August, 2009

Cayley-Dixon Projection Operator for Multi-Univariate Composed Polynomials

Arthur Chtcherba
Deepak Kapur, *University of New Mexico - Main Campus*
Manfred Minimair, *Seton Hall University*

Cayley-Dixon Projection Operator for Multi-Univariate Composed Polynomials

Arthur D. Chtcherba [a,1]

[a] *Bloomberg LP, 731 Lexington, New York, NY, USA*

Deepak Kapur [b,1]

[b] *University of New Mexico, Dept. of Computer Science, Albuquerque, NM, USA*

Manfred Minimair [c,1,2]

[c] *Seton Hall University,Dept. of Mathematics and Computer Science, South Orange, NJ, USA*

Abstract

The Cayley-Dixon formulation for multivariate projection operators (multiples of resultants of multivariate polynomials) has been shown to be efficient (both experimentally and theoretically) for simultaneously eliminating many variables from a polynomial system. In this paper, the behavior of the Cayley-Dixon projection operator and the structure of Dixon matrices are analyzed for composed polynomial systems constructed from a multivariate system in which each variable is substituted by a univariate polynomial in a distinct variable. Under some conditions, it is shown that a Dixon projection operator of the composed system can be expressed as a power of the resultant of the *outer* polynomial system multiplied by powers of the leading coefficients of the univariate polynomials substituted for variables in the outer system. A new resultant formula is derived for systems where it is known that the Cayley-Dixon construction does not contain any extraneous factor. The complexity of constructing Dixon matrices and roots at toric infinity of composed polynomials are analyzed.

Key words: Cayley-Dixon, resultant, polynomial composition

1 Introduction

Problems in many application domains, including engineering and design, graphics, CAD-CAM, geometric modeling, etc. can be modeled using polynomial systems (Sederberg and Goldman, 1986; Hoffman, 1989; Morgan, 1987; Kapur and Lakshman, 1992; Chionh, 1990; Zhang, 2000; Bajaj et al., 1988; Ponce and Kriegman, 1992; Kapur et al., 1994; Emiris and Mourrain, 1999; Coutsias et al., 2004; Busé et al., 2003; Emiris, 2005; Culver et al., 2004). Often a polynomial system arising from such an application has a structure. Particularly in engineering and design applications and in geometric modeling, a polynomial system can be expressed as a composition of two distinct polynomial systems, each of which is of much lower degree in comparison to the original system. If the structure of a given polynomial system is not known a priori, one can efficiently check if they can be decomposed (Rubio, 2000).

This paper addresses the resultant computation for such composed polynomial systems (Jouanolou, 1991; Cheng et al., 1995; Minimair, 2003a, 2002, 2001, 2003b,c, 2004; Hong and Minimair, 2002; Kapur and Saxena, 1997). The resultant of a polynomial system with symbolic parameters is a necessary and sufficient condition on its parameters for the polynomial system to have a common solution [4] . Resultant computations have been found useful in many application domains including engineering and design, robotics, inverse kinematics, manufacturing, design and analysis of nano devices in nanotechnology, image understanding, graphics, solid modeling, implicitization, CAD-CAM design, geometric construction, drug-design, and control theory.

The focus in this paper is on the Cayley-Dixon formulation for multivariate projection operators which has been shown to be efficient (both experimentally and theoretically) for simultaneously eliminating many variables from a polynomial system (Kapur and Saxena, 1995). The behavior of the Cayley-Dixon projection operator construction and the structure of Dixon matrices are analyzed for composed polynomial systems constructed from a multivariate system in which each variable is substituted by a univariate polynomial in a distinct variable, referred to as *multi-univariate* composition in (Rubio, 2000).

A new resultant formula is derived for multi-univariate composed polynomials where

Email addresses: arthur@cherba.org (Arthur D. Chtcherba), kapur@cs.unm.edu (Deepak Kapur), manfred@minimair.org (Manfred Minimair).

[1] Supported by NSF grants no. CCF-0729097, CCR-0203051 and a grant from the Computer Science Research Institute at the Sandia National Labs.

[2] Also supported by NSF grant CCF 0430741.

[3] Partial results of this paper appeared in proceedings of CASC'05.

[4] A resultant of a given polynomial system depends on an algebraic set in consideration in which the common solutions of the polynomial system are sought. (Buse et al., 2000).

it is known that the Cayley-Dixon projection operator formulation does not produce any extraneous factors for an outer system. The derivation unifies all known related results about resultants for multi-univariate composed polynomials in the literature (Kapur and Saxena, 1997; McKay and Wang, 1989). Such systems include n-degree, (Kapur et al., 1994), bivariate corner cut (Zhang and Goldman, 2000) and generalized corner cut systems (Chtcherba, 2003). Even when extraneous factors are present, a similar formula is derived showing that the extraneous factor of the outer system will be "amplified" in the extraneous factor of the composed system. Hence exploiting the composed structure of a polynomial system can reduce the extraneous factors in the resultant computation. Furthermore, it demonstrates that the resultant of a composed system can be effectively calculated by considering only the resultant of the outer system. For practical applications, that is what is needed. Since the complexity of a resultant computation is typically determined by the degree (and support) of the polynomial system, resultants of composed systems can be computed much faster by focusing only on the outer system.

Another byproduct of the above results is a new resultant formula for a multi-univariate composed polynomial system where the *outer* polynomial system is n-degree n-variable system (Saxena, 1997): it is proved that the resultant of the composed system is a power of the resultant of the outer system, multiplied by the powers of the leading coefficients of the univariate polynomials substituted for variables in the outer system. This formula generalizes the formula for the univariate case in (McKay and Wang, 1989) and considers a case not covered by the formula for the multivariate case in (Jouanolou, 1991; Cheng et al., 1995). It is important to point out that the techniques used for deriving the resultant formulas in this paper are different from the techniques used in previous works (Cheng et al., 1995; Jouanolou, 1991; Minimair, 2003a; Hong, 1997; Minimair, 2002, 2001, 2003b,c, 2004; Hong and Minimair, 2002; Kapur and Saxena, 1997). Previous techniques seemed to be not applicable in our setting.

Particularly, in case, the resultant matrix is singular for an outer system, the resultant matrix of the composed system is also singular. However, the rank sub-matrix construction used in Kapur et al. (1994) works on the resultant matrix of the composed system as well, giving a *projection operator* (see also Buse et al. (2000)). The rank of the resultant matrix of the composed system can be shown to be $\prod_{i=1}^{n} k_i$ times the rank of the outer polynomial system, where k_i is the degree of the univariate polynomial substituted for the respective variable in the outer system. Furthermore, the extraneous factor arising from the gcd of the determinant of the maximal minors of the Dixon matrix of the outer system appears as an extraneous factor in the determinant of the maximal minor of the Dixon matrix of the composed system (but raised to the power $\prod_{i=1}^{n} k_i$). Since the Dixon matrix of the composed system can be larger, there can be additional extraneous factors as well arising from each maximal minor of the Dixon matrix of the outer system.

In Chtcherba and Kapur (2003); Chtcherba (2003); Foo and Chionh (2004), conditions on the support of generic unmixed polynomial systems have been identified for which the Cayley-Dixon formulation generates resultants exactly (without any extraneous factor). The class of polynomial systems for which resultants can be computed exactly can be broadened by composing polynomial systems. More interestingly, it can be shown that the composed system of mixed supports can be generated from a unmixed outer system when univariate substitutions are made for distinct variables, thus establishing a class of mixed supports for which Dixon-Cayley construction yields resultants (without extraneous factors). This result about computing resultants of mixed systems without extraneous factors appears to be the first of its kind. Furthermore, it is also possible to compute resultants exactly for other outer polynomial systems obtained by functional decomposition of composed systems whose resultants can be computed exactly. This construction is illustrated using an example. Such an approach for identifying polynomial systems for which resultants can be computed exactly is novel and seems promising.

Below, we first state the main results of the paper. This is followed by section 2 where the generalized Cayley-Dixon formulation as proposed in Kapur et al. (1994) is briefly reviewed and discusses how the Cayley-Dixon resultant computation of a composed system obtained by composing $n + 1$ polynomials in n variables with a system of univariate polynomials is related to the various polynomials appearing in the composed system. The case when the Dixon matrix is singular or non-square is analyzed. Then, in section 4, a new resultant formula is derived for n-degree polynomials systems composed in multi-univariate manner. This is followed by a brief section where the example of a mixed composed system is discussed whose resultant can be computed exactly.

Since the Cayley-Dixon formulation involves two disjoint sets of variables, the bilinear form representation of a polynomial in disjoint sets of variables is useful. Extensive formalism is developed in Appendix A, where we discuss how bilinear forms are affected by polynomial operations, particularly when two polynomials are multiplied, a polynomial is composed with other polynomials by substituting variables by polynomials etc. To express these relations among bilinear forms, a series of matrix operations is introduced. Section A.2 illustrates in detail these operations in the case of univariate composed system, i.e., how the Cayley-Dixon resultant computation of a composed system obtained by composing two polynomials in a single variable with another univariate polynomial is related to the various polynomials appearing in the composed system. This construction is later stated in general terms for the multivariate case.

The discussion of the paper is self contained and only the proofs are dependent on the material in the Appendix A. Hence all of the proofs are presented in Appendix B. Detail oriented reader is encouraged to follow the discussion in the appendix A before proceeding with section 2.2.

Summarizing, the focus and scope of the current paper is given by the diagram in Figure 1. The paper studies how composition of a list of polynomials F with a list of polynomials G interacts with the constructions of Dixon polynomial, Dixon matrix, Projection operator and Resultant. The theorems in the current paper provide the dashed arrows making the diagram commute or certain fundamental properties of these arrows. A formula representing the first dashed arrow, relating the Dixon polynomials of composed polynomials $F \circ G$ and outer polynomials F, can be found on Page 10. Furthermore, Theorem 2.1 represents the second dashed arrow, relating the Dixon matrices, and Theorem 2.4 studies its complexity. Moreover, Theorems 2.2 and 2.3 provide the third dashed arrow, relating the projection operators. Finally, Theorems 2.5 and 4.1 provide properties of the fourth dashed arrow, connecting the resultants, and a precise representation for n-degree systems.

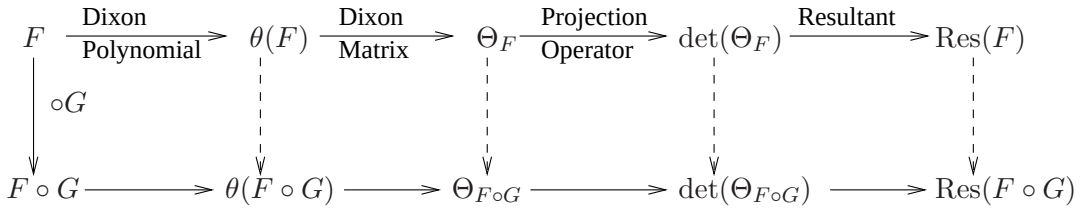


Fig. 1. How is this diagram commuting ?

It is important to point out that this paper uses some terms, including "resultant" and "projection operators", in a way as they are commonly used in the Dixon Resultant literature. This usage may differ in a subtle way from other literature, such as toric resultant literature. Therefore the next section elaborates on these and related terms.

1.1 Notions of resultant, projection operator, resultant matrix and exactness as used in Dixon resultant literature

Before we summarize the main results, we define the notions of resultant, projection operator and resultant matrix, as they are commonly used in the Dixon Resultant literature.

Let us start with resultant. The resultant of the multi-variate polynomial system $F = (f_0, f_1, \ldots, f_n)$ is usually defined with respect to a suitable variety, as is well-known. This variety is of significance because the existence of any common roots of the polynomials inside this variety implies that the resultant vanishes. Examples of such varieties are the projective space and, more generally, toric varieties and certain parametrized varieties (see e.g. Cox et al. (1998); Buse et al. (2000)). Thus we denote, up to a constant factor, the resultant of a polynomial system F, *with respect to a suitable variety*, by the symbol $\mathrm{Res}(F)$. In this sense, we also use the phrase "a quantity

r is Res(F)". That is, we mean that for a suitable variety the quantity r is the resultant of the polynomial system F. The reader might wonder why there is need for not stating the variety explicitly. The need comes from the nature of the construction of the Cayley-Dixon operator which is a multiple of or, as in certain cases, equal to the resultant. That is, the construction algorithm does not explicitly depend on the variety (see Section 2.1). Therefore in the literature it is common not to state the variety explicitly. However, whenever needed, we will make the variety explicit.

Next, a projection operator of the polynomial system F is a (not necessarily constant) multiple of the resultant. In this paper we study the Cayley-Dixon projection operator whose construction is described in Section 2.1.

Furthermore, a resultant matrix of the polynomial system F is a matrix having a minor whose determinant is a projection operator of F. Thus, for example, the Dixon matrix (see Section 2.1) is a resultant matrix.

If a resultant matrix is square with determinant being Res(F) then it is called exact. In this case, in order to emphasize that this resultant matrix is a Dixon matrix one uses the phrase Dixon-exact.

1.2 Main Results

Consider a polynomial system $F = (f_0, f_1, \ldots, f_n)$ with symbolic coefficients, where $F \subset \mathbb{K}[\mathbf{c}][y_1, \ldots, y_n]$ and

$$f_i = \sum_{\alpha \in \mathcal{F}_i} c_{i,\alpha} \mathbf{y}^\alpha \qquad \text{for} \quad i = 0, \ldots, n,$$

where $\mathbf{y}^\alpha = y_1^{\alpha_1} \ldots y_n^{\alpha_n}$ and $\mathcal{F}_i$ is the set of exponent vectors corresponding to the terms appearing in f_i, also called the *support* of f_i. The list $\mathbf{c}$ consists in "other" variables in terms of which the polynomial coefficients $c_{i,\alpha} \in \mathbb{K}[\mathbf{c}]$ are defined. They are also sometimes referred as the *parameters* of the polynomial system.

A polynomial system is called *generic* if there is no algebraic relation among the coefficients $c_{i,\alpha}$ of F.

Let $G = (g_1, \ldots, g_n)$ be another polynomial system in which each g_j, $j = 1, \ldots, n$, is a *univariate* polynomial in x_j, i.e.,

$$g_j = d_{j,k_j} x_j^{k_j} + d_{j,k_j-1} x_j^{k_j-1} + \cdots + d_{j,0}.$$

Let $k = (k_1, \ldots, k_n)$ be the degree vector of G.

It is possible to construct another polynomial system by **composing** F with G, written as $F \circ G$, which is the list of polynomials obtained from the list F of polynomials by replacing each y_j by g_j, respectively. The operator $\circ$ is called *functional composition* on polynomial systems.

The main results of this paper are:

(i) The Dixon matrix $\Theta_{F\circ G}$ of a composed system $F \circ G$ is shown to be a product of three matrices:

$$\Theta_{F\circ G} = A_{\rm L} \times {\rm Diag}_{k_1\cdots k_n}(\Theta_F) \times A_{\rm R},$$

where Θ_F is the Dixon matrix of the outer system F and the matrices $A_{\rm L}$ as well as $A_{\rm R}$ have triangular shape and contain only polynomials in terms of the coefficients of the polynomials in G. The matrix ${\rm Diag}_{k_1\cdots k_n}(\Theta_F)$ is block diagonal, where Θ_F, the Dixon matrix of F is repeated $k_1 \cdots k_n$ times along the diagonal. (Theorem 2.1)

(ii) For a polynomial system F, whenever the Dixon matrix determinant is ${\rm Res}(F)$, then

$$ {\rm Res}(F \circ G) = d_{1,k_1}^{\epsilon_1} \cdots d_{n,k_n}^{\epsilon_n} {\rm Res}(F)^{\delta},$$

where ϵ_j's depend on the degrees of G as well as F but δ depends only on the degrees of G. (Theorem 2.2)

(iii) The resultant of a composed n-degree system, with degrees $(m_1, \ldots, m_n)$, is

$$ {\rm Res}(F \circ G) = \left(d_{1,k_1}^{m_1} \cdots d_{n,k_n}^{m_n}\right)^{\frac{(n+1)!}{2}\, m_1\cdots m_n\, k_1\cdots k_n} {\rm Res}(F)^{k_1\cdots k_n}.$$

(Theorem 4.1; see also Theorem 5.1)

(iv) It is shown that one can construct, by composition, mixed systems of polynomials, for which the Cayley-Dixon construction yields their exact resultant without extraneous factors. (Section 3)

(v) Even if Θ_F is not square or is singular, the *rank sub-matrix construction (RSC)* introduced in Kapur et al. (1994) also works for composed systems (see also Buse et al. (2000)). In particular, the projection operator extracted from Θ_F is a factor of the projection operator extracted from $\Theta_{F\circ G}$ raised to the appropriate power; in addition to the leading coefficients d_{j,k_j} of the polynomials in G, there are also additional factors introduced in the projection operator extracted from $\Theta_{F\circ G}$. (Theorem 2.3)

(vi) Suppose that all the outer polynomials f_i have the same n-dimensional Newton polytope and the inner polynomials g_j's have the same degree k. Then

$$ {\rm O}_{\rm FG} = k^{n^2} \cdot {\rm O}_{\rm F},$$

where ${\rm O}_F$ and ${\rm O}_{FG}$ denote the complexity of constructing the Dixon matrix of F and respectively of the expanded composed polynomials $F \circ G$. This indicates that, under the assumptions of Item (ii) (Theorem 2.2), not making use of the

composition structure would result in a great loss in efficiency when constructing Dixon matrices. (Theorem 2.4)

(vii) A factor d_{j,k_j} in Items (ii)-(v) divides the toric (sparse) resultant of $F \circ G$ if all polynomials f_i contain all variables y_j. Thus the vanishing of any d_{j,k_j} implies that the composed polynomials have a common zero at toric infinity and thus none of the factors d_{j,k_j} are redundant under suitable conditions. (Theorem 2.5)

2 Cayley-Dixon Formulation of multi-univariate composition

2.1 The Cayley-Dixon Formulation

Dixon (1908) extended the Bézout-Cayley's construction for computing the resultant of two univariate polynomials to the bivariate case for three polynomials. Furthermore, Kapur et al. (1994) generalized this construction to the multivariate case. The concepts of a Dixon polynomial and a Dixon matrix were introduced. Below, the generalized multivariate Dixon formulation for simultaneously eliminating many variables from a polynomial system and computing its resultant are briefly reviewed. Let $\pi_i(\mathbf{y}^\alpha) = \overline{y}_1^{\alpha_1} \cdots \overline{y}_i^{\alpha_i} y_{i+1}^{\alpha_{i+1}} \cdots y_n^{\alpha_n}$, where $i \in \{0, 1, \ldots, n\}$, and $\overline{y}_i$'s are new variables; $\pi_0(\mathbf{y}^\alpha) = \mathbf{y}^\alpha$ and $\pi_n(\mathbf{y}^\alpha) = \overline{\mathbf{y}}^\alpha$. π_i is extended to polynomials in a natural way as: $\pi_i(f_j(y_1, \ldots, y_n)) = f_j(\overline{y}_1, \ldots, \overline{y}_i, y_{i+1}, \ldots, y_n)$.

Definition 2.1 *Given a n-variate polynomial system $F = (f_0, f_1, \ldots, f_n)$, where polynomial $f_i \subset \mathbb{K}[\mathbf{c}][y_1, \ldots, y_n]$, define its* ***Dixon polynomial*** *as*

$$\theta(F) = \prod_{i=1}^{n} \frac{1}{\overline{y}_i - y_i} \det \begin{pmatrix} \pi_0(f_0) & \pi_0(f_1) & \cdots & \pi_0(f_n) \\ \pi_1(f_0) & \pi_1(f_1) & \cdots & \pi_1(f_n) \\ \vdots & \vdots & \ddots & \vdots \\ \pi_n(f_0) & \pi_n(f_1) & \cdots & \pi_n(f_n) \end{pmatrix}.$$

Hence, $\theta(f_0, f_1, \ldots, f_n) \in \mathbb{K}[\mathbf{c}][y_1, \ldots, y_n, \overline{y}_1, \ldots, \overline{y}_n]$, where $\overline{y}_1, \ldots, \overline{y}_n$ are new variables. The matrix above is called the cancellation *matrix.*

The order in which original variables in $y_1, \ldots, y_n$ are replaced by new variables in $\overline{y}_1, \ldots, \overline{y}_n$ is significant in the sense that the computed Dixon polynomial can be different for two different orderings (see Dixon (1908); Kapur et al. (1994); Saxena (1997); Chtcherba (2003); Buse et al. (2000)).

It is well-known that polynomials, like the Dixon polynomial, whose variables are divided into two groups $y_1, \ldots, y_n$ and $\overline{y}_1, \ldots \overline{y}_n$, can naturally be represented by a matrix

("representation in bilinear form"). Subsequently, the notion of Dixon matrix is defined based on this observation.

Definition 2.2 *Let the Dixon polynomial* $\theta(f_0, f_1, \ldots, f_n)$ *be represented as the bilinear form*

$$\theta(F) \;=\; \overline{Y}^T \times \Theta_F \times Y,$$

where

$$\overline{Y} = \begin{pmatrix} \overline{\mathbf{y}}^{\beta_1} \\ \vdots \\ \overline{\mathbf{y}}^{\beta_k} \end{pmatrix} \quad \textit{and} \quad Y = \begin{pmatrix} \mathbf{y}^{\alpha_1} \\ \vdots \\ \mathbf{y}^{\alpha_l} \end{pmatrix}$$

are column vectors, where $\mathbf{y} = (y_1, \ldots, y_n)$ *and* $\overline{\mathbf{y}} = (\overline{y}_1, \ldots, \overline{y}_n)$ *are lists containing all the variables* y_i *and* $\overline{y}_i$ *of* $\theta(f_0, f_1, \ldots, f_n)$*, and where the* α_i*'s and* β_i*'s are the exponent vectors of all monomials in the variables in* $\mathbf{y}$ *and respectively in* $\overline{\mathbf{y}}$ *occurring in* $\theta(f_0, f_1, \ldots, f_n)$*. Then the* $k \times l$ *matrix* Θ_F *is called the* ***Dixon matrix****.*

Obviously, the matrix Θ_F is defined relative to the specific orderings of the monomials in $\overline{Y}$ and Y. Furthermore, notice that the entries in Θ_F are polynomials in the coefficients of the polynomials in F.

As shown in Kapur et al. (1994) and Buse et al. (2000), Θ_F is a resultant matrix. However, it can be singular especially for non-generic polynomial systems. In such a case, the resultant is extracted from the determinant of some maximal minor of Θ; this determinant is a *projection operator*, i.e. non-trivial multiple of the resultant (Kapur et al., 1994; Saxena, 1997; Buse et al., 2000).

2.2 Dixon matrix decomposition

Consider a polynomial system $F = (f_0, f_1, , \ldots, f_n)$, in variables $y_1, \ldots, y_n$. Let $G = (g_1, \ldots, g_n)$ be a list of *univariate* polynomials defined as

$$g_i = d_{i,k_i}\, x_i^{k_i} \;+\; d_{i,k_i-1}\, x_i^{k_i-1} \;+ \cdots +\; d_{i,0}, \qquad \text{for} \quad i = 1, \ldots, n,$$

of degrees $k_1, \ldots, k_n$, respectively, and let $\overline{G} = (\overline{g}_1, \ldots, \overline{g}_n)$, where $\overline{g}_j$ is obtained from g_j by replacing x_j with $\overline{x}_j$.

The **Cayley-Dixon Construction** of the composed polynomials $F \circ G$ is a generalization of the Cayley-Bézout construction from the univariate case. The Dixon polynomial

of the composed system

$$\theta_{F\circ G} = \frac{\det \begin{bmatrix} f_0 \circ (\pi_0(G)) & \dots & f_n \circ (\pi_0(G)) \\ \vdots & \ddots & \vdots \\ f_0 \circ (\pi_n(G)) & \dots & f_n \circ (\pi_n(G)) \end{bmatrix}}{\prod_{i=1}^{n} (x_i - \overline{x}_i)}$$

$$= \frac{\det \begin{bmatrix} f_0 \circ (\pi_0(G)) & \dots & f_n \circ (\pi_0(G)) \\ \vdots & \ddots & \vdots \\ f_0 \circ (\pi_n(G)) & \dots & f_n \circ (\pi_n(G)) \end{bmatrix}}{\prod_{i=1}^{n} (g_i - \overline{g}_i)} \times \frac{\prod_{i=1}^{n} (g_i - \overline{g}_i)}{\prod_{i=1}^{n} (x_i - \overline{x}_i)}$$

$$= \theta_F \circ (\overline{G}, G) \times \prod_{i=1}^{n} \frac{g_i - \overline{g}_i}{x_i - \overline{x}_i}.$$

In the Appendix A, detailed analysis of bilinear forms and effects of polynomial product and substitution are discussed. Using these results, given that the above is a product of two bilinear forms, by Lemmas A.1, A.2 and A.3, the following main theorem can be derived.

Theorem 2.1 *Let $F = (f_0, f_1, \dots, f_n)$ and $G = (g_1, \dots, g_n)$ be lists of generic polynomials. Then, the Dixon matrix $\Theta_{F\circ G}$ is*

$$A_{\mathrm{L}} \times \mathrm{Diag}_{k_1 \cdots k_n}(\Theta_F) \times A_{\mathrm{R}},$$

where $\mathrm{Diag}_{k_1 \cdots k_n}(\Theta_F)$ is block diagonal with $k_1 \cdots k_n$ blocks of Θ_F and moreover, matrices A_{L} and A_{R} are step-triangular (see Definition A.3) matrices (up to row/column permutation), whose entries depend only on coefficients of g.

In particular, for the generic n-degree polynomial system F and a generic system G of n polynomials used to substitute for variables $y_1, \cdots, y_n$ in F, the factors A_{L}, A_{R} and Θ_F in the above Theorem can be proved to be square and non-singular matrices (Kapur et al., 1994; Saxena, 1997). We investigate this in the next section.

More generally, if the factors are square in the above theorem, then we can derive a precise expression for the determinant of the Dixon matrix.

Let Δ_F be the support of the Dixon polynomial of the polynomial system F in terms of variables $x_1, \dots, x_n$, and similarly $\overline{\Delta}_F$ support in terms of $\overline{x}_1, \dots, \overline{x}_n$. Notice that the size of the A_{L} matrix is $|\Delta_{F\circ G}| \times |\overline{\Delta}_F| \cdot \prod_{j=1}^{n} k_j$ and size of A_{R} is $|\Delta_F| \cdot \prod_{j=1}^{n} k_j \times |\Delta_{F\circ G}|$.

The next lemma uses the same notation as in Theorem 2.1, and derives the determinants

of the factors of the Dixon matrix of the composed system.

Lemma 2.1 *If* $|\overline{\Delta}_F| \cdot \prod_{j=1}^{n} k_j = |\overline{\Delta}_{F\circ G}|$, *i.e.,* $A_{\rm L}$ *is square, then*

$$\det(A_{\rm L}) = \pm (d_{1,k_1}, \ldots, d_{n,k_n})^{(\sum_{\alpha\in\overline{\Delta}_F} \alpha)\, k_1\cdots k_n};$$

if $|\Delta_F| = |\overline{\Delta}_F|$, *i.e.,* Θ_F *is square, then*

$$\det\left(\mathrm{Diag}_{|\overline{\mathcal{Q}}|}(\Theta_F)\right) = (\det(\Theta_F))^{k_1\cdots k_n};$$

and if $|\Delta_F| \cdot \prod_{j=1}^{n} k_j = |\Delta_{F\circ G}|$, *i.e.,* $A_{\rm R}$ *is square, then*

$$\det(A_{\rm R}) = \pm (d_{1,k_1}, \ldots, d_{n,k_n})^{(|\Delta_F| + \sum_{\beta\in\Delta_F} \beta) k_1\cdots k_n}.$$

The above results holds even when the generic coefficients of f_i's and g_j's are specialized as long as the sizes of the matrices and their ranks are not lower than for generic coefficients. The interested reader can find technical derivation of this theorem in the appendix, (page 42). It is based on analyzing how composition and bilinear polynomial multiplication can be expressed in matrix forms.

Polynomial systems for which the Dixon matrix produces a projection operator without extraneous factors have been a topic of active research. Resultants for n-degree systems defined by (Kapur et al., 1994), multigraded (Chtcherba and Kapur, 2000), and corner-cut (Zhang and Goldman, 2000), Chionh (2001) can be computed efficiently using the Cayley-Dixon formulation. For such systems, by Theorem 2.1 and the above lemma, we have another main result of the paper.

Theorem 2.2 *Let* F *be a polynomial system such that* $\det(\Theta_F) = \mathrm{Res}(F)$. *Then under the multi-univariate polynomial composition* $F \circ G$,

$$\mathrm{Res}(F\circ G) = (d_{1,k_1}, \ldots, d_{n,k_n})^{(\sum_{\alpha\in\overline{\Delta}_F} \alpha + |\Delta_F| + \sum_{\beta\in\Delta_F} \beta)\, k_1\cdots k_n} \mathrm{Res}(F)^{k_1\cdots k_n}.$$

In other words if the polynomial system F is such that Cayley-Dixon construction computes the resultant without extraneous factors then also for the composition $F \circ G$ Cayley-Dixon produces no extraneous factors.

2.3 Rank Submatrix Construction

This subsection examines the cases when the Dixon matrix of the composed polynomials (or any of its factors in Lemma 2.1) is not square or when the Dixon matrix

is rank deficient. In such cases, one can extract a projection operator from the Dixon matrix by computing the determinant of any maximal minor (Kapur et al., 1994; Buse et al., 2000). Since the Dixon matrix $\Theta_{F\circ G}$ can be factored into a product, one obtains a similar factorization of a maximal minor,

$$\begin{aligned}\det_{\max}\left[A_{\mathrm{L}} \times \mathrm{Diag}_{k_1\cdots k_n}(\Theta_F) \times A_{\mathrm{R}}\right] \\ = \det\left[\max_{\mathrm{row}}(A_{\mathrm{L}}) \times \mathrm{Diag}_{k_1\cdots k_n}(\Theta_F) \times \max_{\mathrm{col}}(A_{\mathrm{R}})\right] \\ = \det\left[M_{\mathrm{L}} \times \mathrm{Diag}_{k_1\cdots k_n}(\Theta_F) \times M_{\mathrm{R}}\right],\end{aligned}$$

by selecting appropriate rows M_{L} of A_{L} and columns M_{R} of A_{R}. Furthermore, the well-known Cauchy-Binet formula allows us to expand the determinant of the minor into a sum of products of the form $l \cdot s \cdot r$, where l ranges over determinants of minors of M_{L}, s ranges over determinants of minors of $\mathrm{Diag}_{k_1\cdots k_n}(\Theta_F)$ and r ranges over determinants of minors of M_{R}.

More formally, given a square matrix M of size $m \times m$, where $M = T_1 \times D \times T_2$, and D is of size $s \times t$ for $m > s$ or $m > t$ then $\det(M) = 0$; otherwise, when $m = s = t$, then $\det(M) = \det(T_1) \cdot \det(D) \cdot \det(T_2)$. If $m < s$ or $m < t$, by the application of the Cauchy-Binet expansion of the determinant of the product of non-square matrices, we get

$$\det(T_1 \times D \times T_2) = \sum_{\substack{\sigma\in \mathtt{C}_m^s,\\ \rho\in \mathtt{C}_m^t}} \det(\mathrm{cols}_\sigma(T_1)) \cdot \det(\mathrm{submatrix}_{\rho,\sigma}(D)) \cdot \det(\mathrm{cols}_\rho(T_2)),$$

where $\mathtt{C}_m^s$ is the set of subsets of size m from set of $\{1, \ldots, s\}$.

The following elementary linear algebra result guarantees that the gcd of the determinants of all maximal minors of the matrix D will be a factor in any maximal minor determinant of M.

Proposition 2.1 *If $M = T_1 \times D \times T_2$ and the rank of T_1 equals the number of columns of T_1 and the rank of T_2 equals the number of rows of T_2, then* $\mathrm{rank}(M) = \mathrm{rank}(D)$.

Let us fix a notation for the remainder of this paper. By $\gcd(\det_{\max}(D))$, we denote the greatest common divisor of the determinants of all maximal minors of the matrix D.

The above proposition implies the following fact. If $\mathrm{rank}(D) = m$ and the matrices T_1 and T_2 are of rank s and t respectively, then $\gcd(\det_{\max}(D))$ is a factor of $\det(T_1 \times D \times T_2)$.

Using the above observation we can compute the determinant of the maximal minor of the Dixon matrix of the composed system by considering the maximal minors of A_L,

Θ_F and A_R. Since A_L and A_R are step-triangular (see Definition A.3) they are of full rank, that is their rank is equal to the minimum number of rows and columns. This leads to a formula similar to the one for the square case.

Theorem 2.3 *For a polynomial system $F = (f_0, f_1, \ldots, f_n)$, composed with univariate polynomials $G = (g_1, \ldots, g_n)$,*

$$\det_{\max}(\Theta_{F\circ G}) = d_{1,k_1}^{\epsilon_1} \cdots d_{n,k_n}^{\epsilon_n} E \left(\gcd \det_{\max}(\Theta_F) \right)^{k_1 \cdots k_n},$$

where E is an extraneous factor dependent not only on the coefficients of G but also that of F.

The above theorem establishes that whenever the resultant can be computed by the Cayley-Dixon construction, the resultant is also decomposable as shown above.

It is an open question what the values of $\epsilon_1, \ldots, \epsilon_n$ are in general and whether the factor E is constant for all the choices of maximal minors of $\Theta_{F\circ G}$.

To illustrate this further, consider an example of a composed polynomial system for which the Dixon matrix is non-square and for which the determinants of the maximal minors has the structure described above.

Example 2.1 [Maximal minor construction] Let

$$\begin{aligned} f_0 &= y_1\, y_2 \,+\, a\, y_1 \,+\, b\, y_2 \,+\, a\, b, \\ f_1 &= y_1\, y_2 \,+\, a\, y_1 \,+\, b\, y_2 \,+\, c, \\ f_2 &= y_1\, y_2 \,+\, y_1 \,+\, b\, y_2 \,+\, a, \end{aligned} \qquad \text{and} \qquad \begin{aligned} g_1 &= d_{1,2}\, x_1^2 \,+\, d_{1,1}\, x_1 \,+\, d_{1,0}, \\ g_2 &= d_{2,1}\, x_2 \,+\, d_{2,0}. \end{aligned}$$

For the composed polynomials $F \circ G$, the Dixon matrix $\Theta_{F\circ G}$ is a 4×2 matrix with rank 2. In this example, the determinant of any maximal minor of 4×2 matrix $\Theta_{F\circ G}$ is

$$-d_{21}^2 d_{12}^2 (a-1)^2 (ab-c)^2 (d_{12}b + d_{10}d_{12} - d_{11}^2).$$

By Theorem 2.1, $\Theta_{F\circ G} = A_{\mathrm{L}} \times \mathrm{Diag}_2(\Theta_F) \times A_{\mathrm{R}}$, where

$$A_{\mathrm{L}} = \begin{bmatrix} d_{1,2} & 0 & 0 & 0 \\ d_{1,1} & 0 & d_{1,2} & 0 \\ d_{1,0} & 1 & d_{1,1} & 0 \\ 0 & 0 & d_{1,0} & 1 \end{bmatrix}, \quad A_{\mathrm{R}} = \begin{bmatrix} 0 & d_{1,2}\, d_{2,1} \\ d_{1,2}\, d_{2,1} & d_{1,1}\, d_{2,1} \end{bmatrix},$$

$$\Theta_F = \begin{bmatrix} -a\, c \,+\, a^2\, b \,+\, c \,-\, a\, b \\ -a\, b\, c \,+\, b\, c - a\, b^2 \,+\, a^2\, b^2 \end{bmatrix} \quad = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}.$$

The 2×2 minor of $\Theta_{F\circ G}$ consisting, for instance, of the second and the third rows of A_{L} and the first two columns of A_{R} factorizes as:

$$M_L \times \mathrm{Diag}_2(\Theta_F) \times M_R = \begin{bmatrix} d_{1,1} & 0 & d_{1,2} & 0 \\ d_{1,0} & 1 & d_{1,1} & 0 \end{bmatrix} \times \begin{bmatrix} w_1 & 0 \\ w_2 & 0 \\ 0 & w_1 \\ 0 & w_2 \end{bmatrix} \times \begin{bmatrix} 0 & d_{1,2}\, d_{2,1} \\ d_{1,2}\, d_{2,1} & d_{1,1}\, d_{2,1} \end{bmatrix}.$$

By the Cauchy-Binet formula, the determinant of the above product is:

$$\begin{aligned} &\det(M_L \times \mathrm{Diag}_2(\Theta_F) \times M_R) \\ &\qquad = \det \begin{bmatrix} 0 & d_{1,2} \\ 1 & d_{1,1} \end{bmatrix} \cdot \det \begin{bmatrix} w_2 & 0 \\ 0 & w_1 \end{bmatrix} \cdot \det \begin{bmatrix} 0 & d_{1,2}\, d_{2,1} \\ d_{1,2}\, d_{2,1} & d_{1,1}\, d_{2,1} \end{bmatrix} \\ &\qquad + \det \begin{bmatrix} d_{1,1} & d_{1,2} \\ d_{1,0} & d_{1,1} \end{bmatrix} \cdot \det \begin{bmatrix} w_1 & 0 \\ 0 & w_1 \end{bmatrix} \cdot \det \begin{bmatrix} 0 & d_{1,2}\, d_{2,1} \\ d_{1,2}\, d_{2,1} & d_{1,1}\, d_{2,1} \end{bmatrix} \\ &\qquad = \gcd(\det_{\max}(\mathrm{Diag}_2(\Theta_F))) \cdot E_g = \gcd(\det_{\max}(\Theta_F))^2 \cdot E_g, \end{aligned}$$

where $E_g = (d_{12}b + d_{10}d_{12} - d_{11}^2)(-d_{21}^2 d_{12}^2)$. Note that the determinants of two maximal minors of Θ_F are

$$(a-1)(ab-c) \quad \text{and} \quad b(a-1)(ab-c),$$

and their greatest common divisor is $(a-1)\,(ab-c)$. The determinant of the maximal minor of $\Theta_{F\circ G}$ is

$$((a-1)\,(ab-c))^2\,(-d_{1,2}^2\, d_{2,1}^2)\,(-d_{1,2}\, b \,+\, d_{1,1}^2 - d_{1,2}d_{1,0}),$$

exhibiting that the greatest common divisor of the determinants of the maximal minors of Θ_F raised by 2 ($= k_1\, k_2$) is a factor.

In general, the factor E_g will contain an extra factor for each maximal minor selected from Θ_F; in this case it is 1 and b. $\square$

2.4 Complexity

In this section we investigate, for the unmixed case, how the results from this paper impact the efficiency of constructing Dixon matrices and projection operators of multi-univariate composed polynomials.

The main theorem of this section is given below. It compares the complexities of constructing the Dixon matrix of the composed polynomials $F \circ G$ and of the outer polynomials F. The complexity for the composed polynomials exceeds the complexity for the outer polynomial by a factor that is exponential in the dimension of the Newton Polytope of the outer polynomials and polynomial in the degree of the inner polynomials G. This indicates that not making use of the composition structure would result in a great loss in efficiency when constructing Dixon matrices.

Theorem 2.4 *Suppose that the outer polynomials f_i have all the same n-dimensional Newton polytope and the inner polynomials g_j's have the same degree k. Then*

$$\mathrm{O_{F\circ G}} = k^{n^2} \cdot \mathrm{O_F},$$

where O_F and $\mathrm{O}_{F\circ G}$ denote respectively the complexity of constructing the Dixon matrix of F and the expanded composed polynomials $F \circ G$. Therefore, under the assumptions of Theorem 2.2 (F is a polynomial system for which the Cayley-Dixon resultant formulation leads to a square and non-singular resultant matrix whose determinant is $\mathrm{Res}(F)$*), making use of the composition structure is by a factor of k^{n^2} more efficient when constructing Dixon matrices than not making use of the structure.*

Proof: Let us first determine O_F. For sufficiently large P (compare Erhart (1967)) the number of integer points in P is of order V, the normalized volume of P. Furthermore, by page 102 of Chtcherba (2003), the complexity of constructing the Dixon matrix of polynomials f_i is of order $n^2\, s^n$, where s is the number of integer points in the Newton polytope of the f_i's. Thus O_F is of order $n^2\, V^n$.

Now let us see how the number of integer points grows for the Newton polytope of the composed polynomials h_i obtained by composing the f_i's with the g_j's. Note that the Newton polytope of the composed polynomials h_i is k times the Newton polytope of the

f_i's. Therefore (Erhart (1967)) the number of integer points in the Newton polytope of the composed polynomials is of order k^n times V, where V is the normalized volume of the Newton polytope of the f_i's.

By the previous paragraphs, we have that $O_{F\circ G}$ is of order $n^2\,(k^nV)^n \;=\; n^2\,k^{n^2}\,V^n$. Therefore $\mathrm{O_{F\circ G}} \;=\; k^{n^2}\,(n^2V^n) \;=\; k^{n^2}\cdot \mathrm{O_F}$.

Furthermore, under the assumptions of Theorem 2.2 (F is a polynomial system for which the Cayley-Dixon resultant formulation leads to a square and non-singular resultant matrix whose determinant is $\mathrm{Res}(F)$), Theorem 2.2 tells that, when computing the resultant of $F \circ G$, one only has to construct the Dixon matrix of F. Therefore making use of the composition structure is by a factor of k^{n^2} more efficient when constructing Dixon matrices than not making use of the structure. □

Moreover, it is important to note that utilizing the composition structure also speeds up computing projection operators from the Dixon matrices. That is, the degrees of the polynomials F in each variable y_j is lower by the factor $k \;=\; k_j$ (under the assumptions of Theorem 2.4) than the corresponding degrees in the variables x_j of the composed polynomials $F \circ G$. Therefore the Dixon matrix of F is smaller by the factor k^n than the Dixon matrix of $F \circ G$. This implies that the projection operators can be extracted more efficiently from the Dixon matrix of F.

2.5 Cayley-Dixon formulation does not introduce new extraneous factors

It is shown in Buse et al. (2000) that the toric (sparse) resultant of a system of polynomials F is a factor of the Dixon resultant of F. Therefore the question arises if the leading coefficients d_{j,k_j} of the g_j's observed as factors of the Dixon resultant of composed polynomials $F \circ G$ in Theorems 2.2 and 2.3 are also factors of the toric resultant of the composed polynomials. The following Theorem 2.5 answers this question affirmatively.

Theorem 2.5 also tells that if a coefficient d_{j,k_j} vanishes then the composed polynomials have a common root at toric infinity, as is discussed later. In Section 4 we give an example of such a root in the context of toric homogenization (Cox et al., 1998) for n-degree polynomials.

Theorem 2.5 *A factor d_{jk_j} of the Dixon projection operator of the composed polynomials $F \circ G$ in Theorems 2.2 and 2.3 divides the toric (sparse) resultant of $F \circ G$ if all polynomials f_i contain all variables y_j. Therefore the vanishing of any d_{j,k_j} implies that the composed polynomials have a common zero at toric infinity and thus none of the factors d_{j,k_j} are redundant.*

Note that the toric resultant in the above theorem is considered with respect to the generic supports of the composed polynomials. Furthermore, the composed polynomials are considered as naturally defined over the toric domains induced by their generic supports (Rojas, 1999a,b).

The proof is provided in the appendix in section B.2 where a few necessary lemmas for proof are also defined.

3 Exact Mixed Systems

As the following example illustrates, resultants of composed systems can be computed exactly under certain conditions using the Cayley-Dixon construction if the outermost system in a composed system is such that its resultant can be computed exactly using the Cayley-Dixon construction. Chtcherba and Kapur (2003) identified a class of generic unmixed polynomial systems for which the resultant can be computed without extraneous factor using the Cayley-Dixon method. By composing such an unmixed system F with G, it is possible to compute resultants without extraneous factors of generic as well as specialized mixed systems as well. This opens up a promising area of research as very little is known about the subclass of mixed or non-generic systems for which resultants can be computed exactly.

Consider the following unmixed bivariate polynomial system with non-generic f_0 and generic f_1 and f_2:

$$f_0 = a\,b + a\,y_1 + b\,y_2 + y_1\,y_2,$$
$$f_1 = c_1 + c_2y_1 + c_3y_2 + c_4y_1y_2,$$
$$f_2 = d_1 + d_2y_1 + d_3y_2 + d_4y_1y_2.$$

This system F has unmixed bidegree support. Moreover, its Dixon polynomial has the same support as the Dixon polynomial of generic bidegree polynomials of the same bidegree as F. Therefore the determinant of the Dixon matrix of F is exactly its resultant $\mathrm{Res}(f_0, f_1, f_2)$ (with respect to the toric variety induced by their supports), given in factored form by the product

$$(-c_3d_1 + bc_3d_2 - bd_4c_1 + b^2d_4c_2 - bd_3c_2 + bc_4d_1 - b^2c_4d_2 + c_1d_3)$$
$$(-c_2d_1 + c_2d_3a - ad_4c_1 + d_4a^2c_3 + ac_4d_1 - c_4d_3a^2 - d_2c_3a + d_2c_1).$$

Consider now a substitution $G = (g_1 = x_1^2 + rx_1 - b, g_2 = x_2 - a)$, where r is an arbitrary symbolic parameter. The composed system is neither generic nor unmixed:

$$
\begin{aligned}
f_0 &= x_1^2 x_2 + r x_1 x_2, \\
f_1 &= c_1 - c_3 a - c_2 b + c_4 ab + (c_2 - c_4 a) x_1^2 + (-c_4 b + c_3) x_2 \\
&\quad + (c_2 r - c_4 ar) x_1 + c_4 x_1^2 x_2 + c_4 r x_1 x_2, \\
f_2 &= d_1 - d_3 a - d_2 b + d_4 ab + (d_2 - d_4 a) x_1^2 + (-d_4 b + d_3) x_2 \\
&\quad + (d_2 r - d_4 ar) x_1 + d_4 x_1^2 x_2 + d_4 r x_1 x_2.
\end{aligned}
$$

The determinant of the Dixon matrix of the above composed system is $\mathrm{Res}(f_0, f_1, f_2)^2$ (with respect to the toric variety induced by their supports). Therefore we get the exact (modulo sign) resultant of $F \circ G$. However, the determinant of the Dixon matrix corresponding to the generic system whose supports is the same as that of $F \circ G$, i.e., the generic system

$$
\begin{aligned}
f_0 &= e_1 x_1^2 x_2 + e_2 x_1 x_2, \\
f_1 &= c_1 + c_2 x_1^2 + c_3 x_2 + c_2 x_1 + c_{1,21} x_1^2 x_2 + c_4 x_1 x_2, \\
f_2 &= d_1 + d_2 x_1^2 + d_3 x_2 + d_2 x_1 + c_{2,21} x_1^2 x_2 + d_4 x_1 x_2,
\end{aligned}
$$

has an extraneous factor

$$
(c_2^2 d_1 d_2 + d_2^2 c_1^2 + d_1^2 c_2^2 + c_2 d_2^2 c_1 - c_1 c_2 d_2 d_2 - 2 d_1 c_2 c_1 d_2 - d_1 c_2 c_2 d_2) \cdot (d_1 c_3 - d_3 c_1)
$$

along with the resultant.

Hence, there exist a condition on the coefficients of the polynomial system to have exact Dixon matrices. So far researchers have investigated only conditions on the support of polynomial system to have Dixon-exact matrices.

This raises a question of whether for any non-exact system, there is a transformation into exact (possibly non-generic) system.

4 Resultant of Composed n-degree polynomial system

In this section, we generalize the McKay and Wang formula A.6 (shown on page 38) for the univariate case to multivariate n-degree polynomials systems. (The interested reader is also referred to Jouanolou (1991); Cheng et al. (1995) who give formulas for projective resultants of multi-variate total-degree composed polynomials.)

Consider $(m_1, \ldots, m_n)$-degree generic polynomials $F = (f_0, f_1, \ldots, f_n)$ where

$$f_j = \sum_{i_1=1}^{m_1} \cdots \sum_{i_n=1}^{m_n} c_{j,i_1,\ldots,i_n} y_1^{i_1} \cdots y_n^{i_n} \qquad \text{for} \qquad j = 0, 1, \ldots, n,$$

with generic coefficients $c_{j,i_1,\ldots,i_n}$ and variables $y_1, \ldots, y_n$.

It is easy to see that the composed polynomials $f_i \circ (g_1, \ldots, g_n)$, $0, \leq i \leq n$, are $(m_1k_1, \ldots, m_nk_n)$-degree as well.

The support of the Dixon polynomial for the n-degree polynomial system F is

$$\overline{\Delta}_F = \{\, \alpha \in \mathbb{N}^n \mid \alpha_i < (n-i+1)m_i \quad \text{for } i = 1, \ldots, n \,\},$$

$$\Delta_F = \{\, \alpha \in \mathbb{N}^n \mid \alpha_i < im_i \quad \text{for } i = 1, \ldots, n \,\},$$

and therefore $|\overline{\Delta}_F| = |\Delta_F| = n!\, m_1 \cdots m_n$.

Applying Lemma 2.1, the sum of all points in the above support for a particular coordinate $i \in \{1, \ldots, n\}$ is

$$\begin{aligned}
\sum_{\alpha \in \overline{\Delta}_F} \alpha_i &= nm_1\, (n-1)m_2 \cdots (n-i+2)m_{i-1} \left(\sum_{j=0}^{(n-i+1)m_i - 1} j \right) (n-i)m_{i+1} \cdots m_n \\
&= n!\, m_1 \cdots m_n \frac{(n-i+1)m_i - 1}{2}, \\
\sum_{\alpha \in \Delta_F} \alpha_i &= m_1\, 2m_2 \cdots (i-1)m_{i-1} \left(\sum_{j=0}^{im_i - 1} j \right) (i+1)m_{i+1} \cdots nm_n \\
&= n!\, m_1 \cdots m_n \frac{im_i - 1}{2}.
\end{aligned}$$

Substituting into Lemma 2.1,

$$\begin{aligned}
&\det \left[\mathrm{Diag}_{\overline{\mathcal{Q}}}(\Theta_F) \right] \;=\; (\det(\Theta_F))^{k_1 \cdots k_n}, \\
&\det [A_L] = \prod_{i=1}^{n} d_{i,\,k_i}^{\,n!\, m_1 \cdots m_n \frac{(n-i+1)m_i - 1}{2}\, k_1 \cdots k_n}, \\
&\det [A_R] = \prod_{j=1}^{n} d_{j,\,n_j}^{\,(n!\, m_1 \cdots m_n + n!\, m_1 \cdots m_n \frac{im_i - 1}{2})\; k_1 \cdots k_n}.
\end{aligned}$$

Note that, if F and G are generic, then the coefficients of $F \circ G$ will still not have any algebraic relations, and therefore the system $F \circ G$ is generic. By Theorem 2.2 and the

fact that the Dixon matrix is exact for generic n-degree systems, we have another main result of the paper.

Theorem 4.1 *For the unmixed n-degree case,*

$$\mathrm{Res}(F \circ G) = \left(d_{1,k_1}^{m_1} \cdots d_{n,k_n}^{m_n}\right)^{\frac{(n+1)!}{2}\, m_1 \cdots m_2\, k_1 \cdots k_n} \mathrm{Res}(F)^{k_1 \cdots k_n}.$$

5 Roots at toric infinity and multi-homogeneous resultants

In this section we study systems of composed polynomials for which a coefficient d_{j,k_j} vanishes. (Recally that d_{j,k_j} is the leading coefficient of the polynomial g_j in the composed polynomials $F \circ G$.) By Theorem 2.5, if $d_{j,k_j} = 0$, for some index j, then the system of composed polynomials have a common root at toric infinity. Section 7 of Cox et al. (1998) shows how to construct such roots as non-trivial roots of toric homogenizations of the composed polynomials. Toric homogenizations of n-degree polynomials are similar to the usual homogenizations of total-degree polynomials. For total-degree polynomials, roots at infinity are those, for which their leading forms vanish, or, equivalently, non-trivial roots of the homogenized polynomials for which the homogenizing variable vanishes.

For toric n-degree polynomials the analogous toric homogenization is obtained by homogenizing with respect to each variable individually, thus, introducing n different homogenizing variables. Roots at toric infinity for n-degree polynomials are non-trivial roots for which a homogenizing variable vanishes. (Here, non-trivial root means a root for which no pair of variable and homogenizing variable vanishes.) It is also interesting to note that this construction shows that the toric resultant of n-degree polynomials equals the so-called multi-homogeneous resultant of the multi-homogeneous polynomials constructed from the n-degree polynomials.

Example 5.1 [Toric homogenization of n-degree polynomial] Let $f = 3\,y_1^2\,y_2 + 7\,y_1^2 + 2\,y_1\,y_2 - 8\,y_1 - 5\,y_2 + 9$. Then the toric homogenization of f is

$$f^{\mathrm{h}_{1,2}} = 3\,y_1^2\,y_2 + 7\,y_1^2\,z_2 + 2\,y_1\,y_2\,z_1 - 8\,y_1\,z_1\,z_2 - 5\,y_2\,z_1^2 + 9\,z_1^2\,z_2,$$

that is, f homogenized with respect to y_1 and homogenizing variable z_1 and with respect to y_2 and homogenizing variable z_2.

As expected, non-trivial roots are those for which no pair (y_j, z_j) vanishes.

Example 5.2 [Trivial and non-trivial root] For instance, $(y_1, y_2, z_1, z_2) = (1, 0, 1, 0)$ is a

trivial root of the toric homogenization in Example 5.1. However any tuple $(y_1, y_2, z_1, 0)$ with $y_2 \neq 0$ and $(y_1, z_1) \neq 0$ that is a root of $f^{\mathrm{h}_{1,2}}(y_1, y_2, z_1, 0) = 3\,y_1^2\,y_2 + 2\,y_1\,y_2\,z_1 - 5\,y_2\,z_1^2$ is a *non-trivial* root.

Similar to total-degree polynomials, roots at toric infinity are those for which a homogenizing variable z_j vanishes. Equivalently, one can consider a root at toric infinity as the root of certain leading forms.

EXAMPLE 5.3 [Roots at infinity] Consider a point $(y_1, y_2, 0, z_2)$. If it is a root of the toric homogenization

$$f^{\mathrm{h}_{1,2}} = 3\,y_1^2\,y_2 + 7\,y_1^2\,z_2 + 2\,y_1\,y_2\,z_1 - 8\,y_1\,z_1\,z_2 - 5\,y_2\,z_1^2 + 9\,z_1^2\,z_2$$

from Example 5.1, then it is a root of $f^{\mathrm{h}_{1,2}}(y_1, y_2, 0, z_2) = 3\,y_1^2\,y_2 + 7\,y_1^2\,z_2$. This polynomial is a "leading form" of f with maximal power in y_1, that is, y_1^2. Similarly, $f^{\mathrm{h}_{1,2}}(y_1, y_2, z_1, 0) = 3\,y_1^2\,y_2 + 2\,y_1\,y_2\,z_1 - 5\,y_2\,z_1^2$ is the "leading form" of f with maximal power in y_2, that is, y_2^1.

For the case of multi-univariate composed polynomials for which the leading coefficient d_{j,k_j} of g_j, for some j, vanishes, we will see in Theorem 5.1 that roots at toric infinity are of the form $(x_1, \ldots, x_n, z_1, \ldots, z_n)$ with $x_j \neq 0$ and $z_j = 0$. The other x_i's and z_i's can be chosen arbitrarily as long as no pair (x_i, z_i) vanishes.

EXAMPLE 5.4 [Root at infinity of composed polynomial] We continue with the polynomial from Example 5.1. Let

$$\begin{aligned} g_1 &= d_{1,2}\,x_1^2 + 3\,x_1 + 1 \\ g_2 &= 7\,x_2 - 8. \end{aligned}$$

Then the homogenizations of g_1 and g_2 are

$$\begin{aligned} g_1^{\mathrm{h}} &= d_{1,2}\,x_1^2 + 3\,x_1\,z_1 + z_1^2 \\ g_2^{\mathrm{h}} &= 7\,x_2 - 8\,z_2. \end{aligned}$$

Since y_j is replaced with g_j in the composed polynomials, above we use the homogenizing variable z_j for y_j also for homogenizing g_j. With this notation,

$$(f \circ (g_1, g_2))^{\mathrm{h}_{1,2}} = f^{\mathrm{h}_{1,2}}(g_1^{\mathrm{h}}, g_2^{\mathrm{h}}, z_1^2, z_2).$$

This formula can easily be verified. See also Lemma 5.3.

By substituting the polynomials, we obtain

$$\begin{aligned}(f \circ (g_1, g_2))^{\mathrm{h}_{1,2}} = {} & 3\,(d_{1,2}\,x_1^2 + 3\,x_1\,z_1 + z_1^2)^2\,(7\,x_2 - 8\,z_2) \\ & + 7\,(d_{1,2}\,x_1^2 + 3\,x_1\,z_1 + z_1^2)^2\,z_2 \\ & + 2\,(d_{1,2}\,x_1^2 + 3\,x_1\,z_1 + z_1^2)\,(7\,x_2 - 8\,z_2)\,z_1^2 \\ & - 8\,(d_{1,2}\,x_1^2 + 3\,x_1\,z_1 + z_1^2)\,z_1^2\,z_2 - 5\,(7\,x_2 - 8\,z_2)\,z_1^4 \\ & + 9\,z_1^4\,z_2.\end{aligned}$$

For $d_{1,2} = 0$, the homogenized composed polynomial has non-trivial roots $(y_1, y_2, 0, z_2)$ with $(y_2, z_2) \neq 0$,

$$\begin{aligned}(f \circ (g_1, g_2))^{\mathrm{h}_{1,2}} = {} & 3\,(0\,x_1^2 + 3\,x_1\,0 + 0^2)^2\,(7\,x_2 - 8\,z_2) \\ & + 7\,(0\,x_1^2 + 3\,x_1\,0 + 0^2)^2\,z_2 \\ & + 2\,(0\,x_1^2 + 3\,x_1\,0 + 0^2)\,(7\,x_2 - 8\,z_2)\,0^2 \\ & - 8\,(0\,x_1^2 + 3\,x_1\,0 + 0^2)\,0^2\,z_2 - 5\,(7\,x_2 - 8\,z_2)\,0^4 \\ & + 9 \cdot 0^4\,z_2 \\ = {} & 0,\end{aligned}$$

because $g_1(y_1, 0) = 0$ and $z_1 = 0$.

Now we formalize our observations. In constructing the toric homogenization for *composed* n-degree polynomials, we first determine the toric homogenization of n-degree (not-necessarily composed) polynomials. Second, we formalize the meaning of non-trivial root. Third, we derive a formula for toric homogenization of composed polynomials. Then we construct a common root with a vanishing homogenizing variable for this toric homogenization.

The following lemma follows immediately from results in Cox et al. (1998).

Lemma 5.1 *The toric homogenization (Cox et al., 1998) of an $(m_1, \ldots, m_n)$-degree polynomial f is $z_1^{m_1} \cdots z_n^{m_n} \cdot f(\frac{y_1}{z_1}, \ldots, \frac{y_n}{z_n})$.*

For the remainder of this section we fix the following notation. Let

$$f^{\mathrm{h}_{1,\ldots,n}} = f^{\mathrm{h}_{1,\ldots,n}}(y_1, \ldots, y_n, z_1, \ldots, z_n)$$

denote the toric homogenization of the n-degree polynomial f as given in the previous lemma. Furthermore, we abbreviate f^{h_1} by f^{h} if f is univariate because in this case the toric homogenization agrees with the usual homogenization usually denoted by a superindex h.

The next lemma specifies when a root of a toric homogenization is called non-trivial.

Lemma 5.2 *A root of the toric homogenization of f is non-trivial iff*

$$(y_1, z_1) \neq 0, \ldots, (y_n, z_n) \neq 0.$$

The proof of Lemma 5.2 is easy and is left to the reader.

The next lemma studies how toric homogenization interacts with composition. It turns out that toric homogenization "commutes" with composition.

Lemma 5.3 *Let f be a $(m_1, \ldots, m_n)$-degree polynomial. Then*

$$(f \circ (g_1, \ldots, g_n))^{\mathrm{h}_{1,\ldots,n}} = f^{\mathrm{h}_{1,\ldots,n}}(g_1^{\mathrm{h}}(x_1, z_1), \ldots, g_n^{\mathrm{h}}(x_n, z_n), z_1^{k_1}, \ldots, z_n^{k_n}),$$

where the homogenizations in the above equation carried out with respect to the generic degrees $(m_1, \ldots, m_n)$ and $k_1, \ldots, k_n$ of f and respectively of $g_1, \ldots, g_n$.

The proof of Lemma 5.3 is easy and is left to the reader.

The following theorem constructs a common root of systems of composed polynomials if $d_{j,k_j} = 0$.

Theorem 5.1 *If $d_{j,k_j} = 0$, for some j, then $(x_1, \ldots, x_n, z_1, \ldots, z_n)$ with $x_j \neq 0$ and $z_j = 0$ is a non-trivial common root of the toric homogenizations of the composed polynomials $F \circ G$, where the homogenizations are carried out with respect to the generic degrees $m_1, \ldots, m_n$ and $k_1, \ldots, k_n$ of F and respectively G.*

Proof: Since $d_{j,k_j} = 0$ and g_j is homogenized with respect to its generic degree k_j, we have that $g_j^{\mathrm{h}}(1, 0) = 0$. Therefore and by Lemma 5.3, for all i we have that

$$f_i^{\mathrm{h}_{1,\ldots,n}}(y_1, \ldots, y_j, \ldots, y_n, z_1, \ldots, z_j, \ldots, z_n) = 0$$

with $y_j = 0 = g_j^{\mathrm{h}}(1, 0)$ and $z_j = 0$. Thus the tuple $(x_1, \ldots, x_n, z_1, \ldots, z_n)$ with $x_j \neq 0$ and $z_j = 0$ is a common root of the toric homogenizations of the composed polynomials $F \circ G$. By Lemma 5.2 this root is non-trivial. □

Note that Theorem 5.1 is shown independently from Theorem 2.5. Furthermore, in the n-degree case, Theorem 5.1 obviously implies Theorem 2.5 and thus provides an alternative strategy for proving Theorem 5.1. However, it seems it is not possible to apply this strategy for proving Theorem 5.1 in general because the technique of toric homogenization is only developed for *unmixed* polynomials, such as n-degree polynomials, in Cox et al. (1998).

6 Conclusion and Future Directions

This paper studied the Cayley-Dixon construction of resultants for multi-univariate composed polynomials. It gave a factorization of the Cayley-Dixon matrix induced by the structure of the composed polynomials and it showed how to efficiently extract the Dixon projection operator utilizing the factorization of the Cayley-Dixon matrix.

In a special case, when $g_i = x_i^k$, the composition problem in the context of Cayley-Dixon construction was analyzed in Kapur and Saxena (1997), where it was studied as support scaling. For this specialized case, the main result of that paper coincides with Theorem 2.2. Results presented here are thus strict generalizations.

A new resultant formula like in McKay and Wang (1989) has been derived for multi-univariate composition of n-degree systems.

This paper also highlighted a class of mixed or non-generic polynomial systems for which the resultant can be computed exactly because of given composition structures in the polynomial systems. It was shown that a composed system of mixed supports can be generated from an unmixed outermost system when univariate polynomials are substituted for distinct variables, thus establishing a class of mixed supports for which Dixon-Cayley construction yields resultants (without extraneous factors). This result about computing resultants of mixed systems without extraneous factors appears to be the first of its kind. Furthermore, it was shown that it is possible to compute resultants exactly for unmixed outer polynomial systems which can be extracted from a composed system by functional decomposition. Such an approach for identifying polynomial systems for which resultants can be computed exactly is novel and seems promising.

Future research includes generalizing the results of the current paper to arbitrary (multivariate and mixed) composing polynomials $g_1, \ldots, g_l$, where l not necessarily equals the number of variables x_i. For such general cases, a Dixon matrix factorization, as in the current paper, may not always exist. Therefore, it is interesting to study conditions under which such a factorization does exist. Furthermore, it is worth investigating if it is still possible to utilize the composition structure of the polynomials in order to efficiently extract their resultant from their Dixon matrix.

References

Bajaj, C., Garrity, T., Warren, J., Nov 1988. On the application of multi-equational resultants. Tech. Rep. CSD-TR-826, Dept. of Computer Science, Purdue University.

Buse, L., Elkadi, M., Mourrain, B., 2000. Generalized resultants over unirational algebraic varieties. J. Symbolic Computation 29, 515–526.

Busé, L., Elkadi, M., Mourrain, B., 2003. Using projection operators in computer aided geometric design. In: Topics in algebraic geometry and geometric modeling. Vol. 334 of Contemp. Math. Amer. Math. Soc., Providence, RI, pp. 321–342.

Cheng, C. C., McKay, J. H., Wang, S. S., April 1995. A chain rule for multivariable resultants. Proceedings of the American Mathematical Society 123 (4), 1037–1047.

Chionh, E., 1990. Base points, resultants, and the implicit representation of rational surfaces. PhD dissertation, University of Waterloo, Department of Computer Science.

Chionh, E.-W., 2001. Rectangular corner cutting and dixon $\mathcal{A}$-resultants. Journal of Symbolic Computation 31, 651–663.

Chtcherba, A., Kapur, D., 2003. Exact resultants for corner-cut unmixed multivariate polynomial systems using the Dixon formulation. Journal of Symbolic Computation 36 (3–4), 289–315.

Chtcherba, A. D., Aug 2003. A new Sylvester-type Resultant Method based on the Dixon-Bézout Formulation. PhD dissertation, University of New Mexico, Department of Computer Science.

Chtcherba, A. D., Kapur, D., 2000. Conditions for Exact Resultants using the Dixon formulation. ISSAC00, 62–70.

Coutsias, E. A., Seok, C., Jacobson, M. P., Dill, K. A., 2004. A kinematic view of loop closure. J Comput Chem 25, 510–528.

Cox, D., Little, J., O'Shea, D., 1998. Using Algebraic Geometry, 1st Edition. Springer-Verlag, New York.

Culver, T., Keyser, J., Manocha, D., 2004. Exact computation of the medial axis of a polyhedron. Comput. Aided Geom. Design 21 (1), 65–98.

Dixon, A., 1908. The eliminant of three quantics in two independent variables. Proc. London Mathematical Society 6, 468–478.

Emiris, I. Z., 2005. Toric resultants and applications to geometric modelling. In: Solving polynomial equations. Vol. 14 of Algorithms Comput. Math. Springer, Berlin, pp. 269–300.

Emiris, I. Z., Mourrain, B., 1999. Computer algebra methods for studying and computing molecular conformations. Algorithmica 25 (2-3), 372–402.

Erhart, E., 1967. Sur un problème de géométrie diophantienne linéaire. I. J. Reine Angew. Math. 226, 1–29.

Foo, M.-C., Chionh, E.-W., 2004. Corner edge cutting and dixon a-resultant quotients. Journal of Symbolic Computation 37, 101–119.

Hoffman, C., 1989. Geometric and Solid modeling. Morgan Kaufmann Publishers, Inc., San Mateo, California 94403.

Hong, H., 1997. Subresultants under composition. J. Symbolic Computation 23 (4), 355–365.

Hong, H., Minimair, M., 2002. Sparse resultant of composed polynomials I. J. Symbolic Computation 33, 447–465.

Jouanolou, J. P., 1991. Le formalisme du résultant. Adv. Math. 90 (2), 117–263.
Kapur, D., Lakshman, Y., 1992. Symbolic and Numeric Computation for Artificial Inteligence. Academic Press, Ch. Elimination Methods: an Introduction, donald, Kapur and Mundy (eds.).
Kapur, D., Saxena, T., July 1995. Comparison of various multivariate resultants. In: ACM ISSAC 95. Montreal, Canada.
Kapur, D., Saxena, T., 1997. Extraneous factors in the Dixon resultant formulation. In: ISSAC. Maui, Hawaii, USA, pp. 141–147.
Kapur, D., Saxena, T., Yang, L., July 1994. Algebraic and geometric reasoning using the Dixon resultants. In: ACM ISSAC 94. Oxford, England, pp. 99–107.
McKay, J. H., Wang, S. S., 1989. A chain rule for the resultant of two polynomials. Arch. Math. 53 (4), 347–351.
Minimair, M., 2001. Sparse resultants of composed polynomials. Ph.D. thesis, North Carolina State University, Raleigh, NC, USA.
Minimair, M., 2002. Sparse resultant of composed polynomials II. J. Symbolic Computation 33, 467–478.
Minimair, M., December 2003a. Dense resultant of composed polynomials. J. Symbolic Computation 36 (6), 825–834.
Minimair, M., 2003b. Factoring resultants of linearly combined polynomials. In: Sendra, J. R. (Ed.), Proceedings of the 2003 International Symposium on Symbolic and Algebraic Computation. ACM, New York, NY, pp. 207–214, iSSAC 2003, Philadelphia, PA, USA, August 3-6, 2003.
Minimair, M., July 2003c. Sparse resultant under vanishing coefficients. J. Algebraic Combinatorics 18 (1), 53–73.
Minimair, M., 2004. Computing resultants of partially composed polynomials. In: Ganzha, V. G., Mayr, E. W., Vorozhtsov, E. V. (Eds.), Computer Algebra in Scientific Computing. Proceedings of the CASC 2004 (St. Petersburg, Russia). TUM München, pp. 359–366.
Morgan, A., 1987. Solving polynomial systems using continuation for Scientific and Engineering problems. Prentice-Hall, Englewood Cliffs, NJ.
Ponce, J., Kriegman, D., 1992. Symbolic and Numerical Computation for Artificial Intelligence. Academic Press, Ch. Elimination Theory and Computer Vision: Recognition and Positioning of Curved 3D Objects from Range, donald, Kapur and Mundy (eds.).
Rojas, J. M., July/August 1999a. Solving degenerate sparse polynomial systems faster. J. Symbolic Computation 28 (1 and 2), 155–186, special Issue Polynomial Elimination – Algorithms and Applications.
Rojas, J. M., March 1999b. Toric intersection theory for affine root counting. J. Pure and Applied Algebra 136 (1), 67–100.
Rubio, R., 2000. Unirational Fields. Theorems, Algorithms and Applications. Ph.D. thesis, University of Cantabria, Santander, Spain.
Saxena, T., 1997. Efficient variable elimination using resultants. Ph.D. thesis, Depart-

ment of Computer Science, State Univeristy of New York, Albany, NY.
Sederberg, T., Goldman, R., 1986. Algebraic geometry for computer-aided design. IEEE Computer Graphics and Applications 6 (6), 52–59.
Zhang, M., May 2000. Topics in resultants and implicitization. Ph.D. thesis, Rice University, Dept. of Computer Science.
Zhang, M., Goldman, R., Aug. 2000. Rectangular corner cutting and sylvester $\mathcal{A}$-resultants. In: Traverso, C. (Ed.), Proc. of the ISSAC. ACM Press, St. Andrews, Scotland, pp. 301–308.

Appendix

A Operations on Bilinear Forms

A multivariate polynomial in terms of two disjoint sets of variables, e.g., the Dixon polynomial in section 2.1, can be represented in a *bilinear form*. For analyzing how the functional composition of two polynomial systems affects the Dixon polynomials and Dixon matrices of the polynomial systems, bilinear form representations turn out to be useful. Below, we discuss various polynomial operations and their effect on bilinear forms.

A bilinear form of a polynomial p in two disjoint sets of variables is expressed as a matrix, post and pre-multiplied by monomial vectors. That is

$$p(x_1, \ldots, x_k, \overline{x}_1, \ldots, \overline{x}_l) = \sum_{\alpha,\beta} p_{\alpha,\beta}\, \overline{\mathbf{x}}^{\alpha} \mathbf{x}^{\beta} \;=\; \overline{X}_p{}^T \times M_p \times X_p,$$

where $\overline{X}_p$ and X_p are vectors with entries being monomials in terms of variables $\{\overline{x}_1, \ldots, \overline{x}_l\}$ and $\{x_1, \ldots, x_k\}$, respectively. M_p is a matrix with the coefficients $p_{\alpha,\beta}$ of terms in p as its entries.

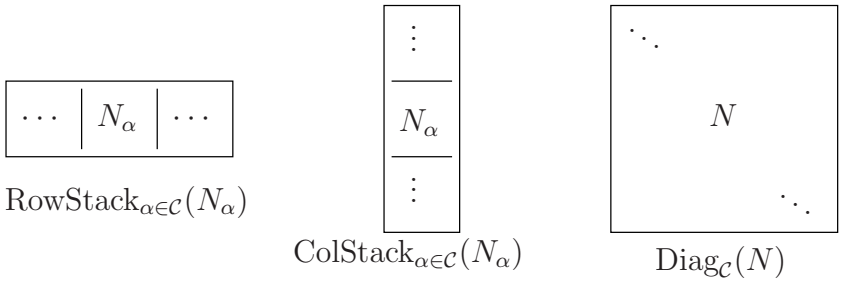
For example, let $p = \overline{xy}x^2 + 2\overline{xy}y - 3x^2$, then

$$p = \overline{X}_p{}^T \times M_p \times X_p = [\ \overline{xy}\ \ 1\ ] \times \begin{bmatrix} 1 & 2 \\ -3 & 0 \end{bmatrix} \times \begin{bmatrix} x^2 \\ y \end{bmatrix}. \tag{A.1}$$

The matrix M_p in the above definition depends on the monomial ordering used. We will assume a total degree ordering on power products, and state explicitly if it is otherwise. Also, implicit in the above definition of M_p are the row labels $\overline{X}_p$ and column labels X_p.

Let $\mathcal{P}$ be the ordered set of the exponent vectors corresponding to X_p; $\mathcal{P}$ is also called the **support** of the polynomial p w.r.t variables $\{x_1, \ldots, x_k\}$. Similarly, let $\overline{\mathcal{P}}$ be the support of p w.r.t. variables $\{\overline{x}_1, \ldots, \overline{x}_l\}$ ($\overline{\mathcal{P}}$ is also the ordered set of the exponent vectors corresponding to $\overline{X}_p$). For the above example, $\mathcal{P} = [(0,1),(2,0)]$ and $\overline{\mathcal{P}} = [(0,0),(1,1)]$.

Consider the following matrix construction operators, $\text{RowStack}_{\alpha\in\mathcal{C}}(N_\alpha)$, $\text{ColStack}_{\alpha\in\mathcal{C}}(N_\alpha)$ and $\text{Diag}_{\alpha\in\mathcal{C}}(N_\alpha)$, respectively, denoting the (block)-row and -column vector/matrix with its (block) indices taken from a support $\mathcal{C}$ and the block-diagonal matrix with as many blocks as elements in the support $\mathcal{C}$. See figure A.1.



$\text{RowStack}_{\alpha\in\mathcal{C}}(N_\alpha)$ $\quad$ $\text{ColStack}_{\alpha\in\mathcal{C}}(N_\alpha)$ $\quad$ $\text{Diag}_{\mathcal{C}}(N)$

Fig. A.1. Matrix constructors.

We will express bilinear polynomial product and composition in terms of the above operators. Consider two polynomials p and q in bilinear forms along with their associated power products. I.e.,

$$p = \overline{X_p}^T \times M_p \times X_p, \quad \text{and} \quad q = \overline{X_q}^T \times M_q \times X_q.$$

The respective supports of p in $\mathbf{x}$ and $\overline{\mathbf{x}}$ are $\mathcal{P}, \overline{\mathcal{P}}$; similarly, the respective supports of q are $\mathcal{Q}, \overline{\mathcal{Q}}$. Let $\mathcal{P} + \mathcal{Q}$ stand for the Minkowski sum of supports $\mathcal{P}$ and $\mathcal{Q}$. (As usual, the Minkowski sum $\mathcal{P} + \mathcal{Q}$ is the set of all sums $\alpha + \beta$ with $\alpha \in \mathcal{P}$ and $\beta \in \mathcal{Q}$.)

A.1 Polynomial Product in Terms of Bilinear Forms

We study the bilinear matrix form of the polynomial product pq of two polynomials p and q. We investigate if it is possible to express the bilinear form representation of pq in terms of the bilinear form representations of p and q. In other words, we investigate if it is possible to express the matrix M_{pq} in terms of the matrices M_p and M_q. This is indeed possible as shown in the following example and in Lemma A.1. Towards this end, we define auxiliary operators, so-called left and right multiplication operators, which are also illustrated in the following example. These operators will be useful in analyzing the Dixon Matrix, because the Dixon polynomial under composition is a

product of two polynomials.

EXAMPLE A.1 [Left and right multiplication operators] Let $p = a_1\,\overline{x}_1\overline{x}_2x_1^2 + a_2\,\overline{x}_1\overline{x}_2x_2 + a_3\,x_1^2$ and $q = b_1\,\overline{x}_1\overline{x}_2x_1^3x_2 + b_2\,\overline{x}_1x_1^3x_2 + b_3x_1^3x_2$, then

$$p = \begin{pmatrix} \overline{x}_1\overline{x}_2 \\ 1 \end{pmatrix}^T \times \begin{pmatrix} a_1 & a_2 \\ a_3 & 0 \end{pmatrix} \times \begin{pmatrix} x_1^2 \\ x_2 \end{pmatrix}, \qquad q = \begin{pmatrix} \overline{x}_1\overline{x}_2 \\ \overline{x}_1 \\ 1 \end{pmatrix}^T \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \times \begin{pmatrix} x_1^3x_2 \end{pmatrix},$$

and

$$\begin{aligned} pq = {} & a_1b_1\overline{x}_1^2\overline{x}_2^2x_1^5x_2 + a_1b_2\overline{x}_1^2\overline{x}_2x_1^5x_2 + a_1b_3\overline{x}_1\overline{x}_2x_1^5x_2 + a_2b_1\overline{x}_1^2\overline{x}_2^2x_1^3x_2^2 \\ & + a_2b_2\overline{x}_1^2\overline{x}_2x_1^3x_2^2 + a_2b_3\overline{x}_1\overline{x}_2x_1^3x_2^2 + a_3b_1\overline{x}_1\overline{x}_2x_1^5x_2 + a_3b_2\overline{x}_1x_1^5x_2 + a_3b_3x_1^5x_2 = \\ & \begin{pmatrix} \overline{x}_1^2\overline{x}_2^2 \\ \overline{x}_1^2\overline{x}_2 \\ \overline{x}_1\overline{x}_2 \\ \overline{x}_1 \\ 1 \end{pmatrix}^T \times \begin{pmatrix} a_1b_1 & a_2b_1 \\ a_1b_2 & a_2b_2 \\ a_1b_3 + a_3b_1 & a_2b_3 \\ a_3b_2 & 0 \\ a_3b_3 & 0 \end{pmatrix} \times \begin{pmatrix} x_1^5x_2 \\ x_1^3x_2^2 \end{pmatrix} = \overline{X}_{pq} \times M_{pq} \times X_{pq}. \end{aligned}$$

The bilinear supports of the polynomials p and q are $\overline{\mathcal{P}} = [(1,1),(0,0)]$, $\mathcal{P} = [(2,0),(0,1)]$, $\overline{\mathcal{Q}} = [(1,1),(1,0),(0,0)]$ and $\mathcal{Q} = [(3,1)]$. Then the left multiplication operator $\mathrm{L}_{\overline{\mathcal{Q}}}(M_p)$ on M_p is implicitly defined by the equality

$$p' = p \cdot \sum_{\overline{\epsilon}_q \in \overline{\mathcal{Q}}} \overline{\mathbf{x}}^{\overline{\epsilon}_q}\mathbf{z}^{\overline{\epsilon}_q} = \begin{pmatrix} \overline{x}_1^2\overline{x}_2^2 \\ \overline{x}_1^2\overline{x}_2 \\ \overline{x}_1\overline{x}_2 \\ \overline{x}_1 \\ 1 \end{pmatrix}^T \times \underbrace{\left(\begin{array}{cc|cc|cc} a_1 & a_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & a_1 & a_2 & 0 & 0 \\ a_3 & 0 & 0 & 0 & a_1 & a_2 \\ 0 & 0 & a_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & a_3 & 0 \end{array}\right)}_{\mathrm{L}_{\overline{\mathcal{Q}}}(M_p)} \times \begin{pmatrix} z_1z_2\,x_1^2 \\ z_1z_2\,x_2 \\ z_1\,x_1^2 \\ z_1\,x_2 \\ x_1^2 \\ x_2 \end{pmatrix},$$

and similarly the right multiplication operator $\mathrm{R}_{\mathcal{P}}(M_q)$ on M_q is implicitly defined by

the equality

$$q' = q \cdot \sum_{\epsilon_p \in \mathcal{P}} \overline{\mathbf{z}}^{\epsilon_p} \overline{\mathbf{x}}^{\epsilon_p} = \begin{pmatrix} \overline{x}_1\overline{x}_2\,\overline{z}_1^2 \\ \overline{x}_1\overline{x}_2\;\;\overline{z}_2 \\ \overline{x}_1\;\overline{z}_1^2 \\ \overline{x}_1\;\;\overline{z}_2 \\ 1\;\overline{z}_1^2 \\ 1\;\;\overline{z}_2 \end{pmatrix}^T \times \underbrace{\begin{pmatrix} b_1 & 0 \\ 0 & b_1 \\ b_2 & 0 \\ 0 & b_2 \\ b_3 & 0 \\ 0 & b_3 \end{pmatrix}}_{\mathrm{R}_{\mathcal{P}}(M_q)} \times \begin{pmatrix} x_1^5 x_2 \\ x_1^3 x_2^2 \end{pmatrix}.$$

Using the above operators, we can express the bilinear form of a polynomial product as matrix multiplication, as shown in Figure A.2. □

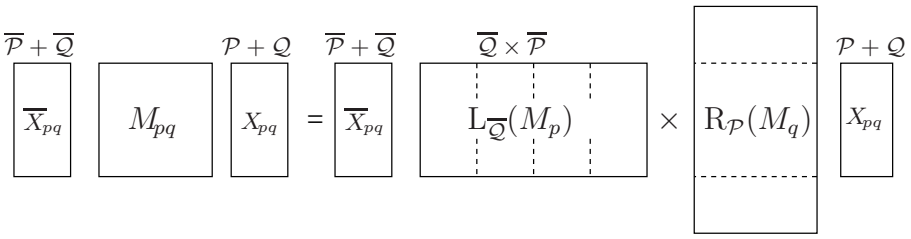


Fig. A.2. Left and right multiplication operators.

Now we formalize what we observed in the preceding example.

Definition A.1 *Given two polynomials p and q admitting bilinear form, consider the following polynomial products*

$$p' = p \cdot \sum_{\overline{\epsilon}_q \in \overline{\mathcal{Q}}} \overline{\mathbf{x}}^{\overline{\epsilon}_q} \mathbf{z}^{\overline{\epsilon}_q} = \overline{X}_{p'} \times M_{p'} \times X_{p'}, \qquad \textit{and}$$
$$q' = q \cdot \sum_{\epsilon_p \in \mathcal{P}} \overline{\mathbf{z}}^{\epsilon_p} \mathbf{x}^{\epsilon_p} = \overline{X}_{q'} \times M_{q'} \times X_{q'},$$

where $\overline{z}_1, \ldots, \overline{z}_n$ and $z_1, \ldots, z_n$ are new variables. These two equalities implicitly define two matrix operators

$$\mathrm{L}_{\overline{\mathcal{Q}}}(M_p) = M_{p'} \qquad \textit{and} \qquad \mathrm{R}_{\mathcal{P}}(M_q) = M_{q'}.$$

The columns of $M_{p'}$ and $M_{q'}$ are ordered according to the ordering of the corresponding monomials in $X_{p'}$ and respectively $X_{q'}$. Likewise, the rows of $M_{p'}$ and $M_{q'}$ are ordered according to the ordering of the corresponding monomials in $\overline{X}_{p'}$ and respectively $\overline{X}_{q'}$. We require that the monomials in $X_{p'}$, $X_{q'}$, $\overline{X}_{p'}$ and $\overline{X}_{q'}$ are ordered with some fixed

monomial orders. To make the matrices $M_{p'}$ and $M_{q'}$ "compatible" with each other, we require that the monomial order used for $X_{p'}$ be the same as for $\overline{X}_{q'}$ after replacing the symbols z, x with the respective symbols $\overline{z}$, $\overline{x}$. Furthermore, the order of $X_{p'}$ and $\overline{X}_{q'}$ is such that $z_k > x_l$ and respectively $\overline{z}_k > \overline{x}_l$ for all indices k, l.

The above matrix operators are defined in such a way that multiplying the left multiplication operator (represented by the matrix $\mathrm{L}_{\overline{\mathcal{Q}}}(M_p)$) and the right multiplication operator (represented by the matrix $\mathrm{R}_{\mathcal{P}}(M_q)$) yields the matrix (binomial form representation) of the polynomial product $p \cdot q$. In order to define the multiplication operators Definition A.1 uses some new distinct auxiliary variables $\overline{z}_1, \ldots, \overline{z}_n$ and $z_1, \ldots, z_n$. While the multiplication operators could be defined without these auxiliary variables, they allow to state the definition and the subsequent lemmas more succinctly and therefore they are employed.

It is important to point out that the row indices of $\mathrm{L}_{\overline{\mathcal{Q}}}(M_p)$ are $\overline{\mathcal{P}} + \overline{\mathcal{Q}}$ and column indices are $\overline{\mathcal{Q}} \times \mathcal{P}$, coming from the monomials $\mathbf{z}^{\overline{\epsilon}_q}\mathbf{x}^{\epsilon_p}$ for $\overline{\epsilon}_q \in \overline{\mathcal{Q}}$ and $\epsilon_p \in \mathcal{P}$. Similarly, the row indices of $\mathrm{R}_{\mathcal{P}}(M_q)$ are $\overline{\mathcal{Q}} \times \mathcal{P}$ (coming from the monomials $\overline{\mathbf{x}}^{\epsilon_q}\overline{\mathbf{z}}^{\epsilon_p}$) and the column indices are $\mathcal{P} + \mathcal{Q}$. Observe that the column indices of the left multiplication operator matrix equals the row indices of the right multiplication operator matrix. Thus it is really possible to multiply the two operator matrices. It is irrelevant how the columns and rows of the left and respectively right multiplication operators are ordered as long as they are ordered identically. Therefore we choose an arbitrary but fixed ordering.

The left and right multiplication operators can also be stated explicitly. In fact, the entry of $\mathrm{L}_{\overline{\mathcal{Q}}}(M_p)$ indexed by row $\overline{\mathbf{x}}^{\overline{\epsilon}_p+\overline{\epsilon}_q}$ and column $\mathbf{z}^{\overline{\epsilon}_q}\mathbf{x}^{\epsilon_p+\epsilon_q}$ is equal to $p_{\overline{\epsilon}_p,\epsilon_p}$. All other entries are 0. Thus, the matrix $\mathrm{L}_{\overline{\mathcal{Q}}}(M_p)$ is quite sparse and its entries are either 0 or the coefficients of p. Also it has block matrix structure

$$\mathrm{L}_{\overline{\mathcal{Q}}}(M_p) = \mathrm{RowStack}_{\alpha \in \overline{\mathcal{Q}}}(N_\alpha \times M_p)\,,$$

where N_α is a matrix which adds zero rows to M_p (depending on α, $\overline{\mathcal{Q}}$ and $\overline{\mathcal{P}}$). $\mathrm{R}_{\mathcal{P}}(M_q)$ also admits a similar block decomposition.

Lemma A.1

$$M_{p\,q} = \mathrm{L}_{\overline{\mathcal{Q}}}(M_p) \times \mathrm{R}_{\mathcal{P}}(M_q)\,.$$

Proof: Directly from the polynomial product of polynomials p and q,

$$(M_{p\,q})_{\alpha,\beta} = \sum_{\substack{\alpha=\overline{\epsilon}_p+\overline{\epsilon}_q,\\ \beta=\epsilon_p+\epsilon_q}} p_{\overline{\epsilon}_p,\epsilon_p} q_{\overline{\epsilon}_q,\epsilon_q},$$

for $\overline{\epsilon}_p \in \overline{\mathcal{P}}$, $\overline{\epsilon}_q \in \overline{\mathcal{Q}}$, $\epsilon_p \in \mathcal{P}$ and $\epsilon_q \in \mathcal{Q}$. On the other hand,

$$\begin{aligned}\left(\mathrm{L}_{\overline{\mathcal{Q}}}(M_p)\times \mathrm{R}_{\mathcal{P}}(M_q)\right)_{\alpha,\beta} &= \mathrm{Row}_\alpha\left(\mathrm{L}_{\overline{\mathcal{Q}}}(M_p)\right)\cdot \mathrm{Col}_\beta\left(\mathrm{R}_{\mathcal{P}}(M_q)\right)\\ &= \sum_{\substack{\overline{\epsilon}_q\in\overline{\mathcal{Q}},\\ \epsilon_p\in\mathcal{P}}} \mathrm{coeff}_{\overline{\mathbf{x}}^\alpha \mathbf{z}^{\overline{\epsilon}_q}\mathbf{x}^{\epsilon_p}}(p')\cdot \mathrm{coeff}_{\mathbf{x}^\beta\overline{\mathbf{z}}^{\epsilon_p}\overline{\mathbf{x}}^{\overline{\epsilon}_q}}(q'),\end{aligned}$$

but

$$\mathrm{coeff}_{\overline{\mathbf{x}}^\alpha \mathbf{z}^{\overline{\epsilon}_q}\mathbf{x}^{\epsilon_p}}(p') = \begin{cases} p_{\overline{\epsilon}_p,\epsilon_p} & \text{if } \alpha=\overline{\epsilon}_p+\overline{\epsilon}_q\\ 0 & \text{otherwise}\end{cases},$$

and

$$\mathrm{coeff}_{\mathbf{x}^\beta\overline{\mathbf{z}}^{\epsilon_p}\overline{\mathbf{x}}^{\overline{\epsilon}_q}}(q') = \begin{cases} q_{\overline{\epsilon}_q,\epsilon_q} & \text{if } \beta=\epsilon_p+\epsilon_q\\ 0 & \text{otherwise}\end{cases}.$$

Therefore

$$\begin{aligned}&\sum_{\substack{\overline{\epsilon}_q\in\overline{\mathcal{Q}},\\ \epsilon_p\in\mathcal{P}}} \mathrm{coeff}_{\overline{\mathbf{x}}^\alpha \mathbf{z}^{\overline{\epsilon}_q}\mathbf{x}^{\epsilon_p}}(p')\cdot \mathrm{coeff}_{\mathbf{x}^\beta\overline{\mathbf{z}}^{\epsilon_p}\overline{\mathbf{x}}^{\overline{\epsilon}_q}}(q')\\ &\qquad = \sum_{\substack{\alpha=\overline{\epsilon}_p+\overline{\epsilon}_q,\\ \beta=\epsilon_p+\epsilon_q}} \mathrm{coeff}_{\overline{\mathbf{x}}^\alpha \mathbf{z}^{\overline{\epsilon}_q}\mathbf{x}^{\epsilon_p}}(p')\cdot \mathrm{coeff}_{\mathbf{x}^\beta\overline{\mathbf{z}}^{\epsilon_p}\overline{\mathbf{x}}^{\overline{\epsilon}_q}}(q') = \sum_{\substack{\alpha=\overline{\epsilon}_p+\overline{\epsilon}_q,\\ \beta=\epsilon_p+\epsilon_q}} p_{\overline{\epsilon}_p,\epsilon_p} q_{\overline{\epsilon}_q,\epsilon_q},\end{aligned}$$

where exponents are chosen $\epsilon_p\in\mathcal{P}$, $\overline{\epsilon}_p\in\overline{\mathcal{P}}$, $q\in\overline{\mathcal{Q}}$ and $\overline{\epsilon}_q\in\overline{\mathcal{Q}}$. □

One of the useful properties of L operator is that the application on matrix product results in the application on one of the matrices times a block diagonal matrix of the other factor.

Lemma A.2 *Given a product of two matrices $A\times B$,*

$$\mathrm{L}_{\mathcal{P}}(A\times B) = \mathrm{L}_{\mathcal{P}}(A)\times \mathrm{Diag}_{\mathcal{P}}(B)\,,$$

where the product of the above matrices is assumed to conform to row and column labels as in Definition A.1.

The same holds for the operator R, but one has to take into account the difference in the row order.

Proof: By definition,

$$\begin{aligned}\mathrm{L}_{\mathcal{P}}(A\times B) &= \mathrm{RowStack}_{\alpha\in\mathcal{P}}(N_\alpha\times(A\times B)) = \mathrm{RowStack}_{\alpha\in\mathcal{P}}((N_\alpha\times A)\times B)\\ &= \mathrm{RowStack}_{\alpha\in\mathcal{P}}(N_\alpha\times A)\times \mathrm{Diag}_{\mathcal{P}}(B) = \mathrm{L}_{\mathcal{P}}(A)\times \mathrm{Diag}_{\mathcal{P}}(B)\,.\end{aligned}$$

□

The following is a simple but useful observation used in proving the main result.

Proposition A.1 *For polynomials* $p(\overline{x}_1, \ldots, \overline{x}_k, x_1, \ldots, x_l)$ *and* $q(\overline{y}_1, \ldots, \overline{y}_k, x_1, \ldots, x_l)$ *which are defined in terms of different sets of variables,*

$$\mathrm{L}_{\overline{\mathcal{Q}}}(M_p) = \mathrm{Diag}_{\overline{\mathcal{Q}}}(M_p)\,.$$

Proof: By definition

$$p' = p \cdot \sum_{\overline{\epsilon}_q \in \overline{\mathcal{Q}}} \mathbf{z}^{\overline{\epsilon}_q} \overline{\mathbf{y}}^{\overline{\epsilon}_q} = \overline{X}_{p'} \times \mathrm{L}_{\overline{\mathcal{Q}}}(M_p) \times X_{p'}.$$

Since the polynomial p does not have terms in variables $\overline{y}_1, \ldots, \overline{y}_k, z_1, \ldots, z_k$, the bilinear form of p, that is, matrix M_p is repeated $|\overline{\mathcal{Q}}|$ times along the diagonal in $\mathrm{L}_{\overline{\mathcal{Q}}}(M_p)$. □

A.1.1 Bilinear Form under Composition with Univariate Polynomials

To express the effect of substituting a univariate polynomial g_i in x_i for y_i in f_j, the following operator is needed. This operator is then used below to express how bilinear forms are affected by functional composition of two polynomial systems.

Definition A.2 *Given a support* $\mathcal{P}$ *and the set of univariate polynomials* $G = (g_1, \ldots, g_n)$*, where each* g_i *is in* x_i*, let*

$$s = \sum_{\alpha \in \mathcal{P}} \overline{\mathbf{x}}^{\alpha} G^{\alpha} = \overline{X}_s \times M_s \times X_s$$

where $G^{\alpha} = \prod_{i=1}^{n} g_i^{\alpha_i}$*. Define an operator* $\mathrm{S}_{\mathcal{P}}(G) = M_s$*.*

$\mathrm{S}_{\mathcal{P}}(G)$ is thus the matrix whose rows are indexed by $\mathcal{P}$ and whose columns are indexed by the union over $\alpha \in \mathcal{P}$ of the supports of $\prod_{j=1}^{n} g_j^{\alpha_j}$. Note that the monomial vector, with support $\mathcal{P}$ composed with G can be expressed as

$$Y_p \circ G = \mathrm{S}_{\mathcal{P}}(G) \times X_s,$$

where X_s is the union of all monomials in G^{α} for all $\alpha \in \mathcal{P}$ and Y_p is monomial vector with support $\mathcal{P}$. Matrix $\mathrm{S}_{\mathcal{P}}(G)$ is also very sparse. More specifically,

$$(\mathrm{S}_{\overline{\mathcal{P}}}(G))_{\overline{\epsilon}_s, \epsilon_s} = \begin{cases} (d_{1,k_1}, \ldots, d_{n,k_n})^{\overline{\epsilon}_s} & \text{if } (\epsilon_s)_i = k_i(\overline{\epsilon}_s)_i,\ \forall i, \\ 0 & \text{if } \exists\, i \text{ s.t. } k_i(\overline{\epsilon}_s)_i < (\epsilon_s)_i, \\ s_{\overline{\epsilon}_s, \epsilon_s} & \text{otherwise, i.e. if } \forall\, i,\ k_i(\overline{\epsilon}_s)_i > (\epsilon_s)_i. \end{cases} \tag{A.2}$$

In particular, in the univariate case if $\overline{\mathcal{P}} = [m-1, \ldots, 0]$, then the support of X_s is $[(m-1)k, \ldots, 0]$, and

$$\left(\mathrm{S}_{\overline{\mathcal{P}}}(G)\right)_{i,j} = \begin{cases} d_k^i & \text{if } j = k \cdot i, \\ 0 & \text{if } j > k \cdot i, \\ s_{i,j} & \text{otherwise,} \end{cases} \tag{A.3}$$

for $i \in [m-1, \ldots, 0]$ and $j \in [(m-1)k, \ldots, 0]$.

Next we illustrate the operator $\mathrm{S}_{\mathcal{P}}(G)$ in the bivariate setting.

Example A.2 [Operator $\mathrm{S}_{\mathcal{P}}(G)$] Let $\mathcal{P} = [(2,0),(1,1),(0,1),(0,0)]$, $g_1 = a_2x_1^2 + a_1x_1 + a_0$ and $g_2 = b_1x_2 + b_0$. Then

$$\mathrm{S}_{\mathcal{P}}(G) = \begin{pmatrix} g_1^2 \\ g_1g_2 \\ g_2 \\ 1 \end{pmatrix} = \begin{array}{c} \begin{matrix} \scriptstyle(4,0) & \scriptstyle(3,0) & \scriptstyle(2,1) & \scriptstyle(2,0) & \scriptstyle(1,1) & \scriptstyle(1,0) & \scriptstyle(0,1) & \scriptstyle(0,0) \end{matrix} \\ \begin{pmatrix} a_2^2 & 2a_2a_1 & 0 & 2a_2a_0 + a_1^2 & 0 & 2a_1a_0 & 0 & a_0^2 \\ 0 & 0 & a_2b_1 & 0 & a_1b_1 & a_1b_0 & a_0b_1 & a_0b_0 \\ 0 & 0 & 0 & 0 & 0 & 0 & b_1 & b_0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array},$$

where the top row shows the support of the X_s, i.e. the column labels of $\mathrm{S}_{\mathcal{P}}(G)$. □

The following lemma states how the bilinear form of a polynomial p in variables sets $\mathbf{y}$ and $\overline{\mathbf{y}}$ is affected when for each $i = 1, \ldots, n$, y_i and $\overline{y}_i$ are, respectively, substituted by g_i and $\overline{g}_i$, where g_i is a univariate polynomial in x_i and $\overline{g}_i$ is the univariate polynomial g_i in which x_i is uniformly replaced by $\overline{x}_i$. Also let $\overline{G} = (\overline{g}_1, \ldots, \overline{g}_n)$.

Lemma A.3 *Let p be a polynomial in the variables $\overline{\mathbf{y}}$, $\mathbf{y}$, and G a set of univariate polynomials g_i in variable x_i, for $i = 1, \ldots, n$. Then*

$$M_{p \circ (\overline{G}, G)} \;=\; \mathrm{S}_{\overline{\mathcal{P}}}\Big(\overline{G}\Big)^{\mathrm{T}} \;\times\; M_p \;\times\; \mathrm{S}_{\mathcal{P}}(G)\,,$$

where $\overline{G} = (\overline{g}_1, \ldots, \overline{g}_n)$, and $\overline{g}_i = g_i(\overline{x}_i)$.

Proof: Since $p = \overline{Y}_p \times M_p \times Y_p$, we have

$$p \circ (\overline{G},\, G) \;=\; (\overline{Y}_p^{\,T} \circ \overline{G}) \times M_p \times (Y_p \circ G)$$

and $Y_p \circ G = \mathrm{S}_{\mathcal{P}}(G) \times X_s$ by definition. □

A.1.2 Properties of Operators L, R and S

Consider the following slight generalization of triangular matrices to non-square matrices.

Definition A.3 *For a $k \times l$ matrix M, let t_i be the column index of first non-zero entry in row i. M is said to be (upper) step-triangular if $t_{i+1} > t_i$ for all $i = 1, \ldots, k-1$. The first non-zero entry in each row is called the diagonal entry, which make up the step diagonal of matrix M.*

Note that if matrix M without zero rows or columns, is square and step-triangular, then it is triangular. A matrix is lower step triangular if its transpose is upper step-triangular.

It is not hard to see that for any $\mathcal{P} \subset \mathbb{N}^d$, the matrix $\mathrm{S}_{\mathcal{P}}(G)$ is upper step triangular, by description of matrix entries in equation (A.2).

A useful property of operators L and S is that in combination, they produce step-triangular matrices for an n-degree support $\overline{\mathcal{Q}}$.

Proposition A.2 *For a polynomial $q = \prod_{i=1}^{n}(g_i - \overline{g}_i)/(x_i - \overline{x}_i)$, the bilinear form matrix M_q is (anti) triangular*[5] *of size $k \times k$, further, the anti-diagonal entries are $d_{1,k_1} \cdots d_{n,k_n}$.*

Proof: In the polynomial

$$\frac{g_i - \overline{g}_i}{x_i - \overline{x}_i} = \sum_{j=0}^{k_i} d_{i,j} \frac{x_i^j - \overline{x}_i^j}{x_i - \overline{x}_i} = \sum_{j=1}^{k_i} d_{i,j} \sum_{l=0}^{j-1} \overline{x}_i^l \, x_i^{j-l-1},$$

monomials $\overline{x}^j \, x^l$ for $j + l \geq k_i$ are not present; if $j + l = k_i - 1$, the coefficient of $\overline{x}_i^j \, x_i^l$ is d_{i,k_i}.

Since q is a product of such polynomials, which are defined in terms of different variables, we can characterize coefficients of q and hence the entries of M_q as

$$q_{\overline{\epsilon}_q, \epsilon_q} = \begin{cases} d_{1,k_1} \cdots d_{n,k_n} & \text{if } \overline{\epsilon}_q + \epsilon_q = k - 1, \\ 0 & \text{if } \exists \, i \text{ s.t. } (\overline{\epsilon}_q)_i + (\epsilon_q)_i > k_i - 1. \end{cases}$$

It can be seen that under lexicographical order on variables $x_1, \ldots, x_n$ and $\overline{x}_1, \ldots, \overline{x}_n$, the matrix M_q will be anti-triangular, i.e. 0 above the anti-diagonal. □

[5] The anti-diagonal of an $n \times n$ matrix consists of elements in the ith-row and $n - i - 1$ column of the matrix. A matrix is called anti-triangular if all entries below (or above) its anti-diagonal are zero.

The support $\overline{\mathcal{Q}}$ of q in the above proposition in terms of variables $\overline{x}_1, \ldots, \overline{x}_n$ is

$$\overline{\epsilon}_q \in \overline{\mathcal{Q}} \qquad \text{iff} \qquad 0 \leq (\overline{\epsilon}_q)_i < k_i \text{ for all } i = 1, \ldots, n.$$

Using the above properties, we can show that the operators L and S in the combination produce step triangular matrices, an important property used in derivation of the main result.

Proposition A.3 *Let $\overline{\mathcal{Q}}$ be a support of $\prod_{i=1}^{n} \frac{g_i - \overline{g}_i}{x_i - \overline{x}_i}$, then for any support $\mathcal{P}$, the matrix $\mathrm{L}_{\overline{\mathcal{Q}}}\Big(\mathrm{S}_{\mathcal{P}}\Big(\overline{G}\Big)^{\mathrm{T}}\Big)$ is (lower) step triangular (after column reordering); moreover, the entry in column $\overline{\epsilon}_q \epsilon_p$ and row α is:*

$$\mathrm{L}_{\overline{\mathcal{Q}}}\Big(\mathrm{S}_{\mathcal{P}}\Big(\overline{G}\Big)^{\mathrm{T}}\Big)_{\alpha, \overline{\epsilon}_q . \epsilon_p} = \begin{cases} (d_{1,k_1}, \ldots, d_{n,k_n})^{\epsilon_p} & \text{if } \alpha = \overline{\epsilon}_p + \overline{\epsilon}_q \text{ and } (\overline{\epsilon}_p)_i = k_i(\epsilon_p)_i, \\ \Big(\mathrm{S}_{\mathcal{P}}\Big(\overline{G}\Big)^{\mathrm{T}}\Big)_{\overline{\epsilon}_p, \epsilon_p} & \text{if } \alpha = \overline{\epsilon}_p + \overline{\epsilon}_q \text{ and } \forall i,\ (\overline{\epsilon}_p)_i < k_i(\epsilon_p)_i, \\ 0 & \text{otherwise} \end{cases}$$

i.e., in every column, the first non-zero entry is the product of the leading coefficients of G, and all these leading non-zero entries are in different rows.

Proof: The columns of $\mathrm{S}_{\mathcal{P}}\Big(\overline{G}\Big)^{\mathrm{T}}$ are labeled by $\mathcal{P}$ and the rows by $\overline{X}_s$, which is the set of all monomials in $\overline{G}^{\alpha} = \overline{g}_1^{\alpha_1} \cdots \overline{g}_n^{\alpha_n}$ for all $\alpha \in \mathcal{P}$. (Since we are considering the transpose $\mathrm{S}_{\mathcal{P}}\Big(\overline{G}\Big)^{\mathrm{T}}$, the $\overline{X}_s$ and X_s are switched as in the definition.)

Consider the following polynomial

$$s = \overline{X}_s \times \mathrm{S}_{\mathcal{P}}\Big(\overline{G}\Big)^{\mathrm{T}} \times X_s \qquad \text{and} \qquad s' = s \cdot \sum_{\overline{\epsilon}_q \in \overline{\mathcal{Q}}} \mathbf{z}^{\overline{\epsilon}_q}\, \overline{\mathbf{x}}^{\overline{\epsilon}_q},$$

as in definition A.1 of $\mathrm{L}_{\overline{\mathcal{Q}}}(M_p)$. We already know that

$$\mathrm{coeff}_{\overline{\mathbf{x}}^{\alpha} \mathbf{z}^{\overline{\epsilon}_q} \mathbf{x}^{e_s}}(s') = \begin{cases} s_{\overline{e}_s, e_s} & \text{if } \alpha = \overline{e}_s + \overline{\epsilon}_q, \\ 0 & \text{otherwise.} \end{cases} \tag{A.4}$$

Since the support of s is $\mathcal{P}$, we will use labels ϵ_p instead of e_s. Putting equations (A.2) and (A.4) together, we get

$$\mathrm{coeff}_{\overline{\mathbf{x}}^{\alpha} \mathbf{z}^{\overline{\epsilon}_q} \mathbf{x}^{\epsilon_p}}(s') = \begin{cases} (d_{1,k_1}, \ldots, d_{n,k_n})^{\epsilon_p} & \text{if } \alpha = \overline{\epsilon}_p + \overline{\epsilon}_q \text{ and } (\overline{\epsilon}_p)_i = k_i(\epsilon_p)_i, \\ s_{\overline{\epsilon}_p, \epsilon_p} & \text{if } \alpha = \overline{\epsilon}_p + \overline{\epsilon}_q \text{ and } \forall i,\ (\overline{\epsilon}_p)_i < k_i(\epsilon_p)_i, \\ 0 & \text{otherwise.} \end{cases}$$

Moreover, from the above we can see that there exist a monomial order on the columns of the matrix such that matrix is lower step-triangular. □

Next, we illustrate how the left multiplication operator interacts with the operator S.

Example A.3 [L and S interaction] Consider a bivariate support composed with the univariate system G from example A.2, where $\mathrm{S}_{\overline{\mathcal{P}}}\left(\overline{G}\right)$ is shown for $\overline{\mathcal{P}} = [(2,0),\ (1,1),\ (0,1),\ (0,0)]$. Let $\overline{\mathcal{Q}} = [(1,0),(0,0)]$, by the previous proposition. Then $\mathrm{L}_{\overline{\mathcal{Q}}}\left(\mathrm{S}_{\overline{\mathcal{P}}}\left(\overline{G}\right)^{\mathrm{T}}\right)$ is a matrix with rows and columns of the following table.

	$\overline{\mathcal{Q}}$	$(1,0)$				$(0,0)$			
	$\overline{\mathcal{P}}$	$(2,0)$	$(1,1)$	$(0,1)$	$(0,0)$	$(2,0)$	$(1,1)$	$(0,1)$	$(0,0)$
	$(5,0)$	a_2^2	0	0	0	0	0	0	0
	$(4,0)$	$2a_2a_1$	0	0	0	a_2^2	0	0	0
$\mathcal{P}$	$(3,1)$	0	a_2b_1	0	0	0	0	0	0
	$(3,0)$	$2a_2a_0 + a_1^2$	0	0	0	$2a_2a_1$	0	0	0
$+$	$(2,1)$	0	a_1b_1	0	0	0	a_2b_1	0	0
	$(2,0)$	$2a_1a_0$	a_1b_0	0	0	$2a_2a_0 + a_1^2$	0	0	0
$\overline{\mathcal{Q}}$	$(1,1)$	0	a_0b_1	b_1	0	0	a_1b_1	0	0
	$(1,0)$	a_0^2	a_0b_0	b_0	1	$2a_1a_0$	a_1b_0	0	0
	$(0,1)$	0	0	0	0	0	a_0b_1	b_1	0
	$(0,0)$	0	0	0	0	a_0^2	a_0b_0	b_0	1

It is important to point out that there is an order on the columns of the matrix, so that matrix is step-triangular. The columns of $\mathrm{L}_{\overline{\mathcal{Q}}}\left(\mathrm{S}_{\overline{\mathcal{P}}}\left(\overline{G}\right)^{\mathrm{T}}\right)$ are ordered by $\{z_1,\ldots,z_n\}$ and then by $\{x_1,\ldots,x_n\}$. The order which makes the above matrix step triangular is the lexicographical order for variables $[x_1,z_1,x_2,z_2,\ldots,x_n,z_n]$, as per Proposition A.3. □

In particular, in the univariate case, when $G=(g)$, g of degree k and $\mathcal{P}=[0,\ldots,m-1]$ then $\overline{\mathcal{Q}}=[0,\ldots,k-1]$, the row support of polynomial s is $[0,\ldots,k(m-1)]$, then the matrix $\mathrm{L}_{\overline{\mathcal{Q}}}\left(\mathrm{S}_{\mathcal{P}}\left(\overline{G}\right)^{\mathrm{T}}\right)$ has km rows labeled by $[0,\ldots,k(m-1)]+[0,\ldots,k-1]$ and km

columns labeled by $[0,\ldots,k-1]\times[0,\ldots,m-1]$. This matrix is square, and more over

$$\Big(\mathrm{L}_{\overline{\mathcal{Q}}}\Big(\mathrm{S}_{\mathcal{P}}\Big(\overline{G}\Big)^{\mathrm{T}}\Big)\Big)_{i,j.l}=\begin{cases}0 & \text{if } i<j,\\ d_k^l & \text{if } i-j=kl,\\ \mathrm{S}_{\mathcal{P}}\Big(\overline{G}\Big)^{\mathrm{T}}_{i-j,l} & \text{if } i-j<kl,\\ 0 & \text{if } i-j>kl,\end{cases} \tag{A.5}$$

for $i\in[0,\ldots,km-1]$, $j\in[0,\ldots,k-1]$ and $l\in[0,\ldots,m-1]$. It is easy to see that for fixed l, we get a lower triangular sub-matrix of size $k\times k$. In fact running indices in $(i,l.j)$ order will result in a triangular matrix, with diagonal entries d_k^l.

In the rest of the appendix, we use the above operators in expressing the manipulations of bilinear forms of various polynomials arising in the Cayley-Dixon construction to show that Dixon matrix of composed system can be decomposed as a matrix product.

Particularly, the next section considers the case when when the outer system F consists of two univariate polynomials in y and G consists of a single univariate polynomial in x.

A.2 Case Study: The Cayley-Dixon construction for a Univariate composed System

The purpose of this section is to illustrate the use of the operators L, R and S and to show in great detail how they can be used to derive a resultant formula for the composed polynomials $F\circ G$ in a special case. The special case considered in this section is when F consists of two univariate polynomials. As the reader will see, this derivation proceeds by relating the Dixon (Bézout) matrix of $F\circ G$ to the Dixon (Bézout) matrix of F.

Consider a general univariate polynomial system $F=(f_0,f_1)$, where

$$f_0=a_{m_0}y^{m_0}+\cdots+a_1y+a_0,\qquad\text{and}\qquad f_1=b_{m_1}y^{m_1}+\cdots+b_1y+b_0,$$

and let $m=\max(m_0,m_1)$. Let $G=(g)$, where

$$g=d_kx^k+d_{k-1}x^{k-1}+\cdots+d_2x^2+d_1x+d_0.$$

McKay and Wang (1989) showed that the resultant of the composed polynomials $F\circ G$, can be factored as follows:

$$\mathrm{Res}(f_0\circ g,\, f_1\circ g)=d_k^{m_0m_1k}\,\mathrm{Res}(f_0,\, f_1)^k. \tag{A.6}$$

Below, we derive the same formula using the matrix techniques introduced in the current paper. This derivation is much longer than the short proof by McKay and

Wang. However, as the reader will see in the next section, this derivation can naturally be generalized to study Dixon matrices for multivariate polynomials. It seems that the techniques used by McKay and Wang cannot be generalized for this purpose.

The **Beźout-Cayley Construction** for the composed univariate polynomials $f_0 \circ g$ and $f_1 \circ g$ is done as follows. Let $\overline{g}$ denote the polynomial obtained from g by replacing x with $\overline{x}$. We get the Bézout polynomial of the composed system

$$\begin{aligned}\theta\left(F \circ G\right) &= \frac{\det \begin{pmatrix} f_0 \circ g & f_1 \circ g \\ f_0 \circ \overline{g} & f_1 \circ \overline{g} \end{pmatrix}}{x - \overline{x}} = \frac{\det \begin{pmatrix} f_0 \circ g & f_1 \circ g \\ f_0 \circ \overline{g} & f_1 \circ \overline{g} \end{pmatrix}}{g - \overline{g}} \cdot \frac{g - \overline{g}}{x - \overline{x}} \\ &= \left(\theta(F) \circ (\overline{g},\, g)\right) \cdot \frac{g - \overline{g}}{x - \overline{x}}.\end{aligned}$$

By Lemma A.1, which factors the bilinear form of a product of two polynomials, we have

$$\theta(F \circ G) \;=\; \mathrm{L}_{\overline{\mathcal{Q}}}(M_p) \times \mathrm{R}_{\mathcal{P}}(M_q)\,, \tag{A.7}$$

where $p = \theta(F) \circ (\overline{g}, g)$ and $q = (g - \overline{g})/(x - \overline{x})$. Moreover the support of p with respect to the variable x is $\mathcal{P} = \{0, \ldots, (m-1)k\}$ and the support of q with respect to the variable $\overline{x}$ is $\overline{\mathcal{Q}} = \{0, \ldots, k-1\}$.

By Lemma A.3, which factors the bilinear form of a composed polynomial $p = \theta(F) \circ (\overline{g}, g)$,

$$M_p \;=\; \mathrm{S}_{\overline{\Delta}_F}(\overline{g})^{\mathrm{T}} \times \Theta_F \times \mathrm{S}_{\Delta_F}(g)\,.$$

By Lemma A.2, which relates the operator L applied to the product of matrices, we can decompose the left side of matrix product in (A.7) as

$$\mathrm{L}_{\overline{\mathcal{Q}}}(M_p) \;=\; \mathrm{L}_{\overline{\mathcal{Q}}}\Big(\mathrm{S}_{\overline{\Delta}_F}(\overline{g})^{\mathrm{T}}\Big) \times \mathrm{Diag}_{\overline{\mathcal{Q}}}(\Theta_F) \times \mathrm{Diag}_{\overline{\mathcal{Q}}}(\mathrm{S}_{\Delta_F}(g))\,.$$

Therefore,

$$\Theta_{F\circ G} = \mathrm{L}_{\overline{\mathcal{Q}}}\Big(\mathrm{S}_{\overline{\Delta}_F}(g)^{\mathrm{T}}\Big) \;\times\; \mathrm{Diag}_{\overline{\mathcal{Q}}}(\Theta_F) \times \Big(\mathrm{Diag}_{\overline{\mathcal{Q}}}(\mathrm{S}_{\Delta_F}(g)) \times \mathrm{R}_{\mathcal{P}}(M_q)\Big)\,.$$

This factorization can also be found in Figure A.3.

Next we compute the determinant of the Bézout matrix $\Theta_{F\circ G}$ of the composed polynomials. Notice that the above factors

$$\mathrm{Diag}_{\overline{\mathcal{Q}}}(\Theta_F)\,, \quad \mathrm{L}_{\overline{\mathcal{Q}}}\Big(\mathrm{S}_{\overline{\Delta}_F}(G)^{\mathrm{T}}\Big) \qquad \text{and} \qquad \mathrm{Diag}_{\overline{\mathcal{Q}}}(\mathrm{S}_{\Delta_F}(G)) \times \mathrm{R}_{\mathcal{P}}(M_q)$$

are all square matrices of size mk.

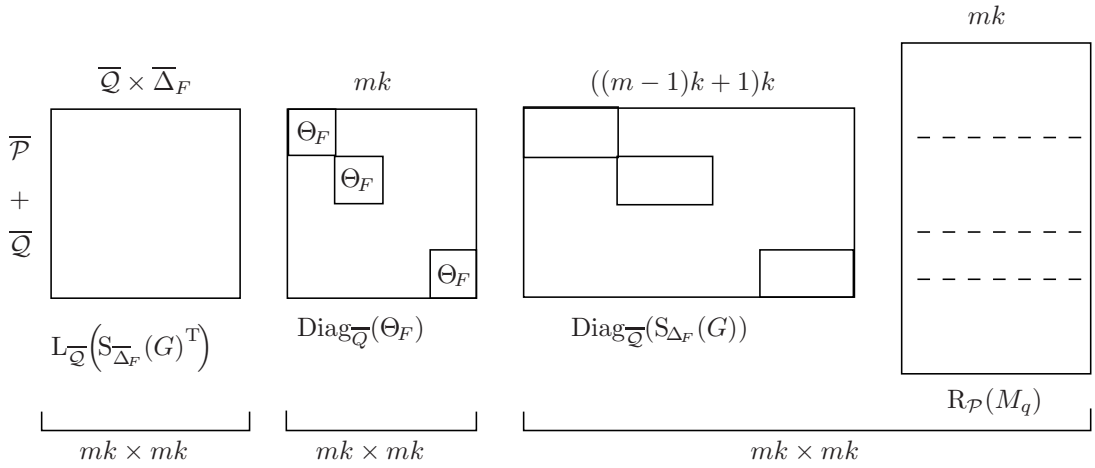


Fig. A.3. Decomposition of $\Theta_{F\circ G}$ in the univariate setting.

Observe also that $\Delta_F = \overline{\Delta}_F = \{0, \ldots, m-1\}$, that is, Θ_F is square of size m and since $\overline{\mathcal{Q}} = \{0, \ldots, k-1\}$, that is $|\overline{\mathcal{Q}}| = k$, the determinant of $\mathrm{Diag}_{\overline{\mathcal{Q}}}(\Theta_F)$ is $\det(\Theta_F)^k$.

By Proposition A.3, the columns of matrix $\mathrm{L}_{\overline{\mathcal{Q}}}\left(\mathrm{S}_{\overline{\Delta}_F}(G)^{\mathrm{T}}\right)$ can be permuted to make the matrix step triangular (see Definition A.3). Moreover, in the univariate case, it is square and hence triangular with entries d_k^i for $i \in \{0, \ldots, m-1\}$. Its determinant is (see equation (A.5))

$$\det\left[\mathrm{L}_{\overline{\mathcal{Q}}}\left(\mathrm{S}_{\overline{\Delta}_F}(G)^{\mathrm{T}}\right)\right] \;=\; \sum_{i=0}^{m-1} d_k^{k\,i} = (d_k)^{km(m-1)/2}.$$

Next consider the matrix $\mathrm{Diag}_{\overline{\mathcal{Q}}}(\mathrm{S}_{\Delta_F}(G)) \times \mathrm{R}_{\mathcal{P}}(M_q)$. Note that, by the proposition A.2, the matrix M_q is anti-triangular of size $k \times k$ with anti-diagonal entries d_k.

Proposition A.4 *The matrix* $\mathrm{Diag}_{\overline{\mathcal{Q}}}(\mathrm{S}_{\Delta_F}(G)) \times \mathrm{R}_{\mathcal{P}}(M_q)$ *is* triangular *with diagonal entries* d_k^{i+1} *in row labeled by* $i.l$ *for all* $i \in \Delta_F = \{0, \ldots, m-1\}$ *and* $l \in \overline{\mathcal{Q}} = \{0, \ldots, k-1\}$. *Therefore,*

$$\det\left[\mathrm{Diag}_{\overline{\mathcal{Q}}}(\mathrm{S}_{\Delta_F}(G)) \times \mathrm{R}_{\mathcal{P}}(M_q)\right] = \prod_{i=0}^{m-1}\prod_{l=0}^{k-1} d_k^{i+1} = d_k^{km(m+1)/2}.$$

Proof: Let $s = \overline{Z}_s \times \mathrm{S}_{\Delta_F}(G) \times X_s$, and set $M_s = \mathrm{S}_{\Delta_F}(G)$, where support of $\overline{Z}_s$ is Δ_F, in terms of new variable $\overline{z}$. Since

$$\mathrm{L}_{\overline{\mathcal{Q}}}(M_s) = \mathrm{Diag}_{\overline{\mathcal{Q}}}(M_s)$$

whenever Δ_F and $\overline{\mathcal{Q}}$ are supports in terms of different variables, by proposition A.1 we have

$$\mathrm{Diag}_{\overline{\mathcal{Q}}}(\mathrm{S}_{\Delta_F}(G)) \times \mathrm{R}_{\mathcal{P}}(M_q) \;=\; \mathrm{L}_{\overline{\mathcal{Q}}}(M_s) \times \mathrm{R}_{\mathcal{P}}(M_q) = M_{sq}.$$

The polynomial product

$$s \cdot q = \sum_{\substack{\overline{\epsilon}_s \in \Delta_F \\ \overline{\epsilon}_q \in \overline{\mathcal{Q}}}} \overline{z}^{\overline{\epsilon}_s}\, \overline{x}^{\overline{\epsilon}_q} \sum_{\beta=\epsilon_s+\epsilon_q} s_{\overline{\epsilon}_s,\epsilon_s}\, q_{\overline{\epsilon}_q,\epsilon_q}\; x^{\beta}.$$

Combining the descriptions of the coefficients of polynomials q and s, which are

$$q_{\overline{\epsilon}_q,\epsilon_q} = \begin{cases} d_k & \text{if } \overline{\epsilon}_q + \epsilon_q = k-1, \\ 0 & \text{if } \overline{\epsilon}_q + \epsilon_q > k-1, \end{cases} \qquad \text{by Proposition A.2,}$$

and also

$$s_{\overline{\epsilon}_s,\epsilon_s} = \begin{cases} d_k^{\overline{\epsilon}_s} & \text{if } \epsilon_s = k\overline{\epsilon}_s, \\ 0 & \text{if } \epsilon_s > k\overline{\epsilon}_s, \end{cases} \qquad \text{by equation (A.3),}$$

we get that

$$(s \cdot q)_{\overline{\epsilon}_s\overline{\epsilon}_q,\beta} = \sum_{\beta=\epsilon_s+\epsilon_q} s_{\overline{\epsilon}_s,\epsilon_s}\, q_{\overline{\epsilon}_q,\epsilon_q}\; x^{\beta} = \begin{cases} d_k^{\overline{\epsilon}_s+1} & \text{if } \epsilon_s = k\overline{\epsilon}_s \text{ and } \overline{\epsilon}_q + \epsilon_q = k-1, \\ 0 & \text{if } \epsilon_s > k\overline{\epsilon}_s \text{ or } \overline{\epsilon}_q + \epsilon_q > k-1, \end{cases}$$

where $\beta = \epsilon_s + \epsilon_q$.

Since $\overline{\epsilon}_s \in \Delta_F = [m-1, \ldots, 0]$, $\epsilon_s \in [(m-1)k, \ldots, 0]$, $\overline{\epsilon}_q \in [k-1, \ldots, 0]$ and $\epsilon_s \in [k-1, \ldots, 0]$, we can rewrite the above as

$$(s \cdot q)_{i.l,j} = \begin{cases} d_k^{i+1} & \text{if } j = ik + l, \\ 0 & \text{if } j < ik + l. \end{cases}$$

It is easy to see that M_{sq} is lower triangular matrix if the rows of are indexed by $[m-1, \ldots, 0] \times [k-1, \ldots, 0]$ and the columns indexed by $[km-1, \ldots, 0]$. $\square$

Hence

$$\begin{aligned} \det(\Theta_{F\circ G}) &= \det\left[\mathrm{L}_{\overline{\mathcal{Q}}}\left(\mathrm{S}_{\overline{\Delta}_F}(G)^{\mathrm{T}}\right)\right] \times \det\left[\mathrm{Diag}_k(\Theta_F)\right] \times \det\left[\mathrm{Diag}_{\overline{\mathcal{Q}}}(\mathrm{S}_{\Delta_F}(G)) \times \mathrm{R}_{\mathcal{P}}(M_q)\right] \\ &= (d_k)^{km(m-1)/2}\; (\det(\Theta_F))^k\; (d_k)^{km(m+1)/2} \\ &= (d_k)^{km^2} \det(\Theta_F)^k. \end{aligned}$$

It is well-known that in the case of $m_0 > m_1$, the determinant of the Bézout matrix constructed for F has $a_{m_0}^{(m_0-m_1)}$ as an extraneous factor. For the composed system $F \circ G$,

$$\begin{aligned} \det(\Theta_{F\circ G}) &= d_k^{km^2} a_{m_0}^{(m_0-m_1)k} \operatorname{Res}(f_0,\, f_1)^k \\ &= (d_k^{m_0} a_{m_0})^{k(m_0-m_1)} \cdot \operatorname{Res}(F \circ G). \end{aligned}$$

In this case, the extraneous factor is $(d_k^{m_0} a_{m_0})^{k(m_0-m_1)}$, which is the extraneous factor arising from F raised to the power $k(m_0-m_1)$ in addition to another extraneous factor which is a power of d_k, the leading coefficient of g.

Most of the above reasoning carries to general multivariate case with a few caveats. First, Dixon matrices are not guaranteed to be square or non-singular; thus their determinant cannot be computed or is 0. The technique introduced in Kapur et al. (1994) for extracting a multiple of the resultant from a matrix minor can be extended to these cases. Second, extra care is required to show that matrices (or their minors) are triangular so that the determinant can be computed and the resultant can be extracted. Moreover, the extraneous factors arising in the multivariate setting are more complex.

B Proofs

This section is a collection of main proofs of the paper, which are dependent on the material presented in Appendix A.

B.1 Proof of Lemma 2.1, page 11

PROOF: When A_L is square, it is triangular (up to column permutation) with diagonal entries

$$(A_L)_{\alpha,\overline{\epsilon}_q\epsilon_p} = (d_{1,k_1}, \ldots, d_{n,k_n})^{\epsilon_p}$$

in column $\overline{\epsilon}_q\epsilon_p$, where $\epsilon_p \in \mathcal{P} = \overline{\Delta}_F$, by Proposition A.3. Note that the size of $\overline{Q}$ is $k_1 \cdots k_n$. Therefore,

$$\det(A_L) = \prod_{\substack{\epsilon_p \in \overline{\Delta}_F \\ \overline{\epsilon}_q \in \overline{\mathcal{Q}}}} (d_{1,k_1}, \ldots, d_{n,k_n})^{\epsilon_p} = (d_{1,k_1}, \ldots, d_{n,k_n})^{\left(\sum_{\alpha \in \overline{\Delta}_F} \alpha\right) k_1 \cdots k_n}.$$

Also, for $A_R = \mathrm{Diag}_{\overline{\mathcal{Q}}}(\mathrm{S}_{\Delta_F}(G)) \times \mathrm{R}_{\mathcal{P}}(\mathrm{M}_q)$, let $s = \overline{Z}_s \times \mathrm{S}_{\Delta_F}(G) \times X_s$, $A_R = M_{sq}$, as in the univariate case. By Proposition A.2, M_q is triangular, where

$$q_{\overline{\epsilon}_q,\epsilon_q} = \begin{cases} d_{1,k_1} \cdots d_{n,k_n} & \text{if } \forall i \text{ s.t. } (\overline{\epsilon}_q)_i + (\epsilon_q)_i = k_i - 1, \\ 0 & \text{if } \exists i \text{ s.t. } (\overline{\epsilon}_q)_i + (\epsilon_q)_i > k_i - 1, \end{cases}$$

and entries of $\mathrm{S}_{\Delta_F}(G)$ by equation (A.2) are

$$s_{\overline{\epsilon}_s,\epsilon_s} = \begin{cases} (d_{1,k_1}, \ldots, d_{n,k_n})^{\overline{\epsilon}_s} & \text{if } \forall\, i \text{ s.t. } (\epsilon_s)_i = k_i(\overline{\epsilon}_s)_i, \\ 0 & \text{if } \exists\, i \text{ s.t. } k_i(\overline{\epsilon}_s)_i < (\epsilon_s)_i, \end{cases}$$

for $\overline{\epsilon}_s \in \Delta_F$ and ϵ_s in support of G^α for all $\alpha \in \Delta_F$. Therefore

$$(s \cdot q)_{\overline{\epsilon}_s \overline{\epsilon}_q, \epsilon_s + \epsilon_q} = \begin{cases} (d_{1,k_1}, \ldots, d_{n,k_n})^{\overline{\epsilon}_s + 1} & \text{if } \forall\, i \text{ s.t. } (\epsilon_s)_i = k_i (\overline{\epsilon}_s)_i \\ & \text{and } \overline{\epsilon}_q + \epsilon_q = k - 1, \\ 0 & \text{if } \exists\, i \text{ s.t. } (\epsilon_s)_i > k_i (\overline{\epsilon}_s)_i \\ & \text{or } (\overline{\epsilon}_q)_i + (\epsilon_q)_i > k_i - 1, \end{cases}$$

i.e., A_{R} is triangular, since the first non-zero entry in each row is in a different column.

In row $\overline{\epsilon}_s \overline{\epsilon}_q$, the diagonal element is $(d_{1,k_1}, \ldots, d_{n,k_n})^{\overline{\epsilon}_s + 1}$. Since $\overline{\epsilon}_s \in \Delta_F$ and $\overline{\epsilon}_q \in \overline{\mathcal{Q}}$, where $|\overline{\mathcal{Q}}| = k_1 \cdots k_n$, we have the determinant of A_{R}

$$\det(A_R) = \prod_{\substack{\overline{\epsilon}_s \in \Delta_F \\ \overline{\epsilon}_q \in \overline{\mathcal{Q}}}} (d_{1,\,k_1}, \ldots, d_{n,k_n})^{\overline{\epsilon}_s + 1} = (d_{1,\,k_1}, \ldots, d_{n,k_n})^{(|\Delta_F| + \sum_{\beta \in \Delta_F} \beta)\, k_1 \ldots k_n}.$$

□

B.2 Proof of Theorem 2.5, page 16

Before we prove the theorem, we state two auxiliary lemmas.

The first lemma shows that, under some mild technical assumption, the Newton polytope of a composed polynomial has a *vertex* with maximal, positive j-th coordinate for any index j and the coefficient of the term corresponding to this vertex is divisible by the leading coefficient of g_j. Such a vertex can be found among the points of the Newton polytope of the composed polynomial with maximal j-th coordinate. For example, if $f = 3\,y_1^2\,y_2^2 + 7\,y_1\,y_2^2 - 20\,y_2^2 - 2\,y_1 + 2$, $g_1 = 3\,x_1^2 - 4$ and $g_2 = 5\,x_2^3 + 2\,x_2^2 + 3\,x$, then, for $j = 2$, the part of the composed polynomial $f \circ (g_1, g_2)$ corresponding to maximal j-th coordinate is $3\,(3\,x_1^2 - 4)^2\,(5\,x^3)^2 + 7\,(3\,x_1^2 - 4)\,(5\,x^3)^2 - 20\,(5\,x^3)^2$. This part is obtained by composing the leading form, with respect to y_2, (the part of f with maximal power in y_2) with g_1 and the leading term of g_2. Furthermore, this part contains the monomials $x_1^4\,x_2^6$, $x_1^2\,x_2^6$, $x_1^2\,x_2^6$ and x_2^6. (For this example, one finds that the coefficient of x_2^6 is 0. Still, we consider it as belonging to the part because it would not vanish for polynomials f and g_j with arbitrary, generic coefficients.) For this example, one can easily verify that the monomials $x_1^4\,x_2^6$ and x_2^6 correspond to vertices in the Newton polytope of the composed polynomial $f \circ (g_1, g_2)$. The proof of Lemma B.1 shows that one can always find a monomial corresponding to a vertex. We also observe that the leading coefficient of g_2, $d_{j,k_j} = 5$, is a factor of the coefficients of the monomials $x_1^4\,x_2^6$ and x_2^6 (and of all other monomials in the part).

The second lemma shows that for fixed index j, we can uniformly find one normal vector for a system of composed polynomials selecting vertices whose corresponding terms have coefficients divisible by d_{j,k_j}. This observation is crucial for the application of Rojas' Vanishing Theorem, as we will see in the proof of Theorem 2.5. The existence of such a normal vector can be derived from the Minkowski sum of the composed polynomials' Newton polytopes. The normal vector is the normal vector of a vertex of the Minkowski sum with maximal j-th coordinate because this (as any) vertex of the Minkowski sum is the sum of vertices of the summands with the same normal vectors. Since the vertex has maximal j-th coordinate, the summands' vertices also have maximal j-th coordinate. By our previous observations, all these vertices correspond to monomial coefficients that are divisible by d_{j,k_j}.

Now we formalize our observations.

Lemma B.1 *For all i and j, the coefficient d_{j,k_j} is a factor of the coefficient C of a term $C \cdot x_1^{\epsilon_{i1}} \cdots x_j^{\overline{\epsilon_{ij}}} \cdots x_n^{\epsilon_{in}}$ in $f_i \circ G$ corresponding to a vertex of the Newton polytope of $f_i \circ G$ with maximal $\overline{\epsilon_{ij}} > 0$.*

PROOF: Let i and j be arbitrary but fixed. Moreover, fix a term of the form $C \cdot x_1^{\epsilon_{i1}} \cdots x_j^{\overline{\epsilon_{ij}}} \cdots x_n^{\epsilon_{in}}$ in $f_i \circ G$ corresponding to a vertex of the Newton polytope of $f_i \circ G$ with maximal $\overline{\epsilon_{ij}} > 0$. (Such a term exists because the Newton polytope is bounded.) We will see that d_{j,k_j} is a factor of the coefficient C by expanding the *composed* polynomial

$$\begin{aligned} f_i \circ G \;&=\; \sum_{\alpha \in \mathcal{F}_i} c_{i,\alpha} \cdot (g_1, \ldots, g_j, \ldots, g_n)^\alpha \\ &= \sum_{\alpha \in \mathcal{F}_i} c_{i,\alpha} \cdot g_1^{\alpha_1} \cdots \; (d_{j,k_j} x_j^{k_j} + \cdots + d_{j,0})^{\alpha_j} \; \cdots g_n^{\alpha_n}. \end{aligned}$$

Observe that $f_i \circ G$ is the sum of the terms $(d_{j,k_j} x_j^{k_j} + \cdots + d_{j,0})^{\alpha_j} \cdot c_{i,\alpha} \cdot (g_1^{\alpha_1} \cdots g_{j-1}^{\alpha_j} \cdot g_{j-1}^{\alpha_{j-1}} \cdots g_{j+1}^{\alpha_{j+1}})$. Since $\overline{\epsilon_{ij}}$ is maximal, the term $C \cdot x_1^{\epsilon_{i1}} \cdots x_j^{\overline{\epsilon_{ij}}} \cdots x_n^{\epsilon_{in}}$ can only occur in terms with maximal $\alpha_j = m_j$, the degree of f_i in y_j. Hence, $\overline{\epsilon_{ij}} = m_j\, k_j$ and thus only powers of d_{j,k_j} and of no other coefficients of g_j can be factors of C. $\square$

The next lemma shows that we can uniformly select vertices from Newton polytopes of a system of composed polynomials with maximal j-th coordinate.

Lemma B.2 *There is one inward normal vector for all composed polynomials $F \circ G$ that selects a vertex $(\epsilon_{i1}, \ldots, \overline{\epsilon_{ij}}, \ldots, \epsilon_{in})$ of the Newton polytopes of $f_i \circ (G)$, for $i = 1, \ldots, n$, with maximal $\overline{\epsilon_{ij}}$.*

PROOF: We observe that $\overline{\epsilon_{1j}} + \cdots + \overline{\epsilon_{nj}}$ is the j-th coordinate of a vertex with maximal j-th coordinate in the Minkowski sum $P_1 + \cdots + P_n$ of the Newton polytopes P_i of

$f_i \circ G$ because the Minkowski sum is bounded. Choose an inward normal vector ω that selects this vertex in the Minkowski sum. Note that any vertex of the Minkowski sum is the sum of vertices of the components P_i. Therefore (see Exercise 12b, p. 325, of Cox et al. (1998)) and by the maximality of the j-th coordinate, this inward normal vector ω selects for all i a vertex $(\epsilon_{i1}, \ldots, \overline{\epsilon_{ij}}, \ldots, \epsilon_{in})$ of the Newton polytopes of $f_i \circ (G)$ with maximal $\overline{\epsilon_{ij}}$. $\square$

Proof of the main Theorem 2.5, on page 16.

PROOF: By Lemma B.1, the vertices $(\epsilon_{i1}, \ldots, \overline{\epsilon_{ij}}, \ldots, \epsilon_{in})$ of the Newton polytopes of $f_i \circ G$ with maximal $\overline{\epsilon_{ij}}$'s correspond to monomials with coefficients of which d_{j,k_j} is a factor. By Lemma B.2, there is one inward normal vector for all composed polynomials selecting such a vertex. Then by Roja's Vanishing Theorem (Rojas, 1999a; Hong and Minimair, 2002), the vanishing of d_{j,k_j} implies that the toric resultant of the composed polynomials $F \circ G$ vanishes. Therefore, by Hilbert's Nullstellensatz, the coefficient d_{j,k_j} is a factor of the toric resultant of the composed polynomials. Moreover, by Rojas' Vanishing Theorem, the vanishing of d_{j,k_j} implies that the composed polynomials have a common zero at toric infinity (Rojas, 1999a,b; Hong and Minimair, 2002) and thus the factor d_{j,k_j} is not redundant. $\square$