

$$\mathrm{Res}_{\mathcal{A}}(p_0, \ldots, p_n) \neq 0.$$

Note that the $\mathcal{A}_i$ are fixed at input, so the inclusion of $\mathrm{supp}(p_i)$ is usually proper.

Definition 3.1 *The toric Generalized Characteristic Polynomial (GCP) associated with the sequence $(p_0, \ldots, p_n)$ (p-GCP) is:*

$$C_p(t) := \mathrm{Res}_{\mathcal{A}}\left(f_0 - t.p_0, \ldots, f_n - t.p_n\right)$$

Proposition 3.2 *$C_p(t)$ is a polynomial in t of degree $r := \deg(\mathrm{Res}_{\mathcal{A}}(f_0, \ldots, f_n))$ and not identically zero. Its leading coefficient is $\mathrm{Res}_{\mathcal{A}}(p_0, \ldots, p_n)$.*

Proof. Because of the homogeneities of the sparse resultant, we have that

$$\begin{aligned} C_p(t) &= \mathrm{Res}_{\mathcal{A}}\left(t(\tfrac{1}{t}f_0 - p_0), \ldots, t(\tfrac{1}{t}f_n - p_n)\right) \\ &= t^r \mathrm{Res}_{\mathcal{A}}\left(\tfrac{1}{t}f_0 - p_0, \ldots, \tfrac{1}{t}f_n - p_n\right). \end{aligned}$$

In order to check that

$$\mathrm{Res}_{\mathcal{A}}\left(\frac{1}{t}f_0 - p_0, \ldots, \frac{1}{t}f_n - p_n\right) \neq 0,$$

set $y := \frac{1}{t}$, then it is clear that $\mathrm{Res}_{\mathcal{A}}\left(y.f_0 - p_0, \ldots, y.f_n - p_n\right) \neq 0$ because at $y = 0$, we have $\mathrm{Res}_{\mathcal{A}}(p_0, \ldots, p_n) \neq 0$. This is the coefficient of t^r in $C_p(t)$. Now the proposition follows straightforwardly. □

Below we examine particular cases where some $p_i = 0$ or 1. More are shown in section 7.

Corollary 3.3 (Dense case) *When $\mathcal{A}_i = \{x \in \mathbb{N}^n : x_1 + \ldots + x_n \leq d_n\}$, set $p_0 = 1$ and $p_i := x_i^{d_i}$, for $i = 1, \ldots, n$. Then $\mathrm{Res}_{\mathcal{A}}(1, x_1^{d_1}, \ldots, x_n^{d_n}) = 1$, and $C_p(t)$ is the GCP constructed in [Can90].*

The u-resultant of system $f_1, \ldots, f_n$ is the resultant of the overconstrained system where $f_0 := u_0 + u_1x_1 + \cdots + u_nx_n$ and $\mathcal{A}_0 = \{0, e_1, \ldots, e_n\}$. We can now significantly relax the original requirement that $\mathrm{Res}_{\mathcal{A}}(p_0, \ldots, p_n) \neq 0$:

Proposition 3.4 (u-resultant) *Consider the u-resultant construction with linear f_0. Require that $\mathrm{Res}_{\mathcal{A}}(f_0, p_1, \ldots, p_n) \neq 0$ and let $p_0 := 0$ in definition 3.1. Then $C_p(t)$ is a polynomial in t of degree $r := \deg(\mathrm{Res}_{\mathcal{A}}(f_0, \ldots, f_n)) - \mathrm{MV}(f_1, \ldots, f_n)$ and not identically zero. Its leading coefficient is $\mathrm{Res}_{\mathcal{A}}(f_0, p_1, \ldots, p_n)$.*

Proof. For any linear p_0 in definition 3.1, polynomial $f_0 - t.p_0$ can be substituted by f_0 by replacing every term $(u_i - tc_i)x_i$ by u_i, since the latter are indeterminates. The proof of proposition 3.2 can be adapted to this case, with the resultant at $y := 0$ becoming $\mathrm{Res}_{\mathcal{A}}(f_0, p_1, \ldots, p_n)$. This is nonzero, unless the parameters u_i are specialized so that system $f_0 = p_1 = \cdots = p_n = 0$. □

Example 3.5 [Rojas' toric GCP] Given are $f_1, \ldots, f_n$. Choose $p_1, \ldots, p_n$ with nonzero coefficients in an *irreducible fill* of the sequence $(\mathcal{A}_1, \ldots, \mathcal{A}_n)$ [Roj99]. Set $p_0 := 0$ and $f_0 := u_0 + u_1x_1 + \cdots + u_nx_n$, where $\{u_a\}$ are parameters, since we are in the case of the u-resultant. Using the main property of irreducible fills, one has $C_p(t) = \pm\mathcal{H}(u, t) \not\equiv 0$ where $\mathcal{H}(u, t)$ is the polynomial called toric GCP in [Roj99].

The following theorem states the main property of polynomials $C_p(t)$, and allows us to regard them as a generalization of the classical GCP [Can90]. We start with $n + 1$ polynomials $f_i\left(y_1, \ldots, y_m, x_1, \ldots, x_n\right)$ with support (viewed as polynomials in $\mathbb{C}\left[x_1^{\pm 1}, \ldots, x_n^{\pm 1}\right]$) contained in $\mathcal{A}_i$.

Let Z be the algebraic set $V\left(f_0, \ldots, f_n\right) \subset \mathbb{C}^m \times \mathcal{T}$, where $\mathcal{T}$ is the toric variety associated to the respective Newton polytopes. Recall that $\mathcal{T}$ is the closure of a monomial parametrization in a projective space $\mathbb{P}^N$ for some $N \in \mathbb{N}$ [GKZ94], so w.l.o.g. we can suppose that Z is an algebraic variety in $\mathbb{C}^m \times \mathbb{P}^N$.

Theorem 3.6 *Let W be a proper component of Z, so that the dimension of W is $m-1$. Let $C_p(y_1, \ldots, y_m)(t)$ be the p-GCP of definition 3.1. Arranging $C_p(t)$ in powers of t, let $C_{p,k}(y_1, \ldots, y_m)$ be its coefficient of lowest degree in t, namely k. If $\pi_y : \mathbb{C}^m \times \mathcal{T} \to \mathbb{C}^m$ denotes projection on the y_i coordinates, then $C_{p,k}\left(\pi_y(s)\right) = 0$, for all $s \in W$.*

Proof. Set $\widehat{f_i} := f_i - t.p_i$, where t is a new parameter. Consider $Z' := V(\widehat{f_0}, \ldots, \widehat{f_n}) \subset \mathbb{C}^{m+1} \times \mathcal{T}$. Since $\widehat{f_i}$ and f_i are identical when $t = 0$, the intersection of Z' and the hypersurface $t = 0$ is exactly $Z \times \{0\}$.

Because of the assumptions made on the sequence of supports $\{\mathcal{A}_0, \ldots, \mathcal{A}_n\}$ at the beginning, the toric variety $\mathcal{T}$ has dimension exactly n (cf [Stu94]), and using the immersion of $\mathcal{T}$ in $\mathbb{P}^N$, a toric version of the dimension lemma [Can90, lem 3.1] follows. Because of this, every component of Z' has dimension at least m. This implies that component W' containing $W \times \{0\}$ must be of dimension at least m.

Every point of W' has an m-dimensional neighborhood whose intersection with $\{t = 0\}$ is $(m-1)$-dimensional, therefore for every $s \in W \times \{0\}$ there is a point sequence $(s_j) \subset W' \setminus (W \times \{0\})$ converging to s. Let π_t denote projection on the t coordinate. Now

$\pi_t(s_j)^k \neq 0$ because we have chosen our s_j with last coordinate not equal to zero. Then $C_p(\pi_y(q))(\pi_t(q)) = 0$ for any point $q \in Z'$, therefore $C_p(\pi_y(s_j))(\pi_t(s_j)) = 0$ for all j. So we can divide the latter polynomial by $\pi_t(s_j)^k$ and obtain

$$C_{p,k}(\pi_y(s_j)) + \sum_{i>k} (\pi_t(s_j))^{i-k} C_{p,i}(\pi_y(s_j)) = 0.$$

$C_{p,i}$ is a polynomial in the coordinates of s_j, hence a continuous function of them, and vanishes for all $s_j \to s$. Then, it is also zero at s. Since $\pi_t(s) = 0$ the summation for $i > k$ vanishes and the theorem is proven. □

The coefficient $C_{p,k}$ of minimum degree in t in the p-GCP $C_p(t)$ of definition 3.1 is a suitable *projection operator,* i.e., a nonzero polynomial in the parameters y (typically the input coefficients), which vanishes when there is a solution of the specialized system (1) in some proper component. $C_{p,k}$ is not identically zero even after the specialization, thus fulfilling the first two conditions set in the Introduction.

4 Mixed Subdivisions and Systems with nonzero Resultant

In this section, we recall the mixed subdivision-based method for computing the sparse resultant from a matrix determinant. This discussion allows us also to construct a family $(p_0, \ldots, p_n)$ of sparse polynomials, with nonzero resultant.

For any $\omega := (\omega_0, \ldots, \omega_n) \in \mathbb{R}^m$, $\omega_i \in \mathbb{R}^{m_i}$, $init_\omega(\mathrm{Res}_{\mathcal{A}})$, the *initial term* of $\mathrm{Res}_{\mathcal{A}}(f_0, \ldots, f_n)$ is a well-defined polynomial in the coefficients of $f_0, \ldots, f_n$ which is a monomial if ω is generic [Stu94]. Let us pick a vector $\omega := (\omega_0, \ldots, \omega_n)$, which corresponds to functions $\omega_i : \mathcal{A}_i \to \mathbb{Q}$, for $i = 0, \ldots, n$. We impose the following condition:

$$\omega_i(a) \geq 0 \text{ for all } a \in \mathcal{A}_i,\ i = 0, \ldots, n, \qquad (2)$$

and suppose, as in [CE93], that ω_i is the restriction of an affine function. This hypothesis can be removed [Stu94].

Consider the lifted polytopes $Q_{i,\omega} := \mathrm{conv}\{(a, \omega_i(a)) : a \in \mathcal{A}_i\} \subset \mathbb{R}^{n+1}$, and their Minkowski sum $Q_\omega := Q_{0,\omega} + \ldots + Q_{n,\omega}$. Taking the *upper* envelope of Q_ω, we get a coherent polyhedral subdivision Δ_ω of the Minkowski sum $Q := Q_0 + \ldots + Q_n$, by means of the natural projection $\mathbb{R}^{n+1} \to \mathbb{R}^n$ of Q_ω. Moreover, this projection, constrained to the upper envelope of Q_ω, becomes a bijection. Suppose that our ω is sufficiently generic ([CE93, Stu94, Emi96]), then we have a tight mixed coherent decomposition, i.e. each facet in the upper envelope is of the form

$$F_\omega = F_{0,\omega} + \ldots + F_{n,\omega} \qquad (3)$$

where F_i is a face of Q_i and $n = \dim(F_0) + \ldots + \dim(F_n)$. This implies that at least one summand is zero-dimensional, i.e., a Newton polytope vertex. Furthermore, sum (3) maximizes the aggregate lifting of any sum of n faces equal to F_ω; so it is sometimes called an optimal sum. The cell F, (the projection of F_ω) is said to be *mixed of type i* or *i-mixed* if $\dim F_i = 0,\ \dim F_j > 0,\ j \neq i$.

We use the matrix construction based on mixed subdivision Δ_ω. Proceeding as in [CE93], we fix a small vector $\delta \in \mathbb{Q}^n$, and set $\mathcal{E} := \mathbb{Z}^n \cap (\delta + Q)$. We shall work with the following

Definition 4.1 *The* row content $RC(F) = (i, a)$ *for any cell of Δ_ω will be a pair (i, a) which satisfies: if $F_\omega = F_{0,\omega} + \ldots + F_{n,\omega}$ is the unique facet on the upper envelope of Q_ω projecting to F expressed by this optimal sum, let i be the* largest *index such that* $\dim(F_i) = 0,\ F_i = \{a\}$. *We also define the row content of all points $p \in F : p + \delta \in \mathcal{E}$ to be* $RC(p) = \mathrm{RC}(F)$.

For $i = 0, \ldots, n$, set:

$$\mathcal{B}_i := \{a \in \mathcal{A}_i : \mathrm{RC}(p) = (i, a),\ p \in \mathcal{E}\}.$$

Applying this construction, we will define a family of polynomials with nonzero resultant.

Definition 4.2 *Let $\lambda \in \mathbb{C}^*$ be a parameter. Then we set*

$$p_{\lambda,i} := \sum_{a \in \mathcal{B}_i} \lambda^{\omega_i(a)} x^a.$$

The following proposition ensures that $p_\lambda := (p_{\lambda,0}, \ldots, p_{\lambda,n})$ has nonzero resultant for generic λ.

Proposition 4.3 *For all but finite $\lambda \in \mathbb{C}^*$, we have that* $\mathrm{Res}_{\mathcal{A}}(p_\lambda) \neq 0$.

Proof. The initial form of $\mathrm{Res}_{\mathcal{A}}(f_0, \ldots, f_n)$ with respect to ω equals the monomial

$$\pm \prod_{i=0}^{n} \prod_{i-\mathrm{mixed} F} c_{ia_F}^{vol(F)}$$

where $F = F_0 + \ldots + \{a_F\} + \ldots + F_n$ and vol denotes ordinary euclidean volume [Stu94, theorem 2.1].

Expanding $\mathrm{Res}_{\mathcal{A}}(\sum_{a \in \mathcal{B}_i} \lambda^{\omega_i(a)} x^a)$ as a polynomial in λ and using condition (2), it is easy to check that the coefficient of its highest term equals ± 1, with exponent

$$\sum_{i=0}^{n} \sum_{i-\mathrm{mixed} F} \omega_i(a_F)\, vol(F) > 0$$

hence, $\mathrm{Res}_{\mathcal{A}}(p_\lambda)$ is not a zero polynomial in λ. Then, if λ is specialized outside the finite set of roots of this polynomial, $\mathrm{Res}_{\mathcal{A}}(p_\lambda) \neq 0$. □

There is a deterministic condition that guarantees the validity of a given λ, namely the computation of the sparse resultant matrix and its determinant for p_λ.

5 Perturbed Matrices

We have established nontrivial projection operators in sparse elimination theory. Now we wish to obtain efficient methods for computing them, namely as (nonzero) determinants of sparse resultant matrices. Deterministic as well as randomized methods are proposed for computing the coefficients in the perturbation polynomials. Lastly, we derive precise bounds on the number of roots for specialized coefficients, and estimate the complexity of our algorithms.

Construct the square matrix $\mathcal{M}_{\omega,\delta} := (m_{p,p'})_{p,p'\in\mathcal{E}}$ with $m_{p,p'} = c_{ia'}$ for $\mathrm{RC}(p) = (i,a)$ and $p - a + a' = p'$. Besides the row function defined above by means of the subdivision, the greedy variant of Canny and Pedersen or the generalized lifting in [Stu94] may be applied.

The matrix determinant is generically a nonzero multiple of $\mathrm{Res}_{\mathcal{A}}$ [CE93]. Its degree in the coefficients of f_i equals the number of rows containing this polynomial's coefficients. The particular definition of $\mathrm{RC}(\cdot)$ implies that the determinant degree in the coefficients of f_0 is $\deg_{f_0} \mathrm{Res}_{\mathcal{A}} = \mathrm{MV}(f_1,\ldots,f_n)$.

Theorem 5.1 *Let $\lambda \in \mathbb{C}^*$ and $p_{\lambda,i}$ as in definition 4.2. Then, for every specialization of $f_0,\ldots,f_n$, $\det \mathcal{M}_{\omega,\delta}\,(f_0 - t.p_{\lambda,0}, f_1 - t.p_{\lambda,1},\ldots,f_n - t.p_{\lambda,n})$ is a nonzero polynomial in t of degree $\#\mathcal{E}$. For almost all $\lambda \in \mathbb{C}^*$, this determinant equals $C_{p_\lambda}(t)$ multiplied by a polynomial in t, not depending on the coefficients of f_0 and $p_{\lambda,0}$, where $C_{p_\lambda}(t)$ is the p_λ-GCP.*

Proof. For any point $q \in Q$, let q_ω be its unique preimage on the upper envelope of Q_ω. We denote by $h(q) \in \mathbb{R}_{\geq 0}$ the non-negative distance (or height) between q_ω and q. For any Newton polytope vertex $a \in Q_i$, $h(a) = \omega_i(a)$. For every row p of $\mathcal{M}_{\omega,\delta} = (m_{p,q})_{p,q\in\mathcal{E}}$, where $\mathrm{RC}(p) = (i,a)$, let $\theta(p) := h(p) - h(a)$ and define

$$\mathcal{M}_\theta := (m^\theta_{p,q})_{p,q\in\mathcal{E}}, \quad m^\theta_{p,q} := \left(m_{p,q}\lambda^{\theta(p)}\right).$$

The key claim is that for every column q of $\mathcal{M}_\theta$, the largest power of λ that multiplies t appears at the diagonal element. By construction, assuming $\mathrm{RC}(q) = (j,b)$,

$$m^\theta_{qq} = c_{jb}\lambda^{\theta(q)} + t\lambda^{\omega_j(b)+\theta(q)}$$

hence the power of λ multiplying t has exponent equal to $h(q)$. Consider another element of the same column, say m^θ_{pq} at row p with $\mathrm{RC}(p) = (i,a)$: it is of the form $c_{ia'}\lambda^{\theta(p)} + t\lambda^{\omega_j(a')+\theta(p)}$ or simply $c_{ia'}\lambda^{\theta(p)}$. The latter does not depend on t so we concentrate on the former: its degree in λ equals $h(a') + h(p) - h(a)$ and $a' + p - a = q$. Thus q_ω and $a'_\omega + p_\omega - a_\omega$ lie on the same vertical in $\mathbb{R}^{n+1}$ but only q_ω can lie on the upper envelope of Q_ω by [CE93, lemma 5.5]. This is essentially due to picking the unique Minkowski sum to represent q which maximizes the aggregate lifting of the summands. Hence $h(q) > h(a') + h(p) - h(a)$ and our key claim is proven.

It is now clear that the product of the leading terms in the diagonal entries of $\mathcal{M}_\theta$ has degree $\sum_{p\in\mathcal{E}} h(p)$ in λ, degree $\#\mathcal{E}$ in t, and coefficient ± 1. Moreover, for $\mathrm{RC}(p) = (i,a)$, $\det \mathcal{M}_\theta(\lambda) = \det \mathcal{M}_{\omega,\delta}(\lambda) \prod_{p\in\mathcal{E}} \lambda^{h(p)-\omega_i(a)}$ which proves the first statement of the theorem.

By [CE93, Emi96, CLO98], $\det(\mathcal{M}_{\omega,\delta})$ is a nonzero multiple of the sparse resultant, and contains the optimal number of rows of the zeroth polynomial. The second statement of the theorem follows by noting that, for all values of λ outside a finite set, $C_{p_\lambda}(t) \neq 0$ by propositions 4.3 and 3.2. □

When $f_0 := u_0 + u_1x_1 + \cdots + u_nx_n$ with all u_i indeterminate, any perturbation can be "absorbed" in these indeterminates by $u_i \mapsto u_i - t\lambda^{\omega_0(a_i)}$ for the corresponding a_i. This demonstrates the following

Corollary 5.2 *In the case of the u-resultant of $(f_1,\ldots,f_n)$, f_0 is not perturbed and the previous theorem holds mutatis mutandis.*

The obvious construction of p_λ consists in picking a random value for $\lambda \in \mathbb{C}^*$. Thus only a few random bits are needed.

Proposition 5.3 (Derandomization) *It is possible to remove randomization completely by replacing $t \cdot p_{\lambda,i}$ in the perturbation by $p_{t,i}$, provided all powers of λ are positive.*

By replacing the value λ by the perturbation parameter t, this construction raises, in general, the perturbation degree in the resultant matrix and the subsequent computation. It thus has primarily theoretical value, but we return to the question of efficient deterministic perturbations in section 6.

Corollary 5.4 *If some very small probability of error is acceptable, then polynomials $p_{\lambda,i}$ do not have to depend on λ and their coefficients may be chosen completely at random. The probability that this choice be invalid equals the probability that the aggregate coefficient vector lies on the hypersurface defined by the vanishing of the determinant of the perturbation matrix. The same holds for the incremental construction of the sparse resultant matrix [Emi94] where supports $\mathcal{B}_i$ are defined as before.*

Let ϕ be the trailing coefficient of $\det \mathcal{M}_{\omega,\delta}(t)$. Then ϕ is a multiple of $C_{p,k}$. It is a polynomial in the u_i indeterminates in the case of the u-resultant. Otherwise, if we have "hidden" some variable x_i in the space of coefficients (as in [Emi96]), ϕ is a polynomial in x_i. Recall that the variety of system f has positive dimension iff the resultant vanishes identically. By theorem 3.6, it follows:

Corollary 5.5 (Root counting) *The degree of ϕ in the hidden variable or, in the case of the u-resultant, in any u_i indeterminate, bounds the number of isolated roots of the* specialized *system, which in some cases is tighter than the generic bound available by mixed volume or Bézout's theorem.*

If the conjecture of [Roj99] is true, this is a bound on the intersection multiplicity of the positive-dimensional components of the variety.

If the degree of ϕ in the coefficients of some polynomial f_j is strictly smaller than the respective degree of $\mathrm{Res}_{\mathcal{A}}$, namely $\mathrm{MV}(f_0, \dots, f_{j-1}, f_{j+1}, \dots, f_n)$, then the variety has positive dimension.

We now outline our algorithm:

1. Construct a sparse resultant matrix by any algorithm. If the determinant is nonzero, there is no need to proceed with a perturbation.

2. Construct sets $\mathcal{B}_i$ and define polynomials p_λ with coefficients which are either random (cor. 5.4), functions of some random λ (thm. 5.1), or deterministic (prop. 5.3).

3. Compute the perturbed matrix and the trailing coefficient ϕ of its determinant. By corollary 5.5, the degree of ϕ in any u_i or in hidden variable x_i bounds the number of isolated roots.

4. In the case of the u-resultant, factorization of $\phi(u)$ gives all the isolated roots. Otherwise, if we have hidden x_i, the zero set of $\phi(x_i)$ contains the coordinates of x_i at all isolated roots.

There are two general rules for *simplifying* the perturbation polynomials:

- By the definition of $\mathrm{RC}(\cdot)$, $\mathcal{M}_{\omega,\delta}$ contains the optimal number of rows of f_0. It is possible to define ω so that the only coefficient of f_0 on the diagonal is that of a specific monomial $a \in \mathcal{A}_0$ [Emi96, lemma 5.3], by requiring $\omega(a) \gg \omega(a')$ for all $a' \in \mathcal{A}_0$.

- Let $\mu_i := \min\{\omega_i(a) \,|\, a \in \mathcal{B}_i\}$. We obtain valid perturbation polynomials if we divide each $p_{\lambda,i}$ by λ^{μ_i}. This implies that when $\mathcal{B}_i = \{a\}$ is unitary, we obtain $p_{\lambda,i} = x^a$.

A more general approach for minimizing the perturbation degree is presented in the next section.

Computation of $\phi()$ is significantly more efficient if we avoid a direct determinant computation as a polynomial in t. To compute only the trailing coefficient, we may specialize all indeterminates to random prime values, including $t \mapsto r_0$. Let the specialized determinant be D. Then $D \equiv 0 \pmod{r_0^l} \Rightarrow l \leq k$ and $(D/r_0^l) \not\equiv 0 \pmod{r_0} \Rightarrow l \geq k$, with high probability. The converse direction of both implications is always true. This gives us a randomized method for computing k.

It is possible to *interpolate* $\phi()$ directly once k has been determined. Let us consider the case of one hidden variable x (several variables are treated analogously) and let $N = \dim \mathcal{M}$. T specializations $\phi(x = r)$ are needed, where $N \geq T \geq \deg_x \phi(x)$. A low bound T can usually be obtained from $\mathcal{M}$ (see section 7). Recovering the coefficients of $\phi(x)$ reduces to Vandermonde matrix computations (for r equal to different powers of some integer), with complexity dominated by the evaluations, so we focus on the latter.

Let $d = \max\{\deg_x f_i\}$ and $M(\cdot)$ express the arithmetic complexity of determinant computation. Then $B(r_0, r) = \det \mathcal{M}(r_0, r) r_0^{-k}$ is computed by *Chinese remaindering* with complexity $O^*(NdM(N))$, where polylogarithmic factors are ignored. $\phi(r) \equiv B(r_0, r) \pmod{r_0}$ can be computed for $O^*(T)$ values r_0 thus yielding $\phi(r)$. The overall *bit complexity* is $O^*(T^2 NdM(N))$.

Theorem 5.6 (Complexity) *All operations described in the above algorithm have asymptotic complexity dominated by the complexity of the existing matrix construction and root approximation algorithms, namely polynomial in* $\deg(\mathrm{Res}_{\mathcal{A}})$ *and exponential with a linear exponent in n [Emi94, Emi96].*

Remark 5.7 *In [Emi94] a conjecture is posed on the existence of a submatrix $\mathcal{M}^{nm}_{\omega,\delta}$ of $\mathcal{M}_{\omega,\delta}$, defined by the points in non-mixed cells, such that the quotient* $\det \mathcal{M}_{\omega,\delta} \,/\, \det \mathcal{M}^{nm}_{\omega,\delta}$ *equals the sparse resultant. This may be regarded as a generalization of a classical formula due to Macaulay for computing the multivariate resultant.*

If this is the case, it is easy to check that the p-GCP can be computed as a quotient of determinants. In other words, the extraneous factor in $\det \mathcal{M}_{\omega,\delta}(f_i - t.p_{\lambda,i})$ *is the determinant of the submatrix* $\mathcal{M}^{nm}_{\omega,\delta}(f_i - t.p_{\lambda,i})$. *This can be exploited for computing the sparse resultant of a specialized system when the determinant of $\mathcal{M}_{\omega,\delta}$ equals zero, by just computing the first non-zero coefficient of the p-GCP, like in [Can90].*

6 Optimal Perturbations

The goal of this section is to further decrease the degree of the perturbation and examine the intrinsic hardness of this question. The filling criterion of [Roj99] contributes in this direction, albeit with a high computational cost. We consider *optimal perturbations*, defined as those perturbations involving one monomial per polynomial, i.e. where every $\mathcal{B}_i$ is unitary.

Optimal perturbations exist in the completely dense case. For certain deterministic choices of the lifting, the row content function and the perturbation vector, the subdivision algorithm yields Macaulay's matrix [Emi94] and the $\mathcal{B}_i$ are unitary.

For the general case, write $Q_{-i} = \sum_{j=0,\dots,n,\, j\neq i} Q_j$ and $Q_{-i-k} = \sum_{j=0,\dots,n,\, j\neq i,k} Q_j$. For a vector $v \in \mathbb{R}^n$ and a polytope $P \subset \mathbb{R}^n$, we define the *extremal boundary* of P with respect to v, denoted $\partial_v P$, as the union of all facets of P which have nonempty intersection with the ray $p + \lambda v$ for $\lambda > 0$ and some p in the interior of P. For a set of vectors $K = \{v_1, \dots, v_k\} \subset \mathbb{R}^n$ (equivalently a non-negative cone in $\mathbb{R}^n$ defined by the vectors in K), the extremal boundary with respect to K is $\partial_K P = \cup_{i=1}^k \partial_{v_i} P$.

A necessary condition for an optimal perturbation is the following; the proof is omitted for lack of space.

Lemma 6.1 (Necessary condition for optimality) *Let $a_i \in \mathcal{A}_i$ be a distinguished element of the i-th support, $i = 0, \dots, n$. Assume $\mathcal{A}_0 \subset K$, where K is a non-negative cone in $\mathbb{R}^n$, and $a_0 = 0 \in \mathcal{A}_0$. If there exist δ and ω such that for every $p \in \mathcal{E}$, $RC(p) = (i, a) \Rightarrow a = a_i$, then $\partial_K Q_{-0} \subset \bigcup_{i=1}^n a_i + Q_{-0-i}$.*

Consider *multihomogeneous* systems defined by a single support. The variables are partitioned into r subsets, each of l_i variables, for $i = 1, \dots, r$. Clearly, $l_1 + \cdots + l_r = n$. Let d_j be the total degree in the variables of the j-th group; then the system is said to be of *type* $(l_1, \dots, l_r; d_1, \dots, d_r)$. We illustrate lemma 6.1 with the following

Example 6.2 Take the system of 3 polynomials of type $(1, 2; d_1, d_2)$ and let $\mathcal{A}_0 = \{0, e_1\}$. K is generated by $(1, 0, 0)$. Then $\partial_K Q_{-0} = \{(3d_1, a_2, a_3) : a_2, a_3 \geq 0, a_2 + a_3 \leq 3d_2\}$ is covered by 3 copies of $2Q_1$ shifted by $a_1 = (d_1, d_2, 0)$, $a_2 = (d_1, 0, d_2)$, $a_3 = (d_1, 0, 0)$. However, $\mathcal{A}_0 = \{0, e_2 + e_3\}$ implies $\partial_K Q_{-0} = [0, 3d_1] \times \{(a_2, a_3) : a_2 + a_3 = 3d_2\}$. Its four vertices $(0, 3d_2, 0), (0, 0, 3d_2), (3d_1, 3d_2, 0), (3d_1, 0, 3d_2)$ cannot be covered by any 3 shifts of $2Q_1$, and indeed the necessary condition fails. This condition again fails for $\mathcal{A}_0 = \{0, e_1 + e_2 + e_3\}$.

There are multihomogeneous systems of n polynomials for which *there is no* f_0 that would lead to an overconstrained system admitting an optimal perturbation. For instance, the multihomogeneous system of type $(1, 2; d_1, d_2)$, for any d_1, d_2, cannot have a mixed subdivision of $Q_{-0} = Q_1 + Q_2 + Q_3$ where the vertices involved in the facets' optimal sums include exactly one vertex per polynomial. This implies that, regardless of Q_{-0}, no mixed subdivision of Q can be defined using a single point per support.

7 Implementation and Examples

The perturbation has been implemented in MAPLE and can be found at:

`http://www.inria.fr/saga/emiris/soft_alg.html`

Example 7.1 Consider the system in [Stu94, exam. 2.1]:

$$\begin{aligned} f_0 &= \alpha_1 + \alpha_2 x_1^2 x_2^2 + \alpha_3 x_1 x_2^3, \\ f_1 &= \beta_1 + \beta_2 x_1^2 + \beta_3 x_1 x_2^2, \\ f_2 &= \gamma_1 x_1^3 + \gamma_2 x_1 x_2. \end{aligned}$$

Setting $\delta := (0, -1/3)$, and the affine lifting functions $\omega_1(a, b) = -a/2 - b/4 + 1$, $\omega_2(a, b) = 3a - 5b + 7$, $\omega_3(a, b) = 0$, this yields

$$\begin{aligned} p_{\lambda,0} &= \lambda + x_1 x_2^3, \\ p_{\lambda,1} &= \lambda^{13} x_1^2, \\ p_{\lambda,2} &= x_1^3 + x_1 x_2. \end{aligned}$$

An explicit computation shows that $\mathrm{Res}_{\mathcal{A}}(p_\lambda) = \pm\lambda^{93}$. The second simplification rule in section 5 allows to take $p_{\lambda,1} = x_1^2$, hence $\mathrm{Res}_{\mathcal{A}}(p_\lambda) = \pm\lambda^2$. It is clear that $\lambda = 1$ is a good choice, and produces a good perturbation. We could also apply proposition 5.3 to avoid using λ altogether and, instead, perturb by $t^2 + t x_1 x_2^3$, $t x_1^2$, $t x_1^3 + t x_1 x_2$.

Consider now specialization $\alpha_1 = \beta_3 = -2, \alpha_2 = \alpha_3 = \gamma_1 = -\gamma_2 = 1$, $\beta_1 = u, \beta_2 = 2 - u$. Determinant $\det \mathcal{M}(t)$ is a scalar multiple of the sparse resultant of the perturbed system. The constant coefficient being zero we conclude that the resultant of the input system vanishes for all u, hence the variety has positive dimension; indeed, it contains line $(1, 1, u) \simeq \mathbb{C}$. The linear coefficient $\phi(u)$ yields the 6 projections of the isolated roots of the system on the u-axis. $\phi(u)$ can be computed by interpolating on $T = 8$ terms, instead of $\dim \mathcal{M} = 21$ terms.

Example 7.2 Consider the system in [CE93]:

$$\begin{aligned} f_0 &= c_{11} + c_{12} x_1 x_2 + c_{13} x_1^2 x_2 + c_{14} x_1, \\ f_1 &= c_{21} x_2 + c_{22} x_1^2 x_2^2 + c_{23} x_1^2 x_2 + c_{24} x_1, \\ f_2 &= c_{31} + c_{32} x_2 + c_{33} x_1 x_2 + c_{34} x_1. \end{aligned}$$

Setting $\delta := (-3/8, -1/8)$, $L \geq l^2 > l \gg 0$, using the following lifting functions: $\omega_0 = (L - l)x_1 + (L - l^2)x_2$, $\omega_1 = (L + l^2)x_1 + (L + 1)x_2$, $\omega_2 = (L - 1)x_1 + (L + l)x_2$, we have that

$$\begin{aligned} p_{\lambda,0} &= 1, \\ p_{\lambda,1} &= \lambda^{4L+2l^2+2} x_1^2 x_2^2 + \lambda^{3L+2l^2+1} x_1^2 x_2, \\ p_{\lambda,2} &= \lambda^{L+l} x_2 + \lambda^{2L+l-1} x_1 x_2 + \lambda^{L-1} x_1; \end{aligned}$$

and

$$\pm\mathrm{Res}_{\mathcal{A}}(p_\lambda) = \lambda^{16L+3l^2+2} - 2\,\lambda^{17L+4l^2+l+2} + \lambda^{18L+5l^2+2l}.$$

Observe that $\lambda = 1$ is a zero of this polynomial, so another value must be chosen, for example, $\lambda = 1/b$ where b is an even integer. Of course, any random value would be valid with very high probability.

In the special case $f_0 = 1+x_1\,x_2+x_1^2\,x_2+x_1,\ f_2 = 1+x_2+x_1\,x_2+x_1$, the sparse resultant vanishes identically due to the fact that the variety $V(f_0, f_1)$ has positive dimension (it is formed by the union of the isolated point $(1, -1)$ and the line $\{-1\} \times \mathbb{C}$). With $L = 10$ and $l = 3$, we have that the coefficient of t^2 in the perturbed determinant is

$$\phi = -(c_{22}c_{23})(c_{24} - c_{21} + c_{22} - c_{23})(c_{24} + c_{21} - c_{22} + c_{23}),$$

and we can recover in the last two factors the value of f_1 at the isolated zero $(1, -1)$ and the point $(-1, -1)$ in the positive-dimensional component.

Example 7.3 This is the example of [Roj99, sect. 3.2]. To the system $1 + 2x - 2x^2y - 5xy + x^2 + 3x^3y,\ 2 + 6x - 6x^2y - 11xy + 4x^2 + 5x^3y$ we add $f_0 := u_1x + u_2y + u_0$, which does not have to be perturbed, by corollary 5.2. We use function `spares` to construct a 16×16 matrix `Mut` in parameters u, t. The number of rows per polynomial are, respectively, $4, 6, 6$, whereas the mixed volumes of the 2×2 subsystems are all equal to 4. For certain ω and δ, we have used $p_{\lambda,1} := -3x^2 + x^3y$, $p_{\lambda,2} := 2 + 5x^2$, not depending on any λ but instead with random coefficients, by corollary 5.4. The perturbed determinant has maximum and minimum degree in t, respectively, 12 and 1. $\phi(u_i)$ gives two factors corresponding to isolated solutions $(1/7, 7/4)$ and $(1, 1)$:

$$(49\,u_2 + 4\,u_1 + 28\,u_0)\,(u_2 + u_1 + u_0)\,.$$

Another two factors give points on the line $\{-1\} \times \mathbb{C}$ of solutions, but the specific points are very sensitive to changes in ω and δ.

$$(-u_0 + u_1)\,(27\,u_2 + 40\,u_1 - 40\,u_0)\,.$$

Rojas obtains a 17×17 matrix whose perturbed determinant has degree only 8 and the lowest term is again linear in t, though different from the one above.

Example 7.4 In the robot motion planning implementation of Canny's *roadmap algorithm* by H. Hirukawa, numerous "degenerate" systems are encountered. Let us examine a 3×3 system, where we hide x_3 to obtain dense polynomials of degrees $3, 2, 1$. Our `Maple` function `spares` applies an optimal perturbation to an identically singular 14×14 matrix in x_3. Now $\det \mathcal{M}_{\omega,\delta}(t)$ is of degree 14 and the trailing coefficient of degree 2, which provides a bound on the number of affine roots, tighter than Bézout's and the mixed volume, both equal to 6. We obtain

$$\phi(x_3) = \left(x_3 - \frac{1434}{625}\right)\left(x_3 - \frac{12815703325}{21336}\right),$$

the first solution corresponding to the unique isolated solution whereas the second one is superfluous, hence the variety has dimension zero and degree 1.

8 Future Work

Our current work studies the questions raised in section 6. One way is to extend work in [VGC96] on mixed cells to minimizing the lifting for the entire subdivision. Another is to exploit the concept of filling the Newton polytopes with simpler polytopes and, ideally, segments [Roj99].

Lastly, we should complete the solution of perturbed systems by matrix methods by applying the eigenvector computation to perturbed matrices. Recall that eigenvectors serve to compute the root coordinates corresponding to the eliminated variables.

REFERENCES

[Can90] J. Canny. Generalised characteristic polynomials. *J. Symb. Computation*, 9:241–250, 1990.

[CE93] J. Canny and I. Emiris. An efficient algorithm for the sparse mixed resultant. In *Proc. AAECC*, LNCS 263, pages 89–104, Berlin, 1993. Springer-Verlag.

[CLO98] D. Cox, J. Little, and D. O'Shea. *Using Algebraic Geometry.* Graduate Texts in Mathematics. Springer-Verlag, New York, 1998.

[Emi94] I.Z. Emiris. *Sparse Elimination and Applications in Kinematics.* PhD thesis, Computer Science, U.C. Berkeley, 1994.

[Emi96] I.Z. Emiris. On the complexity of sparse elimination. *J. Complexity*, 12:134–166, 1996.

[GKZ94] I. Gelfand, M. Kapranov, A. Zelevinsky. *Discriminants and resultants.* Birkhäuser, 1994.

[KS96] D. Kapur and T. Saxena. Sparsity considerations in Dixon resultants. In *Proc. ISSAC*, pages 187–194, ACM Press, 1995.

[Roj99] J.M. Rojas. Solving degenerate sparse polynomial systems faster. *J. Symbolic Computation*, 28:155–186, 1999.

[Stu94] B. Sturmfels. On the Newton polytope of the resultant. *J. Algebr. Combinatorics*, 3:207–236, 1994.

[VGC96] J. Verschelde, K. Gatermann, R. Cools. Mixed volume computation by dynamic lifting applied to polynomial system solving. *Discr. and Comput. Geometry*, 16(1):69–112, 1996.